On the Complexity of Bandit and Derivative-Free Stochastic Convex Optimization

Ohad Shamir   OHAD.SHAMIR@WEIZMANN.AC.IL

Microsoft Research and the Weizmann Institute of Science

Abstract

The problem of stochastic convex optimization with bandit feedback (in the learning community) or without knowledge of gradients (in the optimization community) has received much attention in recent years, in the form of algorithms and performance upper bounds. However, much less is known about the inherent complexity of these problems, and there are few lower bounds in the literature, especially for nonlinear functions. In this paper, we investigate the attainable error/regret in the bandit and derivative-free settings, as a function of the dimension d and the available number of queries T. We provide a precise characterization of the attainable performance for strongly-convex and smooth functions, which also imply a non-trivial lower bound for more general problems. Moreover, we prove that in both the bandit and derivative-free setting, the required number of queries must scale at least quadratically with the dimension. Finally, we show that on the natural class of quadratic functions, it is possible to obtain a "fast" $\mathcal{O}(1/T)$ error rate in terms of T, under mild assumptions, even without having access to gradients. To the best of our knowledge, this is the first such rate in a derivative-free stochastic setting, and holds despite previous results which seem to imply the contrary.

Keywords: Stochastic Convex Optimization; Derivative-Free Optimization; Bandit Convex Optimization; Regret

1. Introduction

This paper considers the following fundamental question: Given an unknown convex function F, and the ability to query for (possibly noisy) realizations of its values at various points, how can we optimize F with as few queries as possible?

This question, under different guises, has played an important role in several communities. In the optimization community, this is usually known as "zeroth-order" or "derivative-free" convex optimization, since we only have access to function values rather than gradients or higher-order information. The goal is to return a point with small optimization error on some convex domain, using a limited number of queries. Derivative-free methods were among the earliest algorithms to numerically solve unconstrained optimization problems, and have recently enjoyed increasing interest, being especial useful in black-box situations where gradient information is hard to compute or does not exist Nesterov (2011); Stich et al. (2011). In a stochastic framework, we can only obtain noisy realizations of the function values (for instance, due to running the optimization process on sampled data). We refer to this setting as *derivative-free SCO* (short for stochastic convex optimization).

In the learning community, these kinds of problems have been closely studied in the context of multi-armed bandits and (more generally) bandit online optimization, which are

powerful models for sequential decision making under uncertainty Cesa-Bianchi and Lugosi (2006); Bubeck and Cesa-Bianchi (2012). In a stochastic framework, these settings correspond to repeatedly choosing points in some convex domain, obtaining noisy realizations of some underlying convex function's value. However, rather than minimizing optimization error, our goal is to minimize the (average) regret: roughly speaking, that the average of the function values we obtain is not much larger than the minimal function value. For example, the well-known multi-armed bandit problem corresponds to a linear function over the simplex. We refer to this setting as *bandit SCO*. As will be more explicitly discussed later on, any algorithm which attains small average regret can be converted to an algorithm with the same optimization error. In other words, bandit SCO is only harder than derivative-free SCO. We note that in the context of stochastic multi-armed bandits, the potential gap between the two settings (under the terms "cumulative regret" and "simple regret") was introduced and studied in Bubeck et al. (2011).

When one is given gradient information, the attainable optimization error / average regret is well-known: under mild conditions, it is $\Theta(1/\sqrt{T})$ for convex functions and $\Theta(1/T)$ for strongly-convex functions, where T is the number of queries Zinkevich (2003); Hazan and Kale (2011); Rakhlin et al. (2012). Note that these bounds do not explicitly depend on the dimension of the domain.

The inherent complexity of bandit/derivative-free SCO is not as well-understood. An important exception is multi-armed bandits, where the attainable error/regret is known to be exactly $\Theta(\sqrt{d/T})$, where d is the dimension and T is the number of queries[1] Auer et al. (2002); Audibert and Bubeck (2009). Linear functions over other convex domains has also been explored, with upper bounds on the order of $\mathcal{O}(\sqrt{d/T})$ to $\mathcal{O}(\sqrt{d^2/T})$ (e.g. Abbasi-Yadkori et al. (2011); Bubeck et al. (2012)). For linear functions over general domains, information-theoretic $\Omega(\sqrt{d^2/T})$ lower bounds have been proven in Dani et al. (2007, 2008); Audibert et al. (2011). However, these lower bounds are either on the regret (not optimization error); shown for non-convex domains; or are implicit and rely on artificial, carefully constructed domains. In contrast, we focus here on simple, natural domains and convex problems.

When dealing with more general, non-linear functions, much less is known. The problem was originally considered over 30 years ago, in the seminal work by Yudin and Nemirovsky on the complexity of optimization Nemirovsky and Yudin (1983). The authors provided some algorithms and upper bounds, but as they themselves emphasize (cf. pg. 359), the attainable complexity is far from clear. Quite recently, Jamieson et al. (2012) provided an $\Omega(\sqrt{d/T})$ lower bound for strongly-convex functions, which demonstrates that the "fast" $\mathcal{O}(1/T)$ rate in terms of T, that one enjoys with gradient information, is not possible here. In contrast, the current best-known upper bounds are $\mathcal{O}(\sqrt[4]{d^2/T}), \mathcal{O}(\sqrt[3]{d^2/T}), \mathcal{O}(\sqrt{d^2/T})$ for convex, strongly-convex, and strongly-convex-and-smooth functions respectively (Flaxman et al. (2005); Agarwal et al. (2010)); And a $\mathcal{O}(\sqrt{d^{32}/T})$ bound for convex functions (Agarwal

---

1. In a stochastic setting, a more common bound in the literature is $\mathcal{O}(d\log(T)/T)$, but the $\mathcal{O}$-notation hides a non-trivial dependence on the form of the underlying linear function (in multi-armed bandits terminology, a gap between the expected rewards bounded away from 0). Such assumptions are not natural in a nonlinear bandits SCO setup, and without them, the regret is indeed $\Theta(\sqrt{d/T})$. See for instance (Bubeck and Cesa-Bianchi, 2012, Chapter 2) for more details.

et al. (2011)), which is better in terms of dependence on T but very bad in terms of the dimension d.

In this paper, we investigate the complexity of bandit and derivative-free stochastic convex optimization, focusing on nonlinear functions, with the following contributions (see also the summary in Table 1):

- We prove that for strongly-convex and smooth functions, the attainable error/regret is exactly $\Theta(\sqrt{d^2/T})$. This has three important ramifications: First of all, it settles the question of attainable performance for such functions, and is the first sharp characterization of complexity for a general nonlinear bandit/derivative-free class of problems. Second, it proves that the required number of queries T in such problems must scale quadratically with the dimension, even in the easier optimization setting, and in contrast to the linear case which often allows linear scaling with the dimension. Third, it formally provides a natural $\Omega(\sqrt{d^2/T})$ lower bound for more general classes of convex problems.

- We analyze an important special case of strongly-convex and smooth functions, namely quadratic functions. We show that for such functions, one can (efficiently) attain $\Theta(d^2/T)$ optimization error, and that this rate is sharp. To the best of our knowledge, it is the first general class of nonlinear functions for which one can show a "fast rate" (in terms of T) in a derivative-free stochastic setting. In fact, this may seem to contradict the result in Jamieson et al. (2012), which shows an $\Omega(\sqrt{d/T})$ lower bound on quadratic functions. However, as we explain in more detail later on, there is no contradiction, since the example establishing the lower bound of Jamieson et al. (2012) imposes an extremely small domain (which actually decays with T), while our result holds for a fixed domain. Although this result is tight, we also show that under more restrictive assumptions on the noise process, it is sometimes possible to obtain better error bounds, as good as $\mathcal{O}(d/T)$.

- We prove that even for quadratic functions, the attainable average *regret* is exactly $\Theta(\sqrt{d^2/T})$, in contrast to the $\Theta(d^2/T)$ result for optimization error. This shows there is a real gap between what can be obtained for derivative-free SCO and bandit SCO, without any specific distributional assumptions. Again, this stands in contrast to settings such as multi-armed bandits, where there is no difference in their distribution-free performance.

We emphasize that our upper bounds are based on the assumption that the function minimizer is bounded away from the domain boundary, or that we can query points slightly outside the domain. However, we argue that this assumption is not very restrictive in the context of strongly-convex functions (especially in learning applications), where the domain is often $\mathbb{R}^d$, and a minimizer always exists.

The paper is structured as follows: In Sec. 2, we formally define the setup and introduce the notation we shall use in the remainder of the paper. For clarity of exposition, we begin with the case of quadratic functions in Sec. 3, providing algorithms, upper and lower bounds. The tools and insights we develop for the quadratic case will allow us to tackle the more general strongly-convex-and-smooth setting in Sec. 4. We end the main part of the paper with a summary and discussion of open problems in Sec. 5. In Appendix A, we demonstrate

Function Type	Optimization Error		Average Regret	
	$\mathcal{O}(\cdot)$	$\Omega(\cdot)$	$\mathcal{O}(\cdot)$	$\Omega(\cdot)$
Quadratic	$\boxed{\dfrac{\mathbf{d^2}}{\mathbf{T}}}$		$\boxed{\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}}$	
Str. Convex and Smooth		$\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}$		
Str. Convex	$\min\left\{\sqrt[3]{\dfrac{d^2}{T}},\sqrt{\dfrac{d^{32}}{T}}\right\}$	$\boxed{\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}}$	$\min\left\{\sqrt[3]{\dfrac{d^2}{T}},\sqrt{\dfrac{d^{32}}{T}}\right\}$	$\boxed{\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}}$
Convex	$\min\left\{\sqrt[4]{\dfrac{d^2}{T}},\sqrt{\dfrac{d^{32}}{T}}\right\}$	$\boxed{\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}}$	$\min\left\{\sqrt[4]{\dfrac{d^2}{T}},\sqrt{\dfrac{d^{32}}{T}}\right\}$	$\boxed{\sqrt{\dfrac{\mathbf{d^2}}{\mathbf{T}}}}$

Table 1: A summary of the complexity upper bounds ($\mathcal{O}(\cdot)$) and lower bounds ($\Omega(\cdot)$), for derivative-free stochastic convex optimization (optimization error) and bandit stochastic convex optimization (average regret), for various function classes, in terms of the dimension d and the number of queries T. The boxed results are shown in this paper. The upper bounds for the convex and strongly convex case combine results from Flaxman et al. (2005); Agarwal et al. (2010, 2011). The table shows dependence on d, T only and ignores other factors and constants.

that one can obtain improved performance in the quadratic case, if we're considering more specific natural noise processes. Additional proofs are presented in Appendix B.

