Lazy Queries Can Reduce Variance in Zeroth-Order Optimization

Quan Xiao, Qing Ling, Senior Member, IEEE, and Tianyi Chen

Abstract—A major challenge of applying zeroth-order (ZO) methods is the high query complexity, especially when queries are costly. We propose a novel gradient estimation technique for ZO methods based on adaptive lazy queries that we term as LAZO. Unlike the classic one-point or two-point gradient estimation methods, LAZO develops two alternative ways to check the usefulness of old queries from previous iterations, and then adaptively reuses them to construct the low-variance gradient estimates. We rigorously establish that through judiciously reusing the old queries, LAZO can reduce the variance of stochastic gradient estimates so that it not only saves queries per iteration but also achieves the regret bound for the symmetric two-point method. We evaluate the numerical performance of LAZO, and demonstrate the low-variance property and the performance gain of LAZO in both regret and query complexity relative to several existing ZO methods. The idea of LAZO is general and can be applied to other variants of ZO methods.

Index Terms—Query efficiency, zeroth-order optimization, machine learning.

I. INTRODUCTION

ZEROth-order (ZO) optimization (also known as gradient-free optimization, or bandit optimization), is useful in complex tasks when the analytical forms of loss functions are not available, only permitting evaluations of function values but not gradients. ZO methods have been applied to reinforcement learning [1], adversarial learning [2], [3], meta-learning [4], [5], hyperparameter optimization [6], [7], and Internet-of-Things [8]; see a survey [9]. However, a major challenge of applying ZO methods to benefit these practical problems is their high query complexity, especially when the queries are costly. To this end, we aim to develop a new ZO method that can inherit the merits of existing ones but also reduce the query complexity.

To illustrate our method, we consider the online convex optimization (OCO) setting [10], [11]. OCO can be viewed as a repeated game between a learner and the nature. Consider the time indexed by \( t \). Per iteration \( t \), a learner selects an action \( x_t \in \mathcal{X} \) and subsequently the nature chooses a loss function \( f_t \), through which the learner incurs a loss \( f_t(x_t) \). We consider the setting that, after the decision \( x_t \) is made, only the value of the loss function \( f_t \) at \( x_t \) is revealed; the gradient is unavailable. Our goal is to minimize the static regret

\[
\mathcal{R}_T(\mathcal{A}) \triangleq \mathbb{E} \left[ \sum_{t=0}^{T} f_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=0}^{T} f_t(x) \right]
\]

(1)

where \( \mathcal{A} \) is an algorithm, \( \mathcal{X} \subseteq \mathbb{R}^d \) is a convex set, \( \{f_t\} \) are convex functions and the expectation is taken over the random queries. This is a common performance measure of OCO, which measures the difference between the cumulative loss of online decisions generated by an algorithm \( \mathcal{A} \) and the loss of applying the best fixed decision chosen in hindsight [12].

While the loss function (or its gradient) is unknown, ZO methods approximate the gradient through values of the loss function based on various gradient estimation techniques. The work of [13] has first applied a ZO method to the OCO problem by querying a single function value at each time (so-called one-point gradient estimate). Going beyond one-point ZO [14], [15], [16], [17] leveraged multiple queries at each iteration (so-called multi-point gradient estimate) to achieve an enhanced performance (i.e., regret) relative to the one-point methods. Indeed, multiple queries at a single iteration are available in many applications, such as black-box adversarial attack [18], online linear regression [19], online advertising auction [20], online boosting [21], supervised page-rank learning [22] and collaborative target tracking [23]. For more applications, see [24] and references therein.

However, there is an essential trade-off between the number of queries per iteration and the number of iterations needed to achieve a target accuracy. The target accuracy is often measured by the average regret per iteration. Intuitively, querying fewer points per iteration may increase the variance of gradient estimation, and thus boost the number of iterations needed to achieve a target average regret [14], [17]. Consequently, it may increase the query complexity that depends on both of the two quantities. In this context, a natural yet critical question is Can we develop ZO methods that save the number of queries per iteration without sacrificing regret? It seems counter-intuitive that such a method does exist, achieving the best of both worlds. Nevertheless, under the conditions that (i) the algorithm has access to a two-point query model whenever it becomes necessary; and, (ii) an algorithm-dependency regularity assumption is satisfied, we provide an
affirmative answer in this article. The key idea that we leverage is to adaptively use the delayed queries that we call lazy queries to construct a low-variance gradient estimator.

A. Related Works

To put our work in context, we review prior contributions that we group in the three categories.

**One-point ZO.** ZO methods based on one-point gradient estimation can be traced back to the control literature, where one of the early approaches is the simultaneous perturbation stochastic approximation method [25]. In the context of OCO, algorithms with one-point feedback have been developed in [13], [26]. Building upon this element and the effective one-point gradient estimation scheme, the multi-agent ZO has developed for a game-theoretic model in [27]. See also a recent survey [28] and references therein. For all the aforementioned one-point ZO methods, the regret bound is still $O(T^{3/2})$, which is much worse than the $O(T^{3})$ regret bound of their full information counterparts [10]. The one-point ZO methods achieving the $O(T^{3})$ regret are the kernel-based [29] and ellipsoid method [30], but they use rather sophisticated gradient estimation techniques and have $O(d^{3.5})$ and $O(d^{2.5})$ dimension dependence, respectively, making them less efficient.

**ZO with delayed feedback.** The recent literature on bandits with delayed feedback is also related to this work, e.g., [31], [32], [33], [34], [35], [36]. However, these methods passively receive delayed feedback, in which delays generally sublinearly increase the regret, while our method actively leverages delayed feedback and uses delay to save queries and reduce the regret. Prior-guided ZO [37] is also relevant, but they only consider the time-invariant case. The work most relevant to ours is the residual one-point ZO method [38], which non-adaptively augments the one-point query with a delayed query to construct a two-point gradient estimator. Theoretically, its regret order is still $O(T^{3/2})$, which is the same as the vanilla one-point ZO method. Very recently, a control-theoretical ZO approach has been developed in [39] that significantly improves the dimension dependence in the regret of residual one-point ZO method [38] in the time-invariant case via high-pass and low-pass filters. However, the suboptimal dependence on the time $T$ still remains. The idea of using delayed queries has also been used in saving communication resources in distributed learning [40], [41], but the LAG algorithm and the analysis there are very different from those in the present article.

**Multi-point ZO.** ZO methods based on two or multiple function value evaluations enjoy better $O(T^{2})$ regret. They have been independently developed from both online learning and optimization communities [14], [15], [16], [17], [42], [43]. Recent advances in this direction include the sign-based ZO method [44], ADMM-based ZO method [18], constrained ZO [45], [46], autoencoder-based ZO [47], ZO for distributed systems [48], [49], [50], [51], diffusion ZO [52] and adaptive sampling ZO by leveraging the inherent sparsity of the gradient [53], [54], [55]. For solving finite-sum optimization problems, variance-reduced ZO methods have also been developed by using various state-of-the-art variance reduction techniques [56].

Our work is relevant but complementing to these multi-point ZO works. In fact, they all suggest promising future directions by applying our lazy query-based gradient estimation technique to them. There are several recent works on sparse-aware query efficient ZO [53], [54], [55]. The merits of them lie in that they can demonstrate theoretical improvement on query complexity, provided that the gradients satisfy the exact sparsity assumptions as outlined in these articles. The merit of our work is that the utilization of low-variance estimators can benefit iteration complexity and can offer performance improvement without the sparsity assumption that is required in sparse-aware ZO works. This is because, unlike sparsity-aware ZO methods that use compression, LAZO does not degrade the quality of the gradient estimator.

B. Our Contributions

In this context, our article puts forward a new ZO gradient estimation method that leverages an adaptive condition to parsimoniously query the loss function at one or two points. Our contributions can be summarized as follows.