2. Preliminaries

Let $\|\cdot\|$ denote the standard Euclidean norm. We let $F(\cdot) : \mathcal{W} \mapsto \mathbb{R}$ denote the convex function of interest, where $\mathcal{W} \subseteq \mathbb{R}^d$ is a (closed) convex domain. We say that F is λ-strongly convex, for $\lambda > 0$, if for any $\mathbf{w}, \mathbf{w}' \in \mathcal{W}$ and any subgradient $\mathbf{g}$ of F at $\mathbf{w}$, it holds that $F(\mathbf{w}') \geq F(\mathbf{w}) + \langle \mathbf{g}, \mathbf{w}' - \mathbf{w}\rangle + \frac{\lambda}{2}\|\mathbf{w}' - \mathbf{w}\|^2$. Intuitively, this means that we can *lower bound* F everywhere by a quadratic function of fixed curvature. We say that F is μ-smooth if for any $\mathbf{w}, \mathbf{w}' \in \mathcal{W}$, and any subgradient $\mathbf{g}$ of F at $\mathbf{w}$, it holds that $F(\mathbf{w}') \leq F(\mathbf{w}) + \langle \mathbf{g}, \mathbf{w}' - \mathbf{w}\rangle + \frac{\mu}{2}\|\mathbf{w}' - \mathbf{w}\|^2$. Intuitively, this means that we can *upper-bound* F everywhere by a quadratic function of fixed curvature. We let $\mathbf{w}^* \in \mathcal{W}$ denote a minimizer of F on $\mathbf{w}$. To prevent trivialities, we consider in this paper only functions whose optimum $\mathbf{w}^*$ is known beforehand to lie in some bounded domain (even if $\mathcal{W}$ is large or all of $\mathbb{R}^d$), and the function is Lipschitz in that domain.

The learning/optimization process proceeds in T rounds. Each round t, we pick and query a point $\mathbf{w}_t \in \mathcal{W}$, obtaining an independent realization of $F(\mathbf{w}) + \xi_{\mathbf{w}}$, where $\xi_{\mathbf{w}}$ is an unknown zero-mean random variable, such[2] that $\mathbb{E}[\xi_{\mathbf{w}}^2] \leq \max\left\{1, \|\mathbf{w}\|^2\right\}$. In the bandit

---

2. We note that this slightly deviates from the more common assumption in the bandits/derivative-free SCO setting that $\mathbb{E}[\xi_{\mathbf{w}}^2] \leq \mathcal{O}(1)$. While such assumptions are equivalent for bounded $\mathcal{W}$, we also wish to consider cases with unrestricted domains $\mathcal{W} = \mathbb{R}^d$. In that case, assuming $\mathbb{E}[\xi_{\mathbf{w}}^2] \leq \mathcal{O}(1)$ may lead to trivialities in the derivative-free setting. For example, consider the case where $F(\mathbf{w}) = \mathbf{w}^\top A\mathbf{w} + \mathbf{b}^\top \mathbf{w}$. Then for any $\mathbf{w}$ and any $\xi_{\mathbf{w}}$ with uniformly bounded variance, we can get a virtually noiseless estimate

SCO setting, our goal is to minimize the *expected average regret*, namely

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}F(\mathbf{w}_t) - F(\mathbf{w}^*)\right],$$

whereas in the derivative-free SCO setting, our goal is to compute, based on $\mathbf{w}_1, \ldots, \mathbf{w}_T$ and the observed values, some point $\bar{\mathbf{w}}_T \in \mathcal{W}$, such that the *expected optimization error*

$$\mathbb{E}\left[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)\right],$$

is as small as possible. We note that given a bandit SCO algorithm with some regret bound, one can get a derivative-free SCO algorithm with the same optimization error bound: we simply run the stochastic bandit algorithm, getting $\mathbf{w}_1, \ldots, \mathbf{w}_T$, and returning $\frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$. By Jensen's inequality, the expected optimization error is at most the expected average regret with respect to $\mathbf{w}_1, \ldots, \mathbf{w}_T$. Thus, bandit SCO is only harder than derivative-free SCO.

In this paper, we provide upper and lower bounds on the attainable optimization error / average regret, as a function of the dimension d and the number of rounds/queries T. For simplicity, we focus here on bounds which hold in expectation, and an interesting point for further research is to extend these to bounds on the actual error/regret, which hold with high probability.

3. Quadratic Functions

In this section, we consider the class of quadratic functions, which have the form

$$F(\mathbf{w}) = \mathbf{w}^\top A \mathbf{w} + \mathbf{b}^\top \mathbf{w} + c$$

where A is positive-definite (with a minimal eigenvalue bounded away from 0). Moreover, to make the problem well-behaved, we assume that A has a spectral norm of at most 1, and that $\|\mathbf{b}\| \leq 1, |c| \leq 1$. We note that if the norms are bounded but larger than 1, this can be easily handled by rescaling the function. It is easily seen that such functions are both strongly convex and smooth. Moreover, this is a natural and important class of functions, which in learning applications appears, for instance, in the context of least squares and ridge regression. Besides providing new insights for this class, we will use the techniques developed here later on, in the more general case of strongly-convex and smooth functions.

3.1. Upper Bounds

We begin by showing that for derivative-free SCO, one can obtain an optimization error bound of $\mathcal{O}(d^2/T)$. To the best of our knowledge, this is the first example of a derivative-free stochastic bound scaling as $\mathcal{O}(1/T)$ for a general class of nonlinear functions, as opposed to $\mathcal{O}(1/\sqrt{T})$. However, to achieve this result, we need to make the following mild assumption:

---

of $\mathbf{w}^\top A \mathbf{w}$ by picking $\mathbf{w}' = c\mathbf{w}$ for some large c and computing $\frac{1}{c^2}\left(F(\mathbf{w}') + \xi_{\mathbf{w}'}\right)$. Variants of this idea will also allow virtually noiseless estimates of the linear term.

Assumption 1 *At least one of the following holds for some fixed $\epsilon \in (0, 1]$:*

- *The quadratic function attains its minimum $\mathbf{w}^*$ in the domain $\mathcal{W}$, and the Euclidean distance of $\mathbf{w}^*$ from the domain boundary is at least ϵ.*

- *We can query not just points in $\mathcal{W}$, but any point whose distance from $\mathcal{W}$ is at most ϵ.*

With strongly-convex functions, the most common case is that $\mathcal{W} = \mathbb{R}^d$, and then both cases actually hold for any value of ϵ. Even in other situations, one of these assumptions virtually always holds. Note that we crucially rely here on the strong-convexity assumption: with (say) linear functions, the domain must always be bounded and the optimum always lies at the boundary of the domain.

With this assumption, the bound we obtain is on the order of $d^2/\epsilon^2 T$. As discussed earlier, Jamieson et al. (2012) recently proved a $\Omega(\sqrt{d/T})$ lower bound for derivative-free SCO, which actually applies to quadratic functions. This does not contradict our result, since in their example the diameter of $\mathcal{W}$ (and hence also ϵ) decays with T. In contrast, our $\mathcal{O}(d^2/T)$ bound holds for fixed ϵ, which we believe is natural in most applications.

To obtain this behavior, we utilize a well-known 1-point gradient estimate technique, which allows us to get an unbiased estimate of the gradient at any point by randomly querying for a (noisy) value of the function around it (see Nemirovsky and Yudin (1983); Flaxman et al. (2005)). Our key insight is that whereas for general functions one must query very close to the point of interest (scaling to 0 with T), quadratic functions have additional structure which allows us to query relatively far away, allowing gradient estimates with much smaller variance.

The algorithm we use is presented as Algorithm 1, and is computationally efficient. It uses a modification $\bar{\mathcal{W}}$ of the domain $\mathcal{W}$, defined as follows. First, we let B denote some known upper bound on $\|\mathbf{w}^*\|$. If the first alternative of assumption 1 holds, then $\bar{\mathcal{W}}$ consists of all points in $\mathcal{W} \cap \{\mathbf{w} : \|\mathbf{w}\| \leq B\}$, whose distance from $\mathcal{W}$'s boundary is at least ϵ. If the second alternative holds, then $\bar{\mathcal{W}} = \mathcal{W} \cap \{\mathbf{w} : \|\mathbf{w}\| \leq B\}$. Note that under any alternative, it holds that $\bar{\mathcal{W}}$ is convex, that $\|\mathbf{w}_t\| \leq B$, that $\mathbf{w}^* \in \bar{\mathcal{W}}$, and that our algorithm always queries at legitimate points. In the pseudocode, we use $\Pi_{\bar{\mathcal{W}}}$ to denote projection on $\bar{\mathcal{W}}$. For simplicity, we assume that $T/2$ is an integer and that $\bar{\mathcal{W}}$ includes the origin $\mathbf{0}$.