C1) We propose a lazy query-based ZO gradient estimation method that we term LAZO and also generalize it to the multi-point version. Different from the classic ZO methods, LAZO adaptively reuses the old yet informative queries from previous iterations to construct low-variance gradient estimates.

C2) We apply LAZO to the stochastic gradient descent (SGD) algorithm and obtain a new SGD algorithm. We rigorously establish the regret of LAZO-based SGD. Surprisingly, through judiciously reusing old queries, LAZO can reduce the variance of gradient estimation so that it not only saves queries per iteration but also achieves the $O(\sqrt{dT})$ regret.

C3) We evaluate the numerical performance of LAZO on various tasks and show that LAZO maintains low variance and has performance gain in terms of regret and query complexity relative to popular methods.

II. LAZY QUERY FOR ZEROTH-ORDER SGD

In this section, we present our new ZO method that adaptively queries the loss function at either one or two points in each iteration, which gives its name Lazy Zeroth-Order gradient (LAZO) method.

A. Preliminaries

Before we present our new algorithm, we review some basics of OCO and ZO. In the full information case, where all information of the loss function including the gradient is available, the “workhorse” OCO algorithm is the online gradient descent method [10], given by $x_{t+1} = \Pi_{\mathcal{X}}(x_{t} - \eta\nabla f_{t}(x_{t}))$, where $\Pi_{\mathcal{X}}$ denotes the projection on to $\mathcal{X}$ and $\eta > 0$ is the stepsize. In the partial information setting, we only have access to the value of the loss function $f_{t}$ at $x_{t}$.

Generically speaking, ZO gradient-based methods first query the function values at one or multiple perturbed points $\{x_{t} + \delta u_{t}\}$, where $u_{t}$ is the unit perturbation vector and $\delta > 0$ is a...
small perturbation factor and following [15], [17], we assume that one can query \( f_t \) at any \( x_t + \delta u_t \), construct a stochastic gradient estimate \( \hat{g}_t(x_t) \) using these function values (see exact forms of \( \hat{g}_t(x_t) \) in Section II-B); and then plug it into the gradient iteration [13]

\[
x_{t+1} = \Pi\mathcal{X} (x_t - \eta \hat{g}_t(x_t))
\]  

(2)

Different from the most commonly used SGD, the stochastic gradient estimate in ZO gradient-based methods is usually biased in the sense that \( \mathbb{E}[\hat{g}_t(x_t)] \neq \nabla f_t(x_t) \).

The rationale behind (2) is that \( \hat{g}_t(x_t) \) is an unbiased estimator for the gradient of a smoothed version of \( f_t \) at \( x_t \). Specifically, defining a smoothed version of \( f_t \) as \( f_{S,t}(x) \) \( \triangleq \) \( \mathbb{E}_{v_t \sim \mathcal{U}(\mathcal{B})} [f_t(x + \delta v_t)] \), where \( v_t \sim \mathcal{U}(\mathcal{B}) \) denotes the uniform sampling \( v_t \) from the unit ball \( \mathcal{B} = \{ x \in \mathbb{R}^d ||x|| \leq 1 \} \), we have

\[
\mathbb{E}[\hat{g}_t(x)] = \nabla f_{S,t}(x) \quad \text{and} \quad f_{S,t}(x) \approx f_t(x).
\]  

(3)

In OCO, we make the following basic assumptions. 

Assumption 1: (Lipschitz continuity). For all \( t \), \( f_t(x) \) is \( L_t \)-Lipschitz continuous, i.e., \( \forall x, y \in \mathcal{X} , |f_t(x) - f_t(y)| \leq L_t ||x - y|| \). Moreover, we define \( L \triangleq \max_{t=0, \ldots, T} L_t \).

Assumption 2: (Bounded set). There exist constants \( R \) such that \( \mathcal{X} \subset \mathbb{R}^d \).

Assumption 3: (Convexity). For all \( t \), \( f_t(x) \) is convex.

Assumptions 1–3 are common in OCO with both full and partial information feedback [11], [13], [14], [17].

B. Observation: A Delicate Trade-Off

For the biased SGD iteration (2), the performance has been well-studied in literature. To present the connection between the regret and the quality of gradient estimation, we need the following lemma, the proof of which is in Section VI-A.

Lemma 1: If Assumptions 1–3 hold, and running the biased SGD (BSGD) iteration (2) with a generic \( \hat{g}_t(x_t) \) satisfying (3), then for any \( x \in \mathcal{X} \), we have that

\[
\mathbb{E}[\mathcal{R}_T(\text{BSGD})] \leq \frac{||x^* - x^*||^2}{2\eta} + 2L_T \delta T + \frac{T}{2} \sum_{t=0}^{T} \mathbb{E}[||\hat{g}_t(x_t)||^2]
\]

where \( x^* \in \arg\min_{x \in \mathcal{X}} \sum_{t=0}^{T} f_t(x) \).

The first term in the right-hand side (RHS) of (1) is the initial distance to \( x^* \); the second term shows the impact of the perturbation \( \delta \), which is due to the bias of the gradient estimator (cf. (3)); and the third term relates to both the bias and variance of \( \hat{g}_t \). Lemma 1 implies that the regret of (2) relative to the best fixed decision \( x^* \) critically depends on the second moment bound of \( \hat{g}_t(x_t) \).

Second moment of the gradient estimator \( \hat{g}_t(x_t) \). We discuss the second moment bounds with respect to the one-point [13], one-point residual [38] and two-point ZO methods [16].

C1) The classic one-point gradient estimator [13] can be written as

\[
\hat{g}_t^{(0)}(x_t) \triangleq \frac{du_t}{\delta} f_t(x_t + \delta u_t)
\]  

(4)

1Otherwise, a standard technique in [13], [14] that running the algorithm on a smaller set \( (1 - \gamma)\mathcal{X} = \{(1 - \gamma)x : x \in \mathcal{X}\} \) can be applied since choosing \( \gamma \) sufficiently small guarantees \( x_t + \delta u_t \in \mathcal{X} \).

where \( d \) is the dimension of \( x_t, u_t \sim U(\mathcal{S}) \) is a random vector from the unit sphere \( \mathcal{S} = \{ x \in \mathbb{R}^d ||x|| = 1 \} \) centered around the origin. Its second moment satisfies \( \mathbb{E} \left[ ||\hat{g}_t^{(0)}(x_t)||^2 \right] \leq \frac{d^2 G^2}{\delta^2} \), where \( G \) is defined as \( G \triangleq \max_{x \in \mathcal{X}} \max_{t} f_t(x) \).

C2) With \( f_{-1} = 0 \), the one-point residual gradient estimator in [38] can be written as

\[
\hat{g}_t^{(1)}(x_t) \triangleq \frac{du_t}{\delta} (f_t(x_t + \delta u_t) - f_{t-1}(x_t - \delta u_{t-1})).
\]  

(5)

Define the variation as \( V_t \triangleq \max_{x \in \mathcal{X}} |f_t(x) - f_{t-1}(x)| \). The second moment is bounded by

\[
\mathbb{E}\left[ ||\hat{g}_t^{(1)}(x_t)||^2 \right] \leq \max \left\{ ||\hat{g}_t^{(0)}(x_0)||^2, d^2 L^2 + \frac{d^2 V_t^2}{\delta^2} \right\}
\]  

(6)

where Lipschitz constant \( L \) is defined in Assumption 1.

C3) The asymmetric two-point gradient estimator and its second moment are respectively [16]

\[
\hat{g}_t^{(2)}(x_t) \triangleq \frac{du_t}{\delta} (f_t(x_t + \delta u_t) - f_t(x_t))
\]

\[
\mathbb{E}\left[ ||\hat{g}_t^{(2)}(x_t)||^2 \right] \leq d^2 L^2
\]  

(7)

and, the symmetric two-point gradient estimator and its bound on the second moment are respectively [17]

\[
\hat{g}_t^{(2)}(x_t) \triangleq \frac{du_t}{2\delta} (f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t))
\]

\[
\mathbb{E}\left[ ||\hat{g}_t^{(2)}(x_t)||^2 \right] \leq dL^2.
\]  

(8)

Performance trade-off. Plugging the bounds C1)–C3) into Lemma 1, we observe that the parameter \( \delta \) would play a crucial role on the regret. If we use more queries per iteration (e.g., two-point ZO in C3)), \( \delta \) does not appear in the bound of \( \mathbb{E}\left[ ||\hat{g}_t^{(1)}(x_t)||^2 \right] \), and one can simply choose an arbitrarily small \( \delta \) to minimize the regret bound. Hence, two-point ZO methods can reach \( O(T^{3/2}) \) regret [16], [17]. On the other hand, if we use fewer queries per iteration (e.g., one-point ZO in C1)–C2)), \( \delta \) does appear in the denominator of the bound on \( \mathbb{E}\left[ ||\hat{g}_t^{(2)}(x_t)||^2 \right] \). A trade-off thus emerges between the bias and variance of the gradient estimator since reducing \( \delta \) (e.g., bias) will increase the variance bound on \( \mathbb{E}\left[ ||\hat{g}_t^{(2)}(x_t)||^2 \right] \) in (1). Thus, the one-point ZO methods [13], [38] only achieve \( O(T^{2}) \) regret.

C. Key Idea: Two Lazy Query Rules

Motivated by this delicate trade-off, we will develop a low-variance ZO method that achieves the best of one- and two-point ZO. Our key idea is to use the adaptive combination of the one-point residual and two-point gradient estimators. In this way, the algorithm will query new points only when one point estimator suffers high variance.

Recalling the second moment bound of the one-point residual method in (6), we notice the degrading term is \( \frac{d^2 V_t^2}{\delta^2} \), so that the regret bound can approach \( O(T^{3/2}) \) correspondingly. We gauge that the requirement on \( V_t \) is stringent because it characterizes the maximum function changes at all points and iterations. Intuitively, we can still reuse old queries when the function variation at particular point and iteration is small.
A valuable revisit. We carefully re-analyze the second moment bound for the one-point residual estimator \( g_t^{(1)}(x_t) \) in (5), and get the following instance-dependent bound
\[
\|\hat{g}_t^{(1)}(x_t)\|^2 \leq \frac{f_t(w_t) - f_{t-1}(w_{t-1})}{\|w_t - w_{t-1}\|^2} \times \left(8d^2 + \frac{2d^3 \eta^2}{\delta^2}\right) ||\hat{g}_{t-1}(x_{t-1})||^2
\]
(9)
where the perturbed point is defined as \( w_t \triangleq x_t + \delta u_t \).

The proof of the instance-dependent bound (9) can be found in Section VI-B. The bound implies that if the instance-wise function value variation \( |f_t(w_t) - f_{t-1}(w_{t-1})| \) is small relative to \( \|w_t - w_{t-1}\| \), the second moment bound of \( \hat{g}_t^{(1)}(x_t) \) can be even smaller than the bound for the two-point ZO method in (7). We substantiate this instance-dependent analysis next.

Lemma 2: (Reduced norms). Under Assumptions 1–2, we run (2) with \( \hat{g}_t^{(1)}(x_t) \) and set \( \eta = \frac{R}{\sqrt{d}} \), \( \delta = R \sqrt{\frac{T}{d}} \). For a given iteration \( t \), if \( ||\hat{g}_t^{(1)}(x_{t-1})||^2 \leq d^2 L^2 \) and
\[
\frac{|f_t(w_t) - f_{t-1}(w_{t-1})|^2}{\|w_t - w_{t-1}\|^2} \leq \frac{L^2}{10d}
\]
then the second moment of the one-point residual gradient estimator (5) satisfies \( ||\hat{g}_t^{(1)}(x_t)||^2 \leq dL^2 \).

We present the proof of Lemma 2 in Section VI-C. Compared with the second moment bound in (6) and (8), Lemma 2 quantitatively shows that if the temporal variation, in the sense of (10), is small relative to \( L^2/d \), then reusing the delayed query can reduce the variance, and thus benefit the regret. Inspired by this, we introduce the following two definitions of the temporal variation.

Definition 1: (Temporal variation). For \( \forall x, y \in \mathcal{X} \), define the two temporal variation at \( t \) as
\[
D_t^x(x, y) \triangleq \frac{|f_t(x) - f_{t-1}(y)|}{\|x - y\|} \\
D_t^y(x, y) \triangleq \frac{|f_t(x) - f_{t-1}(y)|}{\eta L}.
\]
(11)

The two definitions in Definition 1 differ in that: the temporal variation \( D_t^x(x, y) \) normalizes the function values by the variation in terms of the query points, and the temporal variation \( D_t^y(x, y) \) approximates the variation in terms of the query points by the stepsize \( \eta \) since \( \|w_t - w_{t-1}\| \approx O(\eta L) \).

Building upon these two definitions, we design two alternative rules that check whether the temporal variations exceed a threshold \( D \) to decide whether to query one or two points. Specifically, using the same ZO iteration (2), we propose the LAZOa/b gradient estimator as
\[
\tilde{g}_t^{a/b}(x_t) = \begin{cases} 
\hat{g}_t^{(1)}(x_t), & \text{if } D_t^{a/b}(w_t, w_{t-1}) \leq D \\
\hat{g}_t^{(2)}(x_t), & \text{else,}
\end{cases}
\]
(12)
We summarize the complete LAZOa and LAZOB algorithms in Algorithm 1. In practice, we use the equivalent rule \( ||\hat{g}_t^{(1)}(x_t)|| \leq D := \frac{D_{\text{max}} L}{\delta} \) of LAZOB instead of directly computing \( D_t^x \), and treat \( \tilde{D} \) for LAZOa and \( \hat{D} \) for LAZOB as the hyperparameter tuned by grid search.

Algorithm 1 LAZO: Lazy query for ZO gradient method:

1. Input: \( x_0 \in \mathbb{R}^d \); \( T, \delta, \eta, \gamma, D > 0 \).
2. Sample \( u_0 \sim U(\mathcal{S}) \).
3. Query \( f_0(w_0) \) and \( f_0(x_0 - \delta u_0) \); \( \triangleright w_0 = x_0 + \delta u_0 \)
4. \( g_0(x_0) = \frac{d_{wu}}{\delta} (f_0(w_0) - f_t(x_0 - \delta u_0)) \)
5. for \( t = 1 \) to \( T \) do
6. Sample \( u_t \sim U(\mathcal{S}) \).
7. Query \( f_t(w_t) \); \( \triangleright w_t = x_t + \delta u_t \)
8. Compute \( \hat{g}_t(x_t) = \frac{d_{wu}}{\delta} (f_t(w_t) - f_{t-1}(w_{t-1})) \)
9. if \( D_t^x(w_t, w_{t-1}) \) or \( D_t^x(w_t, w_{t-1}) > \hat{D} \) then
10. Query \( f_t(w_t - \delta u_t) \).
11. \( \hat{g}_t(x_t) = \frac{d_{wu}}{\delta} (f_t(w) - f_t(x_t - \delta u_t)) \)
12. end if
13. Update \( x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta \hat{g}_t(x_t)) \).
14. end for

LAZOa and LAZOB have trade-off between performance and computation. As defined in Definition 1, \( D_t^x \) can be viewed as an approximation to \( D_t^y \), so LAZOa performs better through a more accurate temporal variation estimation. This is also shown in the experiments where LAZOa has faster convergence or achieves lower loss given a fixed iteration \( T \). However, LAZOB has lower computation overhead, which will be shown in Section V.

III. THEORETICAL ANALYSIS

This section provides theoretical guarantee for LAZO. Before we proceed, we first highlight the technical challenge of analyzing the regret of LAZO.