Algorithm 1 Derivative-Free SCO Algorithm for Strongly-Convex Quadratic Functions

---

Input: Strong convexity parameter $\lambda > 0$; Distance parameter $\epsilon \in (0, 1]$
Initialize $\mathbf{w}_1 = \mathbf{0}$.
for $t = 1, \ldots, T - 1$ **do**
    Pick $\mathbf{r} \in \{-1, +1\}^d$ uniformly at random
    Query noisy function value v at point $\mathbf{w}_t + \frac{\epsilon}{\sqrt{d}}\mathbf{r}$
    Let $\tilde{\mathbf{g}} = \frac{\sqrt{d}v}{\epsilon}\mathbf{r}$
    Let $\mathbf{w}_{t+1} = \Pi_{\bar{\mathcal{W}}}\left(\mathbf{w}_t - \frac{1}{\lambda t}\tilde{\mathbf{g}}\right)$
end for
Return $\bar{\mathbf{w}}_T = \frac{2}{T}\sum_{t=T/2}^{T} \mathbf{w}_t$.

The following theorem quantifies the optimization error of our algorithm.

Theorem 1 *Let $F(\mathbf{w}) = \mathbf{w}^\top A\mathbf{w} + \mathbf{b}^\top \mathbf{w} + c$ be a λ-strongly convex function, where $\|A\|_2, \|\mathbf{b}\|, |c|$ are all at most 1, and suppose the optimum $\mathbf{w}^*$ has a norm of at most B. Then under Assumption 1, the point $\bar{\mathbf{w}}_T$ returned by Algorithm 1 satisfies*

$$\mathbb{E}\left[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)\right] \leq \frac{4(4 + 5\log(2))(B+1)^4}{\lambda \epsilon^2} \frac{d^2}{T}.$$

Note that returning $\bar{\mathbf{w}}_T$ as the average over the last $T/2$ iterates (as opposed to averaging over all iterates) is necessary to avoid $\log(T)$ factors Rakhlin et al. (2012).

As an interesting side-note, we conjecture that a gradient-based approach is crucial here to obtain $\mathcal{O}(1/T)$ rates (in terms of T). For example, a different family of derivative-free methods (see for instance Nemirovsky and Yudin (1983); Agarwal et al. (2011); Jamieson et al. (2012)) is based on a type of noisy binary search, where a few strategically selected points are repeatedly sampled in order to estimate which of them has a larger/smaller function value. This is used to shrink the feasible region where the optimum $\mathbf{w}^*$ might lie. Since it is generally impossible to estimate the mean of noisy function values at a rate better than $\mathcal{O}(1/\sqrt{T})$, it is not clear if one can get an optimization rate faster than $\mathcal{O}(1/\sqrt{T})$ with such methods.

The proof of the theorem relies on the following key lemma, whose proof appears in the appendix.

Lemma 2 *For any $\mathbf{w}_t$, we have that*

$$\mathbb{E}_{\mathbf{r},v}[\tilde{\mathbf{g}}] = \nabla F(\mathbf{w}_t)$$

and

$$\mathbb{E}_{\mathbf{r},v}[\|\tilde{\mathbf{g}}\|^2] \leq \frac{4d^2(B+1)^4}{\epsilon^2}.$$

This lemma implies that Algorithm 1 essentially performs stochastic gradient descent over the strongly-convex function $F(\mathbf{w})$, where the gradient estimates are unbiased and with bounded second moments. The returned point is a suffix-average of the last $T/2$ iterates. Using a convergence analysis for stochastic gradient descent with suffix-averaging (Rakhlin et al., 2012, Theorem 5), and plugging in the bounds of Lemma 2, we get Thm. 1.

3.2. Lower Bounds

In this subsection, we prove that the upper bound obtained in Thm. 1 is essentially tight: namely, up to constants, the worst-case error rate one can obtain for derivative-free SCO of quadratic functions is order of d^2/T. Besides showing that the algorithm above is essentially optimal, it implies that even for extremely nice strongly-convex functions and domains, the number of queries required to reach some fixed accuracy scales *quadratically* with the dimension d. This stands in contrast to the case of linear functions, where the provable query complexity often scales linearly with d.

Theorem 3 *Let the number of rounds T be fixed. Then for any (possibly randomized) querying strategy, there exists a quadratic function of the form $F(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 - \langle \mathbf{e}, \mathbf{w} \rangle$, which is minimized at $\mathbf{e}$ where $\|\mathbf{e}\| \leq 1$, such that the resulting $\bar{\mathbf{w}}_T$ satisfies*

$$\mathbb{E}[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)] \geq 0.01 \min\left\{1, \frac{d^2}{T}\right\}.$$

Note that since $\|\mathbf{e}\| \leq 1$, we know in advance that the optimum must lie in the unit Euclidean ball. Despite this, the lower bound holds even if we do not restrict at all the domain in which we are allowed to query - i.e., it can even be all of $\mathbb{R}^d$.

Proof The proof technique is inspired by a lower bound which appears in Arias-Castro et al. (2011), in the different context of compressed sensing. The argument also bears some close similarities to the proof of Assouad's lemma (see Cybakov (2009)).

We will exhibit a distribution over quadratic functions F, such that in expectation over this distribution, any querying strategy will attain $\Omega(d^2/T)$ optimization error. This implies that for any querying strategy, there exists some deterministic F for which it will have this amount of error.

The functions we shall consider are

$$F_{\mathbf{e}}(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 - \langle \mathbf{e}, \mathbf{w} \rangle,$$

where $\mathbf{e}$ is drawn uniformly from $\{-\mu, \mu\}^d$, with $\mu \in (0, 1/\sqrt{d})$ being a parameter to be specified later. Moreover, we will assume that the noise $\xi_{\mathbf{w}}$ is a Gaussian random variable with zero mean and standard deviation $\max\{1, \|\mathbf{w}\|^2\}$.

By definition of 1-strong convexity, it is easy to verify that $F_{\mathbf{e}}(\mathbf{w}) - F_{\mathbf{e}}(\mathbf{e}) \geq \frac{1}{2}\|\mathbf{w} - \mathbf{e}\|^2$. Thus, the expected optimization error (over the querying strategy) is at least

$$\mathbb{E}[F_{\mathbf{e}}(\bar{\mathbf{w}}_T) - F_{\mathbf{e}}(\mathbf{e})] \geq \mathbb{E}\left[\frac{1}{2}\|\bar{\mathbf{w}}_T - \mathbf{e}\|^2\right] \geq \mathbb{E}\left[\frac{1}{2}\sum_{i=1}^{d}(\bar{w}_i - e_i)^2\right] \geq \mathbb{E}\left[\frac{\mu^2}{2}\sum_{i=1}^{d}\mathbf{1}_{\bar{w}_i e_i < 0}\right]. \quad (1)$$

We will assume that the querying strategy is deterministic: $\mathbf{w}_t$ is a deterministic function of the previous query values $v_1, v_2, \ldots, v_{t-1}$ at $\mathbf{w}_1, \ldots, \mathbf{w}_{t-1}$. This assumption is without loss of generality, since any random querying strategy can be seen as a randomization over deterministic querying strategy. Thus, a lower bound which holds uniformly for any deterministic querying strategy would also hold over a randomization.

To lower bound Eq. (1), we use the following key lemma, which relates this to the question of how informative are the query values (as measured by Kullback-Leibler or KL divergence) for determining the sign of $\mathbf{e}$'s coordinates. Intuitively, the more similar the query values are, the smaller is the KL divergence and the harder it is to distinguish the true sign of each e_i, leading to a larger lower bound. The proof appears in the appendix.

Lemma 4 *Let $\mathbf{e}$ be a random vector, none of whose coordinates is supported on 0, and let $v_1, v_2, \ldots, v_T$ be a sequence of query values obtained by a deterministic strategy returning a point $\bar{\mathbf{w}}_T$ (so that the query location $\mathbf{w}_t$ is a deterministic function of $v_1, \ldots, v_{t-1}$, and $\bar{\mathbf{w}}_T$ is a deterministic function of $v_1, \ldots, v_T$). Then we have*

$$\mathbb{E}\left[\sum_{i=1}^{d}\mathbf{1}_{\bar{w}_i e_i < 0}\right] \geq \frac{d}{2}\left(1 - \sqrt{\frac{1}{d}\sum_{i=1}^{d}\sum_{t=1}^{T}U_{t,i}}\right),$$

where

$$U_{t,i} = \sup_{\{e_j\}_{j \neq i}} D_{kl}\left(\Pr\left(v_t | e_i > 0, \{e_j\}_{j \neq i}, \{v_l\}_{l=1}^{t-1}\right) \,\|\, \Pr\left(v_t | e_i < 0, \{e_j\}_{j \neq i}, \{v_l\}_{l=1}^{t-1}\right)\right)$$

and D_{kl} represents the KL divergence between two distributions.

Using Lemma 4, we can get a lower bound for the above, provided an upper bound on the $U_{t,i}$'s. To analyze this, consider any fixed values of $\{e_j\}_{j\neq i}$, and any fixed values of $v_1, \ldots, v_{t-1}$. Since the querying strategy is assumed to be deterministic, it follows that $\mathbf{w}_t$ is uniquely determined. Given this $\mathbf{w}_t$, the function value v_t equals

$$F_{\mathbf{e}}(\mathbf{w}_t) = \left(\frac{1}{2}\|\mathbf{w}_t\|^2 + \sum_{j\neq i} e_j w_{t,j} \right) + \mu w_{t,i} + \xi_{\mathbf{w}_t} \tag{2}$$

conditioned on $e_i > 0$, and

$$F_{\mathbf{e}}(\mathbf{w}_t) = \left(\frac{1}{2}\|\mathbf{w}_t\|^2 + \sum_{j\neq i} e_j w_{t,j} \right) - \mu w_{t,i} + \xi_{\mathbf{w}_t} \tag{3}$$

conditioned on $e_i < 0$. Comparing Eq. (2) and Eq. (3), we notice that they both represent a Gaussian distribution (due to the $\xi_{\mathbf{w}_t}$ noise term), with standard deviation $\max\left\{1, \|\mathbf{w}_t\|^2\right\}$ and means seperated by $2\mu w_{t,i}$. To bound the divergence, we use the following standard result on the KL divergence between two Gaussians Kullback (1959):