A. Challenge: Lazy Query Introduces Bias

In a high level, the difficulty comes from that the lazy query breaks down the unbiased property of the ZO gradient estimator. For simplicity, we use the LAZOB estimator as an example to illustrate this, and the same argument also holds for the LAZOa estimator.

Let \( A_t \triangleq \{u_t | D_t^x(w_t, w_{t-1}) \leq D \} \) denote the region where LAZOB uses the one-point query and \( A_t \) denote the complementary set of \( A_t \). To see the potential bias, we condition on the iterate \( x_t \) and take the conditional expectation of the LAZOB estimator, given by
\[
E_{u_t} \left[ \tilde{g}_t^{(1)}(x_t) \right] x_t \\
= E_{u_t} \left[ \hat{g}_t^{(1)}(x_t) 1_{A_t} + \hat{g}_t^{(2)}(x_t)1_{A_t} \right] x_t \\
\begin{align}
\triangleq E_{u_t} \left[ \hat{g}_t^{(1)}(x_t) 1_{A_t} + \hat{g}_t^{(2)}(x_t)1_{A_t} \right] x_t \\
+ E_{u_t} \left[ \hat{g}_t^{(1)}(x_t) 1_{A_t} - \hat{g}_t^{(2)}(x_t)1_{A_t} \right] x_t \\
= E_{u_t} \left[ \hat{g}_t^{(1)}(x_t) 1_{A_t} + \hat{g}_t^{(2)}(x_t)1_{A_t} \right] x_t \\
+ E_{u_t} \left[ \hat{g}_t^{(1)}(x_t) - \hat{g}_t^{(2)}(x_t) \right] 1_{A_t} x_t \\
= \mathbb{E}_{u_t} \left[ \left( \hat{g}_t^{(1)}(x_t) - \hat{g}_t^{(2)}(x_t) \right) 1_{A_t} \right] x_t \\
\end{align}
\]
(13)
where (a) holds since $A^e_t \subseteq A_t$ denotes the largest symmetric subset of $A_t$, i.e., $A^e_t = \text{sup}\{A | A|_{x_t}(u) = 1_{A|_{x_t}}(-u), A \subseteq A_t\}$ so that $1_{A^e_t} = 1_{A^e_t} + 1_{A^e_t \setminus A^e_t}$ and $A^e_t$ denotes the complementary set of $A^e_t$ so that $A_t \subseteq A^e_t$ and $\tilde{A}_t = 1_{A^e_t} - 1_{A^e_t \setminus A^e_t}$; (b) is due to $1_{A^e_t \setminus A_t} = 1_{A_t \setminus A^e_t}$.

For the two terms in (13), taking expectations of $\bar{g}_t^{(1)}(x_t)$ and $\bar{g}_t^{(2)}(x_t)$ over a symmetric distribution will give $\nabla f_{\delta,t}(x_t)$, and the remaining term will be treated as the bias; see details in our online version [59, Appendix A.4.1]. This way of decomposition ensures that when the region of one-point query is symmetric, the bias will be none; and the bias will diminish as $A_t$ is close to a symmetric set. This type of asymmetric bias also emerges in gradient clipping [57] and signSGD [58], where a symmetric gradient distribution is often assumed. We make a similar but weaker assumption in ZO below.

**Assumption 4:** Denote $A_t$ as the indicator function and $A|x_t$ as $A$ conditioned on $x_t$. Let $A^e_t = \text{sup}\{A | A|_{x_t}(u) = 1_{A|_{x_t}}(-u), A \subseteq A_t\}$ be the maximum symmetric space in $A_t$. Assume that the asymmetric area satisfies $\sum_{t=1}^T \mathbb{P}(A_t \setminus A^e_t) = O(\sqrt{T/d})$.

Assumption 4 is satisfied even if the summation of the probability series does not converge. For example, if $\mathbb{P}(A_t \setminus A^e_t) \sim O(\frac{1}{t^2})$, then $\sum_{t=1}^T \mathbb{P}(A_t \setminus A^e_t) \sim O(\log T) < O(\sqrt{T/d})$. Similarly, we can denote $A^e_t$ to be the region triggering one-point in LAZOa, that is $A^e_t \triangleq \{u_t | D^2_t(w_t, w_{t-1}) \leq D\}$ and replace $A_t$ in Assumption 4 by $A^e_t$.

**Discussion on symmetry.** Assumption 4 indicates that the active region of lazy rule $A_t$ is asymmetrically symmetric, which turns out to nearly hold throughout our simulations. Since projection preserves symmetry of a symmetric space. We gauge that if for all random projections, the projected areas are symmetric, the original $A_t$ is symmetric. We verify this via the linear quadratic regulator (LQR) and resource allocation experiment in Section V and show in Fig. 1. We randomly choose an iteration for each experiment, sample $4 \times 10^4$ random perturbations $u_t$ and generate 4 random $d \times 2$ Gaussian matrices to project those $u_t \in A_t$ onto a random 2-dimensional space and visualize them, from which we can see that all of the 8 projections are almost symmetric. More visualization of the symmetric projections of both LAZOa and LAZOb at different iteration are included in our online version [59] and a similar phenomenon occurs. Besides illustrating the projection of $A_t$

$\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Iteration } t & 2/2 & 10/100 & 50/200 & 100/1000 \\
\hline
\hat{p}_t \text{ for LAZOa} & 0.02/0.26 & 0.01/0.15 & 0.01/0.04 & 0.00/0.02 \\
\hline
\end{array}$

In Fig. 1, we have also provided the estimate of asymmetric probability $\hat{p}_t = \sum_{i=1}^N 1\{u_i^t \in A_t\}/N$ in Table I. We can see that the asymmetric rate estimator for LAZOa decays quicker while both LAZOa and LAZOb are almost symmetric at around 1000 iteration. Since the number of total iterations is 10000, the decaying speed for LAZOa/b is fast.

**B. Results in Online Convex Optimization**

We first verify that the LAZO estimator is an asymptotically unbiased gradient estimator of the smoothed function $f_{\delta,t}$. The proof of results in this subsection are presented in Sections VI-D and VI-E.

**Lemma 3:** Under Assumptions 1–4, the bias of the LAZO gradient estimator

$$
\mathbb{E} \left[ \tilde{g}_t^a(x_t) | x_t \right] = \nabla f_{\delta,t}(x_t) + \mathbb{E} \left[ b_t | x_t \right]
$$

and $\mathbb{E} \left[ \tilde{g}_t^b(x_t) | x_t \right] = \nabla f_{\delta,t}(x_t) + \mathbb{E} \left[ b_t | x_t \right]
$$

(14)

where $b_t$ is defined in (13), while $b_t$ is defined similarly by replacing $A_t, A^e_t$ with $\tilde{A}_t \triangleq \{u_t | D^2_t(w_t, w_{t-1}) \leq D\}, A^e_t = \text{sup}\{A | A|_{x_t}(u) = 1_{A|_{x_t}}(-u), A \subseteq A_t\}$. Moreover, if we use $\eta = \frac{R}{L \sqrt{d t}}$, $\delta = R \sqrt{\frac{d}{T}}$, $D = O\left(\frac{b_t}{\sqrt{d}}\right) < \frac{b_t}{\sqrt{dT}}$ for LAZOa and $D = O\left(\sqrt{dL}\right)$ for LAZOb, then both $\sum_{t=0}^T \mathbb{E} \left[ ||b_t|| \right] = O\left(\sqrt{d T}D\right)$ and $\sum_{t\leq t=0}^T \mathbb{E} \left[ ||b_t|| \right] = O\left(\sqrt{d T}D\right)$.

Then we establish the second moment bound of the gradient estimator in LAZO.

**Lemma 4:** Under Assumptions 1–3, the second moment bound of the gradient estimator $\bar{g}_t^a(x_t)$ and $\bar{g}_t^b(x_t)$ satisfy that there exists a constant $c = O(1)$ such that

$$
\mathbb{E} \left[ ||\tilde{g}_t^a(x_t) ||^2 | x_t \right] \leq c d L^2 + \frac{4d^2 D^2 c^2}{\delta^2} ||\tilde{g}_{t-1}(x_{t-1})||^2 + 8d^2 D^2;
$$

$$
\mathbb{E} \left[ ||\tilde{g}_t^b(x_t) ||^2 | x_t \right] \leq \frac{D^2 c^2 d L^2}{\delta^2} + c d L^2.
$$

(15)

In (15), one can show that by properly choosing $D$ and $\eta$, the second moment bound is $O(d)$, which is better than the bound of the one-point residual estimator (6) and the asymmetric two-point estimator (7) in terms of the $d$-dependence. In addition, one can choose $\eta, \delta = O\left(1/\sqrt{T}\right)$, so comparing with the one-point and one-point residual estimator, the dependence on the term $d^{-2}$ will be cancelled out in (15).

With the above two lemmas and the biased variant of Lemma 1, we can get the following regret bound of LAZO.

**Theorem 1:** (LAZO in the convex case). Under Assumptions 1–3, we run LAZO for $T$ iterations with $\eta = \frac{R}{L \sqrt{d t}}$, $\delta = R \sqrt{\frac{d}{T}}$. With $D = O\left(\frac{b_t}{\sqrt{d}}\right) < \frac{b_t}{\sqrt{dT}}$ for LAZOa and $D = O\left(\sqrt{dL}\right)$ for LAZOb, the regrets for LAZOa and LAZOb satisfy

---

*Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.*
Moreover, when Assumption 4 is satisfied, the regrets for LAZOa and LAZOb satisfy

\[ E[R_T(\text{LAZOa})] = O(\sqrt{dT}), \quad E[R_T(\text{LAZOb})] = O(\sqrt{dT}). \]

Theorem 1 states that thanks to the δ-independent variance in (15), the regret bound of LAZOa and LAZOb is \( O(\sqrt{dT}) \) with a bias term that is bounded (Table I shows it is small). Moreover, if the Assumption 4 holds, both LAZOa and LAZOb can achieve \( O(\sqrt{dT}) \) regret, which improves the \( O(T^{3/4}) \) and \( O(d\sqrt{V}) \) regret bounds of one-point (residual) methods in [13], [38], and the asymmetric two-point gradient estimator in [16], respectively. It also matches the optimal \( O(\sqrt{dT}) \) regret bound for the symmetric two-point ZO method in [17].

C. Results in Nonconvex Stochastic Optimization

LAZO can be also applied to the stochastic optimization setting. Due to the popularity of nonconvex learning applications, we present the results of the nonconvex setting.

Consider the function \( F(x; \xi) \) that depends on the random variable \( \xi \) and the stochastic problem \( \min_{x \in \mathbb{R}^d} F(x; \xi) \). In this case, instead of minimizing the regret (1), the goal is to minimize the average gradient norm as \( R_{T,F}(A) \) \( \sum_{t=0}^{T} E[||\nabla f(x_t)||^2] \).

Regarding algorithms, we can still implement LAZO in Algorithm 1 by replacing \( f_t(x) = F(x, \xi_t) \) and leave out the projection step in the projected ZO gradient descent since the feasible set is \( \mathbb{R}^d \). We make the following assumptions in additional to Assumption 1 in the OCO setting.

Assumption 5: Assume that \( f(x) \) is \( \mu \)-smooth, i.e. \( \forall x, y \in \mathcal{X}, ||\nabla f(x) - \nabla f(y)|| \leq \mu ||x - y|| \).

Theorem 2: (LAZO in the nonconvex case). Under Assumptions 1 and 5, we run LAZO for \( T \) iterations with \( \eta = \frac{1}{\sqrt{dT}}, \delta = \sqrt{\frac{\mu}{T}} \) and use \( D = O(\frac{L}{\sqrt{d}}) \) for LAZOa and \( D = O(\sqrt{dT}) \) for LAZOb. The regrets for LAZOa and LAZOb satisfy

\[ E[R_{T,F}(\text{LAZOa})] = O(\sqrt{dT} + d\sum_{t=0}^{T} E[\left\langle \tilde{A}_t, \nabla f_t \right\rangle]) \]
\[ E[R_{T,F}(\text{LAZOb})] = O(\sqrt{dT} + d\sum_{t=0}^{T} E[\left\langle A_t, \nabla f_t \right\rangle]) \]

Moreover, when Assumption 4 is satisfied, the regrets for LAZOa and LAZOb satisfy

\[ E[R_{T,F}(\text{LAZOa})] \leq O(\sqrt{dT}), E[R_{T,F}(\text{LAZOb})] \leq O(\sqrt{dT}). \]

Theorem 2 shows that the regret bound of LAZOa and LAZOb is \( O(\sqrt{dT}) \) with a bounded bias term (Table I shows it is small). Moreover, under Assumption 4, to achieve \( E[R_{T,F}(\text{LAZO})]/T \leq \varepsilon \), LAZOa and LAZOb require \( O(d\varepsilon^{-2}) \) iterations, and the iteration complexity is the same as that of the two-point ZO method in [43]. The proof of Theorem 2 and the result of convex case can be found in our online version [59, Appendixes B and C].

IV. EXTENSION TO MULTI-POINT QUERY RULES

To further reduce the variance of gradient estimation, we can construct a \( 2K \)-point LAZO gradient estimator with \( K > 1 \), akin to [14], [17]. In this case, we can apply our idea of lazy queries to existing multi-point ZO methods by expanding the reusing horizon from one round to \( H \) rounds with \( H > 1 \) to benefit the query complexity.

A. Algorithm Development

Correspondingly, the temporal variation in Definition 1 need to be extended to multiple previous horizons case.

Definition 2: (Temporal variation). For \( \forall x, y \in \mathcal{X}, t \in \mathbb{N} \), define the two temporal variation between \( t \) and \( t - \tau \) as

\[ D^a_{t,\tau}(x,y) = \frac{||f_t(x) - f_{t-\tau}(y)||}{||x - y||} \]
\[ D^b_{t,\tau}(x,y) = \frac{||f_t(x) - f_{t-\tau}(y)||}{\eta L} \] (16)

Likewise, the value of two temporal variation in Definition 2 can be served as the indicator of whether the queries at time \( t - \tau \) is informative.

Moreover, in [17], two-point symmetric ZO gradient estimator can be extended to \( 2K \)-points estimator to reduce estimation variance as follows (\( K > 1 \)):

\[ g_{t}^{(2K)}(x_t) = \frac{d}{2K} \sum_{k=1}^{K} u^k_t(f_t(x_t + \delta u^k_t) - f_t(x_t - \delta u^k_t)) \] (17)

where \( u^k_t \) are i.i.d randomly sampled from the unit sphere. We can also use the same idea in LAZO to construct multiple points gradient estimator for robustness.

To attain a \( 2K \)-points LAZO estimator using queries in \( H > 1 \) previous steps, at time \( t \), we first search on the previous \( H \) steps for valuable queries and reuse them, or if no old query is useful, we query a new point. Continuing this procedure until we obtain a \( 2K \)-points estimator. Specifically, we can define \( 2K \)-points LAZOa estimator as follows:

\[ g_{t}^{(2K)}(x_t) = \frac{d}{\delta K} \sum_{k=1}^{K} u^k_t \left\{ \frac{1}{2} T_{t,k} = \emptyset \left( f_t(u^k_t) - f_t(x_t - \delta u^k_t) \right) \right\} + \sum_{(\ell, t) \in T_{t,k}} \left( f_t(u^k_t) - f_{t-\ell}(u^k_{t-\ell}) \right) \] (18)

where \( u^k_t \) and \( T_{t,k} \) are defined as in (17) and (18), respectively.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
and $K^b_t$ is the integer such that $\sum_{k=1}^{K^b_t} \left( |T_{t,k}^b| + 1 \right) = K$.