Lemma 5 *Let $\mathcal{N}(\mu, \sigma^2)$ represent a Gaussian distribution variable with mean μ and variance σ^2. Then*

$$D_{kl}\left(\mathcal{N}(\mu_1, \sigma^2)\|\mathcal{N}(\mu_2, \sigma^2)\right) = \frac{(\mu_1 - \mu_2)^2}{2\sigma^2}$$

Using this lemma, it follows that

$$D_{kl}\left(P(v_t|v_1, \ldots, v_{t-1})\|Q(v_t|v_1, \ldots, v_{t-1})\right) \leq \frac{(2\mu w_{t,i})^2}{2\max\{1, \|\mathbf{w}_t\|^4\}} = \frac{2\mu^2 w_{t,i}^2}{\max\{1, \|\mathbf{w}_t\|^4\}}.$$

Plugging this upper bound on the $U_{t,i}$'s in Lemma 4, we can further lower bound on the expected optimization error from Eq. (1) by

$$\frac{d\mu^2}{4}\left(1 - \sqrt{\frac{1}{d}\sum_{t=1}^{T}\sum_{i=1}^{d}\frac{2\mu^2 w_{t,i}^2}{\max\{1, \|\mathbf{w}_t\|^4\}}}\right) = \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{2\mu^2}{d}\sum_{t=1}^{T}\frac{\|\mathbf{w}\|_t^2}{\max\{1, \|\mathbf{w}_t\|^4\}}}\right)$$

$$= \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{2\mu^2}{d}\sum_{t=1}^{T}\min\left\{\|\mathbf{w}_t\|^2, \frac{1}{\|\mathbf{w}_t\|^2}\right\}}\right) \geq \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{2T\mu^2}{d}}\right). \tag{4}$$

Finally, we choose $\mu = \min\{1/\sqrt{d}, \sqrt{d/4T}\}$, and obtain a lower bound of

$$\frac{1}{4}\left(1 - \frac{1}{\sqrt{2}}\right)\min\left\{1, \frac{d^2}{4T}\right\} > 0.01\min\left\{1, \frac{d^2}{T}\right\}$$

as required. ∎

The theorem above applies to the optimization error for derivative-free SCO. We now turn to deal with the case of bandit SCO and regret, showing an $\Omega(\sqrt{d^2/T})$ lower bound.

Since the derivative-free SCO bound was $\Theta(d^2/T)$, the result implies a real gap between what can be obtained in terms of average regret, as opposed to optimization error, without any specific distributional assumptions. This stands in contrast to settings such as multi-armed bandits, where the construction implying the known $\Omega(\sqrt{d/T})$ lower bound (e.g. Cesa-Bianchi and Lugosi (2006)) applies equally well to derivative-free and bandit SCO (see Bubeck et al. (2011)).

Theorem 6 *Let the number of rounds T be fixed. Then for any (possibly randomized) querying strategy, there exists a quadratic function of the form $F(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 - \langle\mathbf{e}, \mathbf{w}\rangle$, which is minimized at $\mathbf{e}$ where $\|\mathbf{e}\| \leq 1/2$, such that*

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} F(\mathbf{w}_t) - F(\mathbf{w}^*)\right] \geq 0.02 \min\left\{1, \sqrt{\frac{d^2}{T}}\right\}.$$

Note that our lower bound holds even when the domain is unrestricted (the algorithm can pick any point in $\mathbb{R}^d$). Moreover, the lower bound coincides (up to a constant) with the $\mathcal{O}(\sqrt{d^2/T})$ regret upper-bound shown for strongly-convex and smooth functions in Agarwal et al. (2010). This shows that for strongly-convex and smooth functions, the minimax average regret is $\Theta(\sqrt{d^2/T})$. Also, the lower bound implies that one cannot hope to obtain average regret better than $\sqrt{d^2/T}$ for more general bandit problems, such as strongly-convex or even convex problems.

The proof relies on techniques similar to the lower bound of Thm. 3, with a key additional insight. Specifically, in Thm. 3, the lower bound obtained actually depends on the norm of the points $\mathbf{w}_1, \ldots, \mathbf{w}_T$ (see Eq. (4)), and the optimal $\mathbf{w}^*$ has a very small norm. In a regret minimization setting the points $\mathbf{w}_1, \ldots, \mathbf{w}_T$ cannot be too far from $\mathbf{w}^*$, and thus must have a small norm as well, leading to a stronger lower bound than that of Thm. 3. The formal proof appears in the appendix.

4. Strongly Convex and Smooth Functions

We now turn to the more general case of strongly convex and smooth functions. First, we note that in the case of functions which are both strongly convex and smooth, (Agarwal et al., 2010, Theorem 14) already provided an $\mathcal{O}(\sqrt{d^2/T})$ average regret bound (which holds even in a non-stochastic setting). The main result of this section is a *matching* lower bound, which holds even if we look at the much easier case of derivative-free SCO. This lower bound implies that the attainable error for strongly-convex and smooth functions is order of $\sqrt{d^2/T}$, and at least $\sqrt{d^2/T}$ for any harder setting.

Theorem 7 *Let the number of rounds T be fixed. Then for any (possibly randomized) querying strategy, there exists a function F over $\mathbb{R}^d$ which is 0.5-strongly convex and 3.5-smooth; Is 4-Lipschitz over the unit Euclidean ball; has a global minimum in the unit ball; And such that the resulting $\bar{\mathbf{w}}_T$ satisfies*

$$\mathbb{E}[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)] \geq 0.004 \min\left\{1, \sqrt{\frac{d^2}{T}}\right\}.$$

Note that we made no attempt to optimize the constant.

The general proof technique is rather similar to that of Thm. 3, but the construction is a bit more intricate. Specifically, letting $\mu > 0$ be a parameter to be determined later, we look at functions of the form

$$F_{\mathbf{e}}(\mathbf{w}) = \|\mathbf{w}\|^2 - \sum_{i=1}^{d} \frac{e_i w_i}{1 + (w_i/e_i)^2},$$

where $\mathbf{e}$ is uniformly distributed on $\{-\mu, +\mu\}^d$. To see the intuition behind this choice, let us consider the one-dimensional case $(d = 1)$. Recall that in the quadratic setting, the function we considered (in one dimension) was of the form

$$F_e(w) = \frac{1}{2}w^2 - ew,$$

where e was chosen uniformly at random from $\{-\mu, +\mu\}$, and μ is a "small" number. Thus, the optimum is at either $-\mu$ or μ, and the difference $|F_\mu(w) - F_{-\mu}(w)|$ at these optima is order of μ^2. However, by picking $w = \Theta(1)$, the difference $|F_\mu(w) - F_{-\mu}(w)|$ is on the order of μ - much larger than the difference close to the optimum, which is order of μ^2. Therefore, by querying for w far from the optimum, and getting noisy values of F_e, it is easier to distinguish whether we are dealing with $e = +\mu$ or $e = -\mu$, leading to a d^2/T optimization error bound. In contrast, the function we consider here (in the one-dimensional case) is of the form

$$F_e(w) = w^2 - \frac{ew}{1 + (w/e)^2}. \tag{5}$$

This form is carefully designed so that $|F_\mu(w) - F_{-\mu}(w)|$ is order of μ^2, not just at the optima of F_μ and $F_{-\mu}$, but for *all* w. This is because of the additional denominator, which makes the function closer and closer to w^2 the larger w is - see Fig. 1 for a graphical illustration. As a result, no matter how the function is queried, distinguishing the choice of μ is difficult, leading to the strong lower bound of Thm. 7. A formal proof is presented in the appendix.

5. Discussion

In this paper, we considered the dual settings of bandit and derivative-free stochastic convex optimization. We provided a sharp characterization of the attainable performance for strongly-convex and smooth functions. The results also provide useful lower-bounds for more general settings. We also considered the case of quadratic functions, showing that a "fast" $\mathcal{O}(1/T)$ rate is possible in a stochastic setting, even without knowledge of derivatives. Our results have several qualitative differences compared to previously known results which focus on linear functions, such as quadratic dependence on the dimension even for extremely "nice" functions, and a provable gap between the attainable performance in bandit optimization and derivative-free optimization.

Our work leaves open several questions. For example, we have only dealt with bounds which hold in expectation, and our lower bounds focused on the dependence on d, T, where other problem parameters, such as the Lipschitz constant and strong convexity parameter, are fixed constants. While this follows the setting of previous works, it does not cover
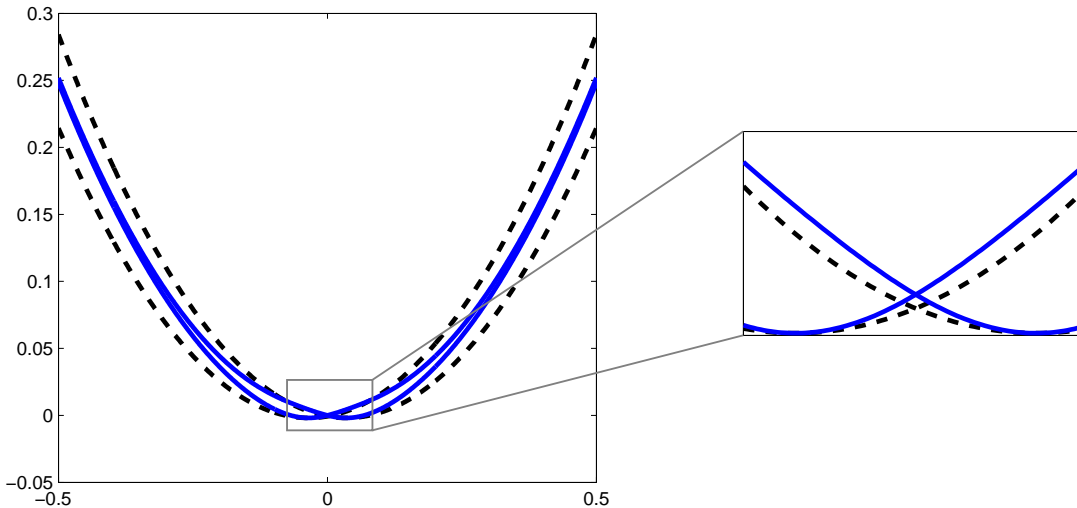
Figure 1: The two solid blue lines represents $F_e(w)$ as in Eq. (5), for $e = 0.1$ and $e = -0.1$, whereas the two dashed black lines represent two quadratic functions with similar minimum points. Close to the minima, $F_e(w)$ and the quadratic functions behave rather similarly. However, as we increase $|w|$, the two quadratic functions become rather distinguishable, whereas $F_e(w)$ become more and more *in*distinguishable for the two choices of e. Thus, distinguishing whether $e = 0.1$ or $e = -0.1$, based only on function values is of $F_e(w)$, is much harder than the quadratic case