The set $T_{t,k}^{a,b}$ only contains $K - \sum_{k=1}^{K^b_t-1} \left( |T_{t,k}^b| + 1 \right)$ elements if its size is bigger.

Similarly, the 2$K$-points LAZOb gradient estimator as

$$g_t^2(x_t) = \frac{d}{dK} \sum_{k=1}^{K^b_t} u_t^k \left\{ \frac{1}{2} \left( f_t(w_t^k) - f_t(x_t - \delta u_t^k) \right) + 1 \right\}$$

where

$$T_{t,k}^b \triangleq \{ (\tau,l) | 1 \leq \tau \leq H, 1 \leq l \leq K^b_{t-\tau}, D_{t,\tau}^b(w_t^k, w_{t-\tau}^l) \leq D \}$$

and $K^b_t$ is the integer such that $\sum_{k=1}^{K^b_t} \left( |T_{t,k}^b| + 1 \right) = K$.

Again if the set $T_{t,k}^{a,b}$ only contains $K - \sum_{k=1}^{K^b_t-1} \left( |T_{t,k}^b| + 1 \right)$ elements if its size is bigger.

V. NUMERICAL EXPERIMENTS

In this section, we empirically evaluate the performance of our LAZO and its multi-point variant on three applications: LQR control, resource allocation and generation of adversarial examples from a black-box deep neural network (DNN).

Throughout this section, we compare LAZO with one-point residual algorithm [38] and two-point ZO gradient descent [17].

The choice of parameters and other details of experiments are presented in [59, Appendix D].

A. Non-Stationary LQR Control

We study a non-stationary version of the classic LQR problem [60] with the time-varying dynamics. At iteration $t$, consider the linear dynamic system described by the dynamic $x_{k+1} = A_t x_k + B_t q_k$, where $x_k \in \mathbb{R}^n$ is the state, $q_k \in \mathbb{R}^p$ is the control variable at step $k$, $A_t \in \mathbb{R}^{n \times n}$ and $B_t \in \mathbb{R}^{n \times p}$ are the dynamic matrices for iteration $t$. Our goal is to minimize the cost which is a fixed quadratic function of state and control given by

$$\min_{K} \mathbb{E} \left[ \frac{1}{T} \sum_{k=1}^{T} (x_k^T Q x_k + q_k^T R q_k) \right]$$

where $\beta \in (0, 1)$ is a discount factor, $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{p \times p}$ are the positive definite matrices, and $H$ is the length of step horizon. We search the control $q_k = K_t x_k$ that linearly depends on the current state $x_k$, where $K_t^* \in \mathbb{R}^{p \times n}$ is the optimal policy in iteration $t$. Thus our optimization variable will be $K$ and the loss function will be

$$\min_{K} f_t(K) = \mathbb{E} \left[ \frac{1}{T} \sum_{k=1}^{T} \frac{1}{2} \sum_{k=1}^{H} \beta^k x_k^T (Q + K^T R K) x_k \right]$$

s.t. $x_{k+1} = (A_t + B_t K) x_k$.

We set $n = p = 6$, $\beta = 0.5$, $\delta = 0.01$, and stepsize $\eta = 10^{-5}$ for all four methods. We generate $A_t$, $B_t$ to mimic the situations where the loss functions encounter intermittent changes, given by

$$A_t = \begin{cases} s_t, & t \text{ mod } 100 = 0 \\ A_{t-1} + s_t + 7 \sin[t \text{ mod } 100], & t \text{ mod } 100 \in [35, 65] \\ A_{t-1} + s_t, & \text{else} \end{cases}$$

where the noise $s_t \sim \mathcal{N}(0, 1)$ if $t \text{ mod } 100 = 0$, and $s_t \sim \mathcal{N}(0, 0.12)$, otherwise. We generate $B_t$ the same way as $A_t$ but replace the sin function with the cos function.

We monitor the average temporal variation, gradient variance and average queries per iteration over 50 iterations and 100 trials in Fig. 2 for LAZoA. We observe that when the loss function varies slowly (e.g., $t \leq 35$), the temporal variation is also small and thus, the average query for LAZoA is relatively small; when the loss function changes rapidly (e.g., $t \geq 35$), the temporal variation is also large and as a result, LAZoA needs more average queries. Note that the actual upper bound of two-point ZO’s variance depends on $L$ and the way we generate $A_t$, $B_t$ will affect not only the temporal variation but also $L_t$, resulting in the change of variance of two-point ZO. Fig. 2 also indicates that thanks to the lazy query, the variance of the LAZoA gradient estimator keeps the lowest. In Fig. 4(a) and 4(e), we report the cost versus iteration and query of the four methods. Here we choose $D = 1$ for LAZoA and $D = 100$ for LAZoB to optimize the performance for them. In Fig. 4(a), LAZoA and LAZoB yield the best convergence performance and LAZoA has the smallest errorbar over random trials. Regarding query complexity in Fig. 4(e), LAZoI still outperforms the other two methods.

Besides, we provide the runtime comparison for one-point residual method, two-point method and LAZoA/b over the same number of iterations* for LQR control in Table II. In fact, we observed that LAZoA/b can even slightly save runtime compared to the two-point ZO method because querying function value usually takes more time than computing temporal variation. This shows the adaptive rules of LAZoA/b actually

Fig. 2. Monitoring the adaptive query condition in LQR at first 50 iterations. The upmost plot records the average temporal variation over 100 trials. The gradient variance condition for one-point residual method (blue), two-point method (green) and LAZoA (red) is shown in the middle plot. The lowermost plot shows the average queries in LAZoA.
TABLE II
RUNTIME FOR ONE-POINT RESIDUAL METHOD, TWO-POINT METHOD AND LAZOa/b in LQR CONTROL

| Method              | one-point residual | two-point   |
|---------------------|--------------------|-------------|
| Time (sec)          | 1548               | 1653        |
| Method              | LAZOa              | LAZOb       |
| Time (sec)          | 1651               | 1640        |

B. Non-Stationary Resource Allocation

We consider a resource allocation problem with 16 agents connected by a ring graph. For each iteration $t$, per step $k$, each node $i$ receives an exogenous data request $b^k_i = \psi_i \sin(\omega_i k + \phi_i)$, stores $y^k_i$ amount of resources and forwards $a_{ij}^k$ fraction of resources to its neighbor node $j \in \mathcal{N}_i$. Then the aggregate (endogenous plus exogenous) workload of each node $i$ evolves by

$$y^{k+1}_i = y^k - \sum_{j \in \mathcal{N}_i} a_{ij}^k y^k_j + \sum_{j \in \mathcal{N}_i} a_{ji}^k y^{k+1}_j - b^k_i.$$

Per iteration $t$, for each node $i$, at each step $k$, the power cost $r_{i,t}^k$ depends on a varying parameter $p^k_i$ as

$$r_{i,t}^k = \begin{cases} 0, & \text{if } y^k_i \geq 0 \\ p^k_i(y^k_i)^2, & \text{else.} \end{cases}$$  

Defining $x^t_k = [y^k_1, \ldots, y^k_n]^T$ and the policy $\pi^t_i(x^t_k, \theta_k) : x^t_k \rightarrow [0,1]^{\mathcal{N}_i}$ at iteration $t$, our goal is to find the optimal policy to allocate $a_{ij}^k$, and thus to minimize the instantaneous accumulated cost $f_t(\theta_k) = \sum_{i=1}^{16} \sum_{k=1}^{H} \beta^k r_{i,t}^k$, where $H$ is the time step length and $\beta$ is the discount factor. The time-varying parameter $p^t_i$ is generated according to $p^t_i = \sin(\frac{\pi}{16}) + s^t_i$ where $s^t_i$ is uniformly distributed over $[0,1]$.

Fig. 4(b) and 4(f) presents the cumulative cost versus iteration and query. LAZO outperforms the other methods in terms of iteration while both LAZOa and LAZOb improve the query complexity. Besides, we can see that both two LAZO variants have narrower errorbar compared with the other methods, and LAZOa is more stable than LAZOb over different trials.

C. Black-Box Adversarial Attacks

In the nonconvex stochastic setting, we study generating adversarial examples from an image classifier given by a black-box DNN on the MNIST dataset. The DNN model is seen as the zeroth-order oracle. Let $(y_0, l_0)$ be the image $y_0$ with true label $l_0 \in \{1, \ldots, n\}$ in $n$ different classes. Assume the target DNN classifier $H(y) = [H_1(y), \ldots, H_n(y)]$ is a well-trained classifier, where $H_i(y)$ means the probability of $y$ being class $i$. Given $H$, an adversarial example $y$ of $y_0$ means that it is visually similar to $y_0$ but $H$ gives a different prediction class to it. Since the pixel value range of images is always bounded, without loss of generality, we can assume $y \in [-0.5, 0.5]^d$. Since the black-box attack is the nonconvex stochastic problem, which we only have access to solve the unconstrained setting, we need to apply the tanh/2 transformation to an unbounded variable $x \in \mathbb{R}^d$ to represent $y$. Then we can adopt the black-box attacking loss function defined in [2], which is given by

$$\min_{x \in \mathbb{R}^d} \max\left\{ \log H_{l_0}(\text{tanh}(x)/2) - \max_{l \neq l_0} \log H_l(\text{tanh}(x)/2), 0 \right\} + \|\text{tanh}(x)/2 - y_0\|_2^2,$$

where the first term represents the maximum difference between the probability of being classified to the true class $l_0$ and the most possible predicted class other than $l_0$, the second term is the $l_2$ distortion, and $\beta$ is the penalty parameter. Here we choose $\beta = 0.5$.

In the experiment, since the one-point residual method is unstable, we choose a smaller stepsize $\eta = 0.1$ and bigger $\delta = 0.5$ for it to ensure stability. To show $\eta = 0.1$ is a reasonable choice for the one-point residual method, we also report the result when $\eta = 0.2$. For the other three methods, we pick the optimal $\eta \in \{0.1, 0.2, 0.3\}$ and $\delta \in \{0.5, 0.1, 0.05, 0.01\}$. Besides, we set $D = 1$ for LAZOa and $D = 1000$ for LAZOb.

Fig. 4(c) and 4(g) shows the attacking loss versus iteration and query for the four methods. LAZOa and LAZOb outperform one-point residual and two-point ZO methods in terms of both iteration and query. In addition, compared to two-point ZO methods, LAZOa and LAZOb only requires nearly 40% and 67% queries to achieve the first successful attack, which is the first iteration or query number when the $l_2$ distortion loss begins to drop [2], [44], respectively.

We also compare multi-point LAZO with multi-point ZO method [17] and the comparison results for $H = 3, K = 3$ setting are shown in Fig. 4(d) and 4(h). It can be seen that both multi-point LAZOa and multi-point LAZOb outperform multi-point ZO in terms of iteration complexity and query complexity. Moreover, multi-point LAZOa and LAZOb requires nearly 38% and 55% queries to achieve the first successful attack,
respectively, which further improve the results for $H = 1, K = 1$ in Fig. 4(c) and 4(g).

VI. PROOFS OF MAIN RESULTS

A. Proof of Lemma 1

Lemma 1 is a standard result of biased SGD. To be self-contained, we provide its proof here.

Proof: Since $f_{\delta,t}$ is convex for all $t$, we have that

$$
 f_{\delta,t}(x_t) - f_{\delta,t}(x) \leq \langle \nabla f_{\delta,t}(x_t), x_t - x \rangle, \forall x \in \mathcal{X}.
$$

Using $E[g_t(x_t)|x_t] = \nabla f_{\delta,t}(x_t)$ in (3) and taking expectation over both sides, we get for all $x \in \mathcal{X}$,

$$
 E[f_{\delta,t}(x_t) - f_{\delta,t}(x)] \leq E[(E[g_t(x_t)|x_t], x_t - x)]
 = E[E[(\tilde{g}_t(x_t), x_t - x)|x_t]]
 = E[(\tilde{g}_t(x_t), x_t - x)].
$$

Since $x_{t+1} = \Pi_\mathcal{X}[x_t - \eta\tilde{g}_t(x_t)]$, for any $x \in \mathcal{X}$ we have that

$$
 \|x_{t+1} - x\|^2 = \|\Pi_\mathcal{X}[x_t - \eta\tilde{g}_t(x_t)] - \Pi_\mathcal{X}[x]\|^2
 \leq \|x_t - \eta\tilde{g}_t(x_t) - x\|^2
 = \|x_t - x\|^2 - 2\eta\langle \tilde{g}_t(x_t), x_t - x \rangle
 + \frac{\eta^2}{2}\|\tilde{g}_t(x_t)\|^2.
$$

where the inequality follows the non-expansive property of the projection. Rearranging the terms in inequality (24) yields

$$
 \langle \tilde{g}_t(x_t), x_t - x \rangle \leq \frac{1}{2\eta} \left( \|x_t - x\|^2 - \|x_{t+1} - x\|^2 \right)
 + \frac{\eta}{2}\|\tilde{g}_t(x_t)\|^2.
$$

Taking expectations on both sides of inequality (25) with respect to $u_t$, substituting the resulting bound into (23) and summing from $t = 0$ to $T$, we obtain that

$$
 E \left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) \right]
 \leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}E \left[ \sum_{t=0}^{T} \|\tilde{g}_t(x_t)\|^2 \right].
$$

Since $f_t$ is Lipschitz for all $t$, we conclude that

$$
 |f_{\delta,t}(x) - f_t(x)| = |E_{v \in \mathcal{V}}[f_t(x + \delta v) - f_t(x)]| \leq LE\|v\| \leq \delta L.
$$

Then we obtain that

$$
 E \left[ \sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x) \right]
 = E \left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) + \sum_{t=0}^{T} (f_t(x_t) - f_{\delta,t}(x_t)) \right]
 \leq \frac{1}{2\eta}\|x_0 - x\|^2 + \frac{\eta}{2}E \left[ \sum_{t=0}^{T} \|\tilde{g}_t(x_t)\|^2 \right] + 2L\delta T.
$$

Then plugging $x^* = \arg \min_{x \in \mathcal{X}} \sum_{t=0}^{T} f_t(x)$ into (27) and according to the definition of regret in (1) yield the conclusion.

B. Derivation of (9)

Proof: From the definition of $\tilde{g}_t^{(1)}(x_t)$, we have
\[ \|g_t\|^2 = \frac{d^2}{\delta^2} (f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|u_t\|^2 \]
\[ \leq \frac{d^2}{\delta^2} (f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|w_t - w_{t-1}\|^2 \]
\[ \leq \frac{2d^2}{\delta^2} (f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|w_t - w_{t-1}\|^2 \]
\[ \times (\|x_t - x_{t-1}\|^2 + \delta^2 \|u_t - u_{t-1}\|^2) \]
\[ \leq \frac{2d^2}{\delta^2} (f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|w_t - w_{t-1}\|^2 \]
\[ \times (\eta^2 \|g_t\|^2 + 4\delta^2) \]
\[ = 2d^2(f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|w_t - w_{t-1}\|^2 \]
\[ \times \left( 4 + \frac{\eta^2}{\delta^2}\|g_t\|^2 \right) \tag{28} \]

where the first inequality is because \(\|u_t\| = 1\), the second inequality comes from the fact that
\[ \|w_t - w_{t-1}\|^2 = \|x_t - x_{t-1} + \delta(u_t - u_{t-1})\|^2 \]
\[ \leq \|x_t - x_{t-1}\|^2 + \delta^2 \|u_t - u_{t-1}\|^2 \] \tag{29}
while the third inequality follows from the update in (2) and the relation \(\|u_t - u_{t-1}\|^2 \leq 4\).