situations where these parameters scale with d. Finally, while this paper settles the case of strongly-convex and smooth functions, we still don't know what is the attainable performance for general convex functions, as well as the more specific case of strongly-convex (possibly non-smooth) functions. Our $\Omega\left(\sqrt{d^2/T}\right)$ lower bound still holds, but the existing upper bounds are much larger: $\min\left\{\sqrt[4]{d^2/T}, \sqrt{d^{32}/T}\right\}$ for convex functions, and $\min\left\{\sqrt[3]{d^2/T}, \sqrt{d^{32}/T}\right\}$ for strongly-convex functions (see table 1). We don't know if the lower bound or the existing upper bounds are tight. However, it is the current upper bounds which seem less "natural", and we suspect that they are the ones that can be considerably improved, using new algorithms which remain undiscovered.

Acknowledgments

References

Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári. Improved algorithms for linear stochastic bandits. In *NIPS*, 2011.

A. Agarwal, O. Dekel, and L. Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, 2010.

A. Agarwal, D. Foster, D. Hsu, S. Kakade, and A. Rakhlin. Stochastic convex optimization with bandit feedback. In *NIPS*, 2011.

E. Arias-Castro, E. Candès, and M. Davenport. On the fundamental limits of adaptive sensing. *CoRR*, abs/1111.4646, 2011.

J.-Y. Audibert and S. Bubeck. Minimax policies for adversarial and stochastic bandits. In *COLT*, 2009.

J.-Y. Audibert, S. Bubeck, and G. Lugosi. Minimax policies for combinatorial prediction games. *COLT*, 2011.

P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. The nonstochastic multiarmed bandit problem. *SIAM J. Comput.*, 32(1):48–77, 2002.

S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *CoRR*, abs/1204.5721, 2012.

S. Bubeck, R. Munos, and G. Stoltz. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science*, 412(19):1832–1852, 2011.

S. Bubeck, N. Cesa-Bianchi, and S. Kakade. Towards minimax policies for online linear optimization with bandit feedback. In *COLT*, 2012.

N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

T. Cover and J. Thomas. *Elements of information theory*. Wiley, 2 edition, 2006.

A.B. Cybakov. *Introduction to nonparametric estimation*. Springer series in statistics. Springer, 2009.

V. Dani, T. Hayes, and S. Kakade. The price of bandit information for online optimization. In *NIPS*, 2007.

V. Dani, T. Hayes, and S. Kakade. Stochastic linear optimization under bandit feedback. In *COLT*, 2008.

A. Flaxman, A. Kalai, and B. McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. In *SODA*, 2005.

E. Hazan and S. Kale. Beyond the regret minimization barrier: an optimal algorithm for stochastic strongly-convex optimization. In *COLT*, 2011.

K. Jamieson, R. Nowak, and B. Recht. Query complexity of derivative-free optimization. *CoRR*, abs/1209.2434, 2012.

S. Kullback. *Information Theory and Statistics*. Dover, 1959.

A. Nemirovsky and D. Yudin. *Problem Complexity and Method Efficiency in Optimization.* Wiley-Interscience, 1983.

Y. Nesterov. Random gradient-free minimization of convex functions. Technical Report 16, ECORE Discussion Paper, 2011.

A. Rakhlin, O. Shamir, and K. Sridharan. Making gradient descent optimal for strongly convex stochastic optimization. In *ICML*, 2012.

S. Stich, C. Müller, and B. Gärtner. Optimization of convex functions with random pursuit. *CoRR*, abs/1111.0194, 2011.

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, 2003.

Appendix A. Improved Results for Quadratic Functions

In Sec. 3, we showed a tight $\Theta(d^2/T)$ bound on the achievable error for quadratic functions, in the derivative-free SCO setting. This was shown under the assumption that the noise $\xi_{\mathbf{w}}$ is zero-mean and has a second moment bounded by $\max\{1, \|\mathbf{w}\|^2\}$. In this appendix, we show how under additional natural assumptions on the noise, one can improve on this result with an efficient algorithm. The main message here is not so much the algorithmic result, but rather to show that the generic noise assumption is important for our lower bounds, and that better algorithms may still be possible for more specific settings.

To give a concrete example, consider the classic setting of ridge regression, where we have labeled training examples $(\mathbf{x}, y)$ sampled i.i.d. from some distribution over $\mathbb{R}^d \times \mathbb{R}$, and our goal is to find some $\mathbf{w} \in \mathbb{R}^d$ minimizing

$$F(\mathbf{w}) = \frac{\lambda}{2}\|\mathbf{w}\|^2 + \mathbb{E}_{(\mathbf{x},y)}\left[\left(\mathbf{w}^\top \mathbf{x} - y\right)^2\right].$$

In a bandit / derivative-free SCO setting, we can think of each query as giving as the value of

$$\hat{F}(\mathbf{w}) = \frac{\lambda}{2}\|\mathbf{w}\|^2 + \left(\mathbf{w}^\top \mathbf{x} - y\right)^2. \tag{6}$$

for some specific example $(\mathbf{x}, y)$, and note that its expected value (over the random draw of $(\mathbf{x}, y)$) equals $F(\mathbf{w})$. Thus, it falls within the setting considered in this paper. However, the noise process is not generic, but has a particular structure. We will show here that one can actually attain an error rate as good as $\mathcal{O}(d/T)$ for this problem.

To formally present our result, it would be useful to consider a more general setting, the ridge regression setting above being a special case. Suppose we can write $F(\mathbf{w})$ as $\mathbb{E}[\hat{F}(\mathbf{w})]$, where $\hat{F}(\mathbf{w})$ decomposes into a deterministic term $R(\mathbf{w})$ and a stochastic quadratic term $\hat{G}(\mathbf{w})$:

$$\hat{F}(\mathbf{w}) \;=\; R(\mathbf{w}) + \hat{G}(\mathbf{w}) \;=\; R(\mathbf{w}) + \left(\mathbf{w}^\top \hat{A}\mathbf{w} + \hat{\mathbf{b}}^\top \mathbf{w} + \hat{c}\right),$$

where $\hat{A}, \hat{\mathbf{b}}, \hat{c}$ are random variables. We assume that whenever we query a point $\mathbf{w}$, we get $\hat{F}(\mathbf{w})$ for some random realization of $\hat{A}, \hat{\mathbf{b}}, \hat{c}$. In general, $R(\mathbf{w})$ can be a strongly-convex regularization term, such as $\frac{\lambda}{2}\|\mathbf{w}\|^2$ in Eq. (6).

The algorithm we consider, Algorithm 2, is a slight variant of Algorithm 1, which takes this decomposition of $F(\mathbf{w})$ into account when constructing its unbiased gradient estimate. Compared to Algorithm 1, this algorithm also queries at random points further away from $\mathbf{w}_t$, up to a distance of $\sqrt{d}$. We will assume here that we can always query at such points[3]. We also let $\bar{\mathcal{W}} = \mathcal{W} \cap \{\mathbf{w} : \|\mathbf{w}\| \leq B\}$ in the algorithm, where we recall that B is some known upper bound on $\|\mathbf{w}^*\|$.

Algorithm 2 Derivative-Free SCO Algorithm for Decomposable-Quadratic Functions

---

Input: Deterministic term $R(\cdot)$; Strong convexity parameter $\lambda > 0$
Initialize $\mathbf{w}_1 = \mathbf{0}$.
for $t = 1, \ldots, T-1$ **do**
  Pick $\mathbf{r} \in \{-1, +1\}^d$ uniformly at random
  Query noisy function value v at point $\mathbf{w}_t + \mathbf{r}$
  Let $\tilde{\mathbf{g}} = (v - R(\mathbf{w}_t + \mathbf{r})) \mathbf{r} + \mathbf{g}_R(\mathbf{w}_t)$, where $\mathbf{g}_R(\mathbf{w})$ is a subgradient of $R(\cdot)$ at $\mathbf{w}$
  Let $\mathbf{w}_{t+1} = \Pi_{\bar{\mathcal{W}}}\left(\mathbf{w}_t - \frac{1}{\lambda t}\tilde{\mathbf{g}}\right)$
end for
Return $\bar{\mathbf{w}}_T = \mathbf{w}_T \frac{2}{T} \sum_{t=T/2}^{T} \mathbf{w}_t$.

We now show that with this algorithm, one can improve on our $\mathcal{O}(d^2/T)$ error upper bound from (Thm. 1).

Theorem 8 *In the setting described above, suppose $\|\hat{A}\|_2, \|\hat{\mathbf{b}}\|, |\hat{c}|$ are all at most 1 with probability 1, the optimum $\mathbf{w}^*$ has a norm of at most B, and $\|\mathbf{g}_R(\mathbf{w})\| \leq N$ for any $\mathbf{w} \in \bar{\mathcal{W}}$. Then under Assumption 1, the point $\bar{\mathbf{w}}_T$ returned by Algorithm 2 satisfies*

$$\mathbb{E}\left[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)\right] \leq 4(4 + 5\log(2))\frac{N^2 + 3d\left((B+1)^4 + \mathbb{E}\left[\|\hat{A}\|_F^2\right]\right)}{\lambda T},$$

where $\|\cdot\|_F$ is the Frobenius norm.

Note that if we only assume $\|\hat{A}\|_2 \leq 1$, then $\|\hat{A}\|_F^2$ can be as high as d, which leads to an $\mathcal{O}(d^2/T)$ bound, same as in Thm. 1. However, it may be much smaller than that. In particular, for the ridge regression case we considered earlier, $\hat{A}$ corresponds to $\mathbf{x}\mathbf{x}^\top$ where $\mathbf{x}$ is a randomly drawn instance. Under the common assumption that $\|\mathbf{x}\| \leq \mathcal{O}(1)$ (independent of the dimension), it follows that $\|\mathbf{x}\mathbf{x}^\top\|_F^2 = \|\mathbf{x}\|^4 = \mathcal{O}(1)$. Therefore, $\|\hat{A}\|_F^2$ is independent of the dimension, leading to an $\mathcal{O}(d/T)$ error upper bound in terms of d, T.

We remark that even in this specific setting, the $\mathcal{O}(d/T)$ bound does not carry over to the bandit SCO setting (i.e. in terms of regret), since the algorithm requires us to query far away from $\mathbf{w}_t$. Also, we again emphasize that this result does not contradict our lower bound in the quadratic case (Thm. 3), since the setting there included a generic noise term, while here the stochastic "noise" has a very specific structure.

As to the proof of Thm. 8, it is very similar to that of Thm. 1, the key difference being a better moment upper bound on the gradient estimate $\bar{\mathbf{g}}^2$, as formalized in the following lemma. Plugging this improved bound into the calculations results in the theorem.

---

3. Similar to Algorithm 1, if one can only query at some distance $\epsilon\sqrt{d}$, where $\epsilon \in (0, 1]$, then one can modify the algorithm to handle such cases, with the resulting error bound depending on ϵ.

Lemma 9 *For any $\mathbf{w}_t$, we have that $\mathbb{E}_{\mathbf{r},v}[\tilde{\mathbf{g}}]$ is a subgradient of $F(\mathbf{w}_t)$, and*

$$\mathbb{E}_{\mathbf{r},v}[\|\tilde{\mathbf{g}}\|^2] \leq 4\left(N^2 + 3d\left((B+1)^4 + \mathbb{E}\left[\|\hat{A}\|_F^2\right]\right)\right).$$

Proof By definition of $F(\mathbf{w}_t)$, we note that

$$\tilde{\mathbf{g}} = \left((\mathbf{w}_t + \mathbf{r})^\top \hat{A}(\mathbf{w}_t + \mathbf{r}) + \hat{\mathbf{b}}^\top(\mathbf{w}_t + \mathbf{r}) + \hat{c}\right)\mathbf{r} + \mathbf{g}_R(\mathbf{w}_t).$$

Using a similar calculation to the one in the proof of Lemma 2, we have that the expected value of this expression over $\mathbf{r}$ and $\hat{A}, \hat{\mathbf{b}}, \hat{c}$ is

$$2\mathbf{w}_t^\top \mathbb{E}[\hat{A}] + \mathbb{E}[\hat{\mathbf{b}}^\top] + \mathbf{g}_R(\mathbf{w}_t),$$

which is a subgradient of $F(\mathbf{w}_t)$. As to the moment bound, we have

$$\mathbb{E}[\|\tilde{\mathbf{g}}\|^2] \leq \mathbb{E}\left[4\left((\mathbf{w}_t + \mathbf{r})^\top \hat{A}(\mathbf{w}_t + \mathbf{r})\right)^2 \|\mathbf{r}\|^2 + 4\left(\hat{\mathbf{b}}^\top(\mathbf{w}_t + \mathbf{r})\right)^2 \|\mathbf{r}\|^2 + 4\hat{c}^2\|\mathbf{r}\|^2 + 4\|\mathbf{g}_R(\mathbf{w}_t)\|^2\right]$$

$$\leq 4d\,\mathbb{E}\left[\left(\|\hat{A}\|_2\|\mathbf{w}_t\|^2 + 2\mathbf{w}_t^\top \hat{A}\mathbf{r} + \mathbf{r}^\top \hat{A}\mathbf{r}\right)^2 + 2\left(\left(\hat{\mathbf{b}}^\top \mathbf{w}_t\right)^2 + \left(\hat{\mathbf{b}}^\top \mathbf{r}\right)^2\right) + 1\right] + 4N^2$$

$$= 4d\,\mathbb{E}\left[\left(B^2 + 2\mathbf{w}_t^\top \hat{A}\mathbf{r} + \mathbf{r}^\top A\mathbf{r}\right)^2 + 2\left(B^2 + \left(\hat{\mathbf{b}}^\top \mathbf{r}\right)^2\right) + 1\right] + 4N^2$$

$$= 12d\left(B^4 + 4\mathbb{E}\left[\left(\mathbf{w}_t^\top \hat{A}\mathbf{r}\right)^2\right] + \mathbb{E}\left[\left(\mathbf{r}^\top A\mathbf{r}\right)^2\right]\right) + 8d\left(B^2 + \mathbb{E}\left[\left(\hat{\mathbf{b}}^\top \mathbf{r}\right)^2\right]\right) + 4d + 4N^2.$$

$$(7)$$

Letting $\hat{a}_{i,j}$ denote entry (i,j) in $\hat{A}$, and recalling that by definition of $\mathbf{r}$, $\mathbb{E}[r_i r_j] = \mathbf{1}_{i=j}$, we have that

$$\mathbb{E}\left[\left(\mathbf{r}^\top A\mathbf{r}\right)^2\right] = \mathbb{E}\left[\left(\sum_{i,j} r_i r_j \hat{a}_{i,j}\right)^2\right] = \mathbb{E}\left[\sum_{i,j,i',j'} r_i r_j r_{i'} r_{j'} \hat{a}_{i,j} \hat{a}_{i',j'}\right]$$

$$= \mathbb{E}\left[\sum_{i,j} r_i^2 r_j^2 \hat{a}_{i,j}^2\right] = \mathbb{E}\left[\sum_{i,j} \hat{a}_{i,j}^2\right] = \mathbb{E}\left[\|\hat{A}\|_F^2\right].$$

Also, using the fact that $\mathbb{E}[\mathbf{rr}^\top]$ is the identity matrix, we have

$$\mathbb{E}\left[\left(\mathbf{w}_t^\top \hat{A}\mathbf{r}\right)^2\right] = \mathbb{E}\left[\mathbf{w}_t^\top \hat{A}\mathbf{rr}^\top \hat{A}^\top \mathbf{w}_t\right] = \mathbb{E}\left[\mathbf{w}_t^\top \hat{A}\hat{A}^\top \mathbf{w}_t\right] \leq \mathbb{E}\left[\|\mathbf{w}_t\|^2\|\hat{A}\|_2^2\right] \leq B^2.$$

Finally, we have

$$\mathbb{E}\left[\left(\hat{\mathbf{b}}^\top \mathbf{r}\right)^2\right] = \mathbb{E}\left[\hat{\mathbf{b}}^\top \mathbf{rr}^\top \hat{\mathbf{b}}\right] = \mathbb{E}\left[\|\hat{\mathbf{b}}\|^2\right] \leq 1.$$

Plugging these inequalities back into Eq. (7), we get that

$$\mathbb{E}[\|\mathbf{g}\|^2] \leq 12d\left(B^4 + 4B^2 + \mathbb{E}[\|\hat{A}\|_F^2]\right) + 8d\left(B^2 + 1\right) + 4d + 4N^2$$

$$= 4d\left(3B^4 + 14B^2 + 3 + 3\,\mathbb{E}\left[\|\hat{A}\|_F^2\right]\right) + 4N^2$$

$$\leq 12d\left((B+1)^4 + \mathbb{E}\left[\|\hat{A}\|_F^2\right]\right) + 4N^2,$$

from which the lemma follows.   ∎

Appendix B. Additional Proofs

B.1. Proof of Lemma 2

By the way $\mathbf{r}$ is picked, we have that $\mathbb{E}_{\mathbf{r}}[\mathbf{r}_i\mathbf{r}_j] = \mathbf{1}_{i=j}$ and that $\mathbb{E}_{\mathbf{r}}[\mathbf{r}_i\mathbf{r}_j\mathbf{r}_k] = 0$ for all i,j,k. Thus, letting $\mathbb{E}$ denote expectation w.r.t. $\mathbf{r}$ and the random function values, we have

$$
\begin{aligned}
\mathbb{E}[\tilde{\mathbf{g}}] &= \mathbb{E}\left[\frac{\sqrt{d}v}{\epsilon}\mathbf{r}\right] \\
&= \mathbb{E}\left[\frac{\sqrt{d}}{\epsilon}\left(\left(\mathbf{w}+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\right)^{\top}A\left(\mathbf{w}+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\right)+\mathbf{b}^{\top}\left(\mathbf{w}+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\right)+c+\xi_{\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}}\right)\mathbf{r}\right] \\
&= \mathbb{E}\left[\frac{\sqrt{d}}{\epsilon}\left(\mathbf{w}^{\top}A\mathbf{w}+\mathbf{b}^{\top}\mathbf{w}+c+\xi_{\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}}\right)\mathbf{r}+\frac{\epsilon}{\sqrt{d}}\left(\mathbf{r}^{\top}A\mathbf{r}\right)\mathbf{r}\right]+\mathbb{E}\left[\left(2\mathbf{w}^{\top}A\mathbf{r}\right)\mathbf{r}+\left(\mathbf{b}^{\top}\mathbf{r}\right)\mathbf{r}\right] \\
&= 0+2\mathbf{w}^{\top}A+\mathbf{b}^{\top}+0 = \nabla F(\mathbf{w}).
\end{aligned}
$$

Also, by the assumptions on $A,\mathbf{b},c$ and the assumptions on the noise $\xi_{\mathbf{w}}$, we have

$$
\begin{aligned}
\mathbb{E}[\|\tilde{\mathbf{g}}\|^2] &= \mathbb{E}\left[\frac{dv^2}{\epsilon^2}\|\mathbf{r}\|^2\right] = \frac{d^2}{\epsilon^2}\mathbb{E}[v^2] = \frac{d^2}{\epsilon^2}\mathbb{E}\left[\left(F\left(\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\right)+\xi_{\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}}\right)^2\right] \\
&\leq \frac{d^2}{\epsilon^2}\mathbb{E}\left[2\left(F\left(\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\right)\right)^2+2\xi^2_{\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}}\right] \\
&\leq \frac{2d^2}{\epsilon^2}\left(\sup_{\mathbf{w}:\|\mathbf{w}\|\leq B+\epsilon}(F(\mathbf{w}))^2+\max\left\{1,\|\mathbf{w}_t+\frac{\epsilon}{\sqrt{d}}\mathbf{r}\|^2\right\}\right) \\
&\leq \frac{2d^2}{\epsilon^2}\left(\sup_{\mathbf{w}:\|\mathbf{w}\|\leq B+\epsilon}\left(\mathbf{w}^{\top}A\mathbf{w}+\mathbf{b}^{\top}\mathbf{w}+c\right)^2+(B+1)^2\right) \\
&\leq \frac{2d^2}{\epsilon^2}\left(\left((B+\epsilon)^2+(B+\epsilon)+1\right)^2+(B+1)^2\right) \\
&\leq \frac{4d^2}{\epsilon^2}(B+1)^4
\end{aligned}
$$

as required.

B.2. Proof of Lemma 4

We have the following:

$$
\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{d} \mathbf{1}_{\bar{w}_i e_i < 0}\right] &= \sum_{i=1}^{d} \Pr\left(\bar{w}_i e_i < 0\right) \\
&= \frac{1}{2} \sum_{i=1}^{d} \left(\Pr(\bar{w}_i < 0 | e_i > 0) + \Pr(\bar{w}_i > 0 | e_i < 0)\right) \\
&= \frac{1}{2}\left(d - \sum_{i=1}^{d}\left(\Pr(\bar{w}_i > 0 | e_i > 0) - \Pr(\bar{w}_i > 0 | e_i < 0)\right)\right) \\
&\geq \frac{d}{2}\left(1 - \frac{1}{d}\sum_{i=1}^{d}\left|\Pr(\bar{w}_i > 0 | e_i > 0) - \Pr(\bar{w}_i > 0 | e_i < 0)\right|\right) \\
&\geq \frac{d}{2}\left(1 - \sqrt{\frac{1}{d}\sum_{i=1}^{d}\left(\Pr(\bar{w}_i > 0 | e_i > 0) - \Pr(\bar{w}_i > 0 | e_i < 0)\right)^2}\right), \quad (8)
\end{aligned}
$$

where the last inequality is by the fact that for any values $a_1, \ldots, a_d$, it holds that $|a_1| + \ldots + |a_d| \leq \sqrt{d}\sqrt{a_1^2 + \ldots + a_d^2}$.

Consider (without loss of generality) the term corresponding to the first coordinate, namely

$$
\left(\Pr(\bar{w}_1 > 0 | e_1 > 0) - \Pr(\bar{w}_1 > 0 | e_1 < 0)\right)^2.
$$

This term equals

$$
\left(\sum_{e_2, \ldots, e_d} \Pr(\{e_j\}_{j=2}^{d}) \left(\Pr\left(\bar{w}_1 > 0 | e_1 > 0, \{e_j\}_{j=2}^{d}\right) - \Pr\left(\bar{w}_1 > 0 | e_1 < 0, \{e_j\}_{j=2}^{d}\right)\right)\right)^2
$$

$$
\leq \sum_{e_2, \ldots, e_d} \Pr(\{e_j\}_{j=2}^{d}) \left(\Pr\left(\bar{w}_1 > 0 | e_1 > 0, \{e_j\}_{j=2}^{d}\right) - \Pr\left(\bar{w}_1 > 0 | e_1 < 0, \{e_j\}_{j=2}^{d}\right)\right)^2
$$

$$
\leq \sup_{e_2, \ldots, e_d} \left(\Pr\left(\bar{w}_1 > 0 | e_1 > 0, \{e_j\}_{j=2}^{d}\right) - \Pr\left(\bar{w}_1 > 0 | e_1 < 0, \{e_j\}_{j=2}^{d}\right)\right)^2
$$

By Pinsker's inequality and the assumption that $\bar{\mathbf{w}}_T$ is a deterministic function of $v_1, \ldots, v_T$, this expression is at most

$$
\frac{1}{2} D_{kl}\left(\Pr\left(v_1, \ldots, v_T | e_1 > 0, \{e_j\}_{j=2}^{d}\right) \| \Pr\left(v_1, \ldots, v_T | e_1 < 0, \{e_j\}_{j=2}^{d}\right)\right),
$$

where $D_{kl}(P\|Q)$ is the Kullback-Leibler divergence between the two distributions. By the chain rule (see e.g. Cover and Thomas (2006)), we can upper bound the above by

$$
\frac{1}{2} \sum_{t=1}^{T} D_{kl}\left(\Pr\left(v_t | e_1 > 0, \{\mathbf{e}_j\}_{j=2}^{d}, \{v_l\}_{l=1}^{t-1}\right) \| \Pr\left(v_t | e_1 < 0, \{\mathbf{e}_j\}_{j=2}^{d}, \{v_l\}_{l=1}^{t-1}\right)\right).
$$

Plugging these bounds back into Eq. (8), the result follows.

B.3. Proof of Thm. 6

We may assume without loss of generality that $T \geq d^2$, and it is enough to show that the expected average regret is at least $0.02\sqrt{d^2/T}$. This is because if there was a strategy with < 0.02 average regret after $T < d^2$ rounds, then for the case of d^2 rounds, we could just run that strategy for T rounds, compute the average $\bar{\mathbf{w}}_T$ of all points played so far, and then repeatedly choose $\bar{\mathbf{w}}_T$ in the remaining rounds. By Jensen's inequality, this would imply a < 0.02 average regret after d^2 rounds, in contradiction.

Let $\bar{\mathbf{w}}_T$ be an arbitrary deterministic function of $\mathbf{w}_1, \ldots, \mathbf{w}_T$. A proof identical to that of Thm. 3, up to Eq. (4), implies that for any $\mu > 0$, there exists a quadratic function of the form

$$F_{\mathbf{e}} = \frac{1}{2}\|\mathbf{w}\|^2 - \langle \mathbf{e}, \mathbf{w} \rangle,$$

with $\mathbf{e} \in \{-\mu, \mu\}^d$, such that

$$\mathbb{E}[F_{\mathbf{e}}(\bar{\mathbf{w}}_T) - F_{\mathbf{e}}(\mathbf{w}^*)] \geq \mathbb{E}\left[\frac{d\mu^2}{4}\left(1 - \sqrt{\frac{\mu^2}{d}\sum_{t=1}^{T}\min\left\{\|\mathbf{w}_t\|^2, \frac{1}{\|\mathbf{w}_t\|^2}\right\}}\right)\right].$$

In particular, letting $\bar{\mathbf{w}}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$, using Jensen's inequality, and discarding the min, we get that

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}F_{\mathbf{e}}(\mathbf{w}_t) - F_{\mathbf{e}}(\mathbf{w}^*)\right] \geq \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{\mu^2}{d}\sum_{t=1}^{T}\|\mathbf{w}_t\|^2}\right). \tag{9}$$

However, we also know that by strong convexity of $F_{\mathbf{e}}$, we have

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}F_{\mathbf{e}}(\mathbf{w}_t) - F_{\mathbf{e}}(\mathbf{w}^*)\right] \geq \frac{1}{2T}\sum_{t=1}^{T}\|\mathbf{w}_t - \mathbf{e}\|^2. \tag{10}$$

Using the fact that

$$\|\mathbf{w}_t\|^2 = \|\mathbf{w}_t - \mathbf{e} + \mathbf{e}\|^2 \leq (\|\mathbf{w}_t - \mathbf{e}\| + \|\mathbf{e}\|)^2 \leq 2\|\mathbf{w}_t - \mathbf{e}\|^2 + 2\|\mathbf{e}\|^2,$$

we get that

$$\|\mathbf{w}_t - \mathbf{e}\|^2 \geq \frac{1}{2}\|\mathbf{w}_t\|^2 - \|\mathbf{e}\|^2 = \frac{1}{2}\|\mathbf{w}_t\|^2 - d\mu^2.$$

Substituting into Eq. (10) and slightly manipulating the resulting inequality, we get

$$\sum_{t=1}^{T}\|\mathbf{w}_t\|^2 \leq 4T\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}F_{\mathbf{e}}(\mathbf{w}_t) - F_{\mathbf{e}}(\mathbf{w}^*)\right] + 2Td\mu^2.$$

For simplicity, denote the average regret term $\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}F_{\mathbf{e}}(\mathbf{w}_t) - F_{\mathbf{e}}(\mathbf{w}^*)\right]$ by R. Substituting the expression above into Eq. (9), we get

$$R \geq \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{\mu^2}{d}\left(4TR + 2Td\mu^2\right)}\right) \geq \frac{d\mu^2}{4}\left(1 - \sqrt{\frac{4\mu^2 TR}{d}} - \sqrt{2T\mu^4}\right).$$

Rearranging and simplifying, we get

$$R + \frac{\sqrt{dT}}{2}\mu^3\sqrt{R} + \frac{d\mu^2}{4}\left(\mu^2\sqrt{2T} - 1\right) \geq 0.$$

The equation above can be seen as a quadratic function of $\sqrt{R}$, with the roots

$$\frac{1}{2}\left(-\frac{\sqrt{dT}}{2}\mu^3 \pm \sqrt{\left(\frac{\sqrt{dT}}{2}\mu^3\right)^2 + d\mu^2\left(1 - \mu^2\sqrt{2T}\right)}\right).$$

Now, recall that μ is a free parameter that we can choose at will. If we choose it so that $1 - \mu^2\sqrt{2T} > 0$, then it is easy to show that we get two roots, one strictly positive and one strictly negative. Since we know $\sqrt{R}$ is a nonnegative quantity, we get that

$$\begin{aligned}
\sqrt{R} &\geq \frac{1}{2}\left(-\frac{\sqrt{dT}}{2}\mu^3 + \sqrt{\left(\frac{\sqrt{dT}}{2}\mu^3\right)^2 + d\mu^2\left(1 - \mu^2\sqrt{2T}\right)}\right) \\
&= \frac{\sqrt{d}\mu}{2}\left(-\frac{\sqrt{T}}{2}\mu^2 + \sqrt{\frac{T}{4}\mu^4 + 1 - \mu^2\sqrt{2T}}\right).
\end{aligned}$$

Finally, choosing $\mu = T^{-1/4}/2$ (which indeed satisfies $1 - \mu^2\sqrt{2T} > 0$), and simplifying, we get

$$\sqrt{R} \geq 0.17\sqrt{\frac{d}{\sqrt{T}}}.$$

Recalling that R is the expected average regret, it only remains to take the square of the two sides. We note that since we assume $T \geq d^2$, then $\|\mathbf{e}\| = \sqrt{d}\mu = \sqrt{\sqrt{d^2/T}}/2 \leq 1/2$, as specified in the theorem statement.

B.4. Proof of Thm. 7

Let $\mu > 0$ be a parameter to be determined later. As discussed in the text, we will look at functions of the form

$$F_{\mathbf{e}}(\mathbf{w}) = \|\mathbf{w}\|^2 - \sum_{i=1}^{d} \frac{e_i w_i}{1 + (w_i/e_i)^2}, \tag{11}$$

where $\mathbf{e}$ is uniformly distributed on $\{-\mu, +\mu\}^d$. Our goal will be to prove a lower bound on the expected optimization error over the randomized choice of $F_{\mathbf{e}}$, with respect to deterministic querying strategies. As explained in the proof of Thm. 3, this would imply the existence of some fixed $F_{\mathbf{e}}$ such that the expected optimization error over a (possibly randomized) querying strategy is the same.

We will need the following properties of $F_{\mathbf{e}}$:

Lemma 10 *For any $\mu > 0$ and any $\mathbf{e} \in \{-\mu, +\mu\}^d$, the function $F_{\mathbf{e}}$ in Eq. (11) is:*

- *0.5-Strongly convex and 3.5-smooth*

- $2 + \sqrt{2d}\mu$-Lipschitz for any $\mathbf{w}$ such that $\|\mathbf{w}\| \leq 1$.

- $F_\mathbf{e}$ is globally minimized at $\mathbf{w}^* = c\mathbf{e}$, where $c = 0.3489... \geq 1/3$

- For any $\mathbf{e}' \in \{-\mu, +\mu\}^d$ which differs from $\mathbf{e}$ in a single coordinate, and for any $\mathbf{w} \in \mathbb{R}^d$, it holds that $|F_\mathbf{e}(\mathbf{w}) - F_{\mathbf{e}'}(\mathbf{w})| \leq \mu^2$.

Proof Note that we can write the function $F_\mathbf{e}(\mathbf{w})$ as $\sum_{i=1}^d g_{e_i}(w_i)$, where

$$g_a(x) = x^2 - \frac{ax}{1 + (x/a)^2}.$$

It is not hard to realize that to prove the lemma, it is enough to prove that:

1. $g_a(x)$ is 0.5-strongly convex and 3.5-smooth;

2. $|g_a'(x)|$ is at most[4] $2|x| + |a|$;

3. For all μ, $|g_\mu(x) - g_{-\mu}(x)| \leq \mu^2$;

4. $g_a(x)$ is minimized at ca where $c = 0.3489....$

To show item 1, we calculate the second derivative of $g_a(x)$, which is

$$2\left(1 + \frac{a^3 x(3a^2 - x^2)}{(a^2 + x^2)^3}\right).$$

By definition of strong convexity and smoothness, it is enough to show that this term is always at least 0.5 and at most 3.5. Substituting $x = ay$ and simplifying, we get

$$2\left(1 + \frac{y(3 - y^2)}{(1 + y^2)^3}\right).$$

It is a straightforward exercise to verify that $\left|\frac{y(3-y^2)}{(1+y^2)^3}\right|$ is at most $3/4$ for all $y \in \mathbb{R}$, hence the expression above is always in $[0.5, 3.5]$ as required.

As to item 2, we note that

$$g_a'(x) = 2x - \frac{a^5 - a^3 x^2}{(a^2 + x^2)^2} = 2x - a\frac{1 - (x/a)^2}{(1 + (x/a)^2)^2}.$$

For any value of x/a, the value of the fraction above is easily verified to be at most 1, hence we can upper bound $|g_a'(x)|$ by $2|x| + |a|$ as required.

As to item 3, we have

$$|g_\mu(x) - g_{-\mu}(x)| = \frac{2|\mu x|}{1 + (x/\mu)^2} = \mu^2 \frac{2|\mu x|}{\mu^2 + x^2} \leq \mu^2,$$

where the last step uses $\mu^2 + x^2 \geq 2|\mu x|$, which follows from the identity $(\mu + |x|)^2 \geq 0$.

---

4. Since this would imply that $\|\nabla F_\mathbf{e}(\mathbf{w})\|$ is at most $\sqrt{\sum_{i=1}^d (2|w_i| + \mu)^2} \leq \sqrt{\sum_{i=1}^d (4w_i^2 + 2\mu^2)} \leq \sqrt{\sum_{i=1}^d (4w_i^2)} + \sqrt{\sum_{i=1}^d (2\mu^2)} = 2\|\mathbf{w}\| + \sqrt{2d}\mu$, which is at most $2 + \sqrt{2d}\mu$ for any $\mathbf{w}$ in the unit ball.

Finally, as to item 4, we note that this function can be equivalently written as

$$g_a(x) = a^2 \left((x/a)^2 - \frac{(x/a)}{1 + (x/a)^2} \right).$$

Substituting $x = ay$, we get $a^2(y^2 - y/(1 + y^2))$. A numerical calculation reveals that the minimizing value of y is $0.3489...$, hence the minimizing value of x is $0.3489... * a$ as required. $\blacksquare$

We now begin to derive the lower bound. Using strong convexity and the lemma, we have

$$\mathbb{E}[F(\bar{\mathbf{w}}_T) - F(\mathbf{w}^*)] \geq \mathbb{E}\left[\frac{1}{4}\|\bar{\mathbf{w}}_T - \mathbf{w}^*\|^2\right] = \frac{1}{4}\mathbb{E}\left[\sum_{i=1}^d (\bar{w}_i - w_i^*)^2\right] \geq \frac{1}{4}\mathbb{E}\left[\sum_{i=1}^d (w_i^*)^2 \mathbf{1}_{\bar{w}_i w_i^* < 0}\right]$$

$$\geq \frac{1}{4}\mathbb{E}\left[\sum_{i=1}^d \left(\frac{e_i}{3}\right)^2 \mathbf{1}_{\bar{w}_i e_i < 0}\right] = \frac{\mu^2}{36}\mathbb{E}\left[\sum_{i=1}^d \mathbf{1}_{\bar{w}_i e_i < 0}\right] \quad (12)$$

We now lower bound this term using Lemma 4. To do so, we need to upper bound the KL divergence of the query values at round t under the two hypotheses $e_i = +\mu$ and $e_i = -\mu$, the other coordinates being fixed. We assume each noise term $\xi_{\mathbf{w}}$ is a standard Gaussian random variable. Thus, the query value that we see is distributed as

$$F_{\mathbf{e}}(\mathbf{w}_t) + \xi_{\mathbf{w}} = \|\mathbf{w}\|^2 - \sum_{j=1}^d \frac{e_j w_j}{1 + (w_j/e_j)^2} + \xi_{\mathbf{w}}.$$

where one of the coordinates i of $\mathbf{e}$ is either $+\mu$ or $-\mu$ and the other coordinates are fixed. This is a Gaussian distribution, with mean $F_{\mathbf{e}}(\mathbf{w}_t)$ and variance 1. By Lemma 10, the difference between the two means under the two cases $e_i = +\mu$, $e_i = -\mu$ is at most μ^2, so by Lemma 5, the KL-divergence is at most $\mu^4/2$. Using Lemma 4, this implies that Eq. (12) is at least

$$\frac{d\mu^2}{72}\left(1 - \sqrt{\frac{1}{d}\sum_{i=1}^d \sum_{t=1}^T \frac{\mu^4}{2}}\right) = \frac{d\mu^2}{72}\left(1 - \sqrt{\frac{T\mu^4}{2}}\right).$$

Picking $\mu = T^{-1/4}$, we get a lower bound of $d/144\sqrt{2T} > 0.004\sqrt{d^2/T}$.

Finally, note that for this choice of μ, by Lemma 10, our function $F_{\mathbf{e}}$ (for any realization of $\mathbf{e}$) is $2 + \sqrt{2d/\sqrt{T}}$- Lipschitz in the unit ball, and has a global minimum with norm at most $0.35\sqrt{d/\sqrt{T}}$. If $T \geq d^2$, the Lipschitz parameter is at most 4 and the global minimum is inside the unit ball, satisfying the requirements in the theorem statement. If $T < d^2$, then the bound cannot be better than what we would obtain for $T = d^2$ (the argument is similar to the one in the proof of Thm. 6), which is 0.004. Thus, for any T, the bound is at least

$$\min\left\{0.004, 0.004\sqrt{\frac{d^2}{T}}\right\} = 0.004\min\left\{1, \sqrt{\frac{d^2}{T}}\right\}$$

as required.