**C. Proof of Lemma 2**

**Proof:** With the choice of \(\eta\) and \(\delta\), we have that \(\frac{\eta^2}{\delta^2} = \frac{1}{\sigma T^2}\).

Then using (9), we obtain that
\[ \|g_t\|^2 \leq 2d^2(f_t(w_t) - f_{t-1}(w_{t-1}))^2 \|w_t - w_{t-1}\|^2 \]
\[ \times \left( 4 + \frac{\eta^2}{\delta^2}\|g_t\|^2 \right) \]
\[ \leq 2d^2 \frac{\|f_t(w_t) - f_{t-1}(w_{t-1})\|^2}{10d} \left( 4 + \frac{1}{\delta d^2}\|g_t\|^2 \right) \]
\[ \leq \frac{4}{5}dL^2 + \frac{1}{5d}\|g_t\|^2 \]
\[ \leq \frac{4}{5}dL^2 + \frac{1}{5d}L^2 = dL^2 \tag{30} \]

where the second inequality comes from the condition.

Next, we prove Lemma 3, Lemma 4 and Theorem 1 by dividing them into Section VI-D for LAZO\(b\) and Section VI-E for LAZO\(a\).

**D. LAZO\(b\) Estimator**

In this section, we present the proof for LAZO\(b\) estimator.

1) Bias for LAZO\(b\): To be self-contained, we restate the LAZO\(b\) part of Lemma 3 as follows.

**Restatement of Lemma 3:** (LAZO\(b\) estimator). Under Assumptions 1–4, the bias of the LAZO\(b\) gradient estimator is
\[ \mathbb{E}[g_t] = \nabla f_{\delta_t}(x_t) + \mathbb{E}[b_t] \tag{31} \]
where \(b_t\) is defined in (13). Moreover, if we use \(D = O(\sqrt{dL})\), then \(\sum_{t=0}^T \mathbb{E}[\|b_t\|] = O(\sqrt{dT})\).

**Proof:** Recall \(A_t \triangleq \{u_t | D_t(u_t, w_{t-1}) \leq D\} \) denote the region where LAZO\(b\) uses the one-point query and \(\bar{A}_t\) denote the complementary set of \(A_t\) where LAZO\(b\) uses the two-point query. We have
\[ \mathbb{E}[g_t] = \mathbb{E}\left[ \frac{d}{\delta_t} \frac{f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)}{2} \right] \]
\[ + \mathbb{E}\left[ \frac{d}{\delta_t} \frac{f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)}{2} \right] \]
\[ + \mathbb{E}\left[ \frac{d}{\delta_t} \frac{f_t(x_t - \delta u_t) - f_t(x_t + \delta u_t)}{2} \right] \]
\[ + \mathbb{E}\left[ \frac{d}{\delta_t} \frac{f_t(x_t - \delta u_t) - f_t(x_t + \delta u_t)}{2} \right] \]
\[ = \mathbb{E}\left[ \frac{d}{\delta_t} f_t(x_t + \delta u_t) \right] \]
\[ + \mathbb{E}\left[ \frac{d}{\delta_t} f_t(x_t - \delta u_t) \right] \]
\[ = \mathbb{E}\left[ \frac{d}{\delta_t} f_t(x_t + \delta u_t) \right] \]
\[ - \mathbb{E}\left[ \frac{d}{\delta_t} f_t(x_t - \delta u_t) \right] \]
\[ \in O(\sqrt{dL}) \tag{32} \]
where \(\bar{A}_t\) and \(A_t \setminus A_t^c\) denote the complementary set of \(A_t\), \(A_t^c\) and the difference of sets \(A_t\) and \(A_t^c\). The first two terms in the third inequality are due to the symmetricity of the sets \(A_t^c\) and \(A_t\).

From equation (32), we can get that
\[ \mathbb{E}[\|b_t\|] \leq d \mathbb{E}[|2f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)|] \]
\[ + \mathbb{E}[|f_t(x_t - \delta u_t) - f_t(x_t + \delta u_t)|] \]
\[ \leq \frac{d}{\delta} \mathbb{E}[|f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)|] \]
\[ + \mathbb{E}[|f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)|] \]
\[ \leq \left( \frac{d\eta dL}{\delta} + dL \right) \mathbb{P}(A_t \setminus A_t^c) \tag{33} \]
where the first inequality is derived from adding and subtracting \(f_t(x_t + \delta u_t)\); the second bias is because \(\mathbb{E}[|X + Y|] \leq \mathbb{E}[|X|] + \mathbb{E}[|Y|]\); the first term in the third inequality is given by the fact that if \(u \in A_t \setminus A_t^c \subseteq A_t\), then \(|f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)| \leq 2d\eta L\) and the second term is due to the Lipschitz condition.

Thus, plugging \(\eta = \frac{R}{2dL}, \delta = \frac{R}{\sqrt{d}}\), \(D = O(\sqrt{dL})\) and \(\sum_{t=0}^T \mathbb{P}(A_t \setminus A_t^c) = O(\sqrt{T} / \sqrt{d})\) to (33), we can get that \(\sum_{t=0}^T \mathbb{E}[\|b_t\|] = O(\sqrt{dT})\).
2) The Second Moment Bound for LAZOb: 

Lemma 5: ([17, Lemma 9]). For any function \( h : \mathbb{R}^d \to \mathbb{R} \) which is \( L \)-Lipschitz continuous, it holds that if \( u \in \mathbb{R}^d \) is uniformly distributed on the Euclidean unit sphere, then there exists a constant \( c = O(1) \) such that

\[
\sqrt{\mathbb{E} [(h(u) - \mathbb{E}[h(u)])^2]} \leq c \frac{L^2}{d}. \tag{34}
\]

Lemma 6: Under Assumption 1, for any symmetric set \( A \subseteq \mathbb{R}^d \) with respect to \( u \in \mathbb{R}^d \), any given \( x \in \mathbb{R}^d \) and any given \( t \), we have that

\[
\mathbb{E}_u \left[ \frac{d^2}{4d^2} (f_t(x + \delta u) - f_t(x - \delta u))^2 \mathbf{1}_A \right] \leq cdL^2 \sqrt{\mathbb{P}(A)}. \tag{35}
\]

Proof: First, it follows that for any \( \alpha \in \mathbb{R} \),

\[
\mathbb{E}_u \left[ \frac{d^2}{4d^2} (f_t(x + \delta u) - f_t(x - \delta u))^2 \mathbf{1}_A \right] = \mathbb{E}_u \left[ \frac{d^2}{4d^2} (f_t(x + \delta u) - \alpha - f_t(x - \delta u) + \alpha)^2 \mathbf{1}_A \right] \leq \frac{d^2}{2d^2} \mathbb{E}_u \left[ (f_t(x + \delta u) - \alpha)^2 + (f_t(x - \delta u) - \alpha)^2 \right] \mathbf{1}_A \leq \frac{d^2}{d^2} \mathbb{E}_u \left[ (f_t(x + \delta u) - \alpha)^2 \right] \mathbf{1}_A \leq \frac{d^2}{d^2} \sqrt{\mathbb{E}_u \left[ (f_t(x + \delta u) - \alpha)^2 \right]} \mathbf{1}_A \leq \frac{d^2}{d^2} \frac{\delta^2 L^2}{d} \mathbf{1}_A \leq \frac{d^2}{d^2} \frac{\delta^2 L^2}{d} \sqrt{\mathbb{P}(A)} = cdL^2 \sqrt{\mathbb{P}(A)}
\]

where the first inequality is due to \( (X + Y)^2 \leq 2(X^2 + Y^2) \); the second inequality is due to the symmetricity of \( u \) and the symmetricity of \( A \); the third inequality is because \( \mathbb{E}[XY] \leq \sqrt{\mathbb{E}[X^2] \mathbb{E}[Y^2]} \). If we choose \( \alpha = \mathbb{E}_u [f_t(x + \delta u)] \) and apply Lemma 5, then the last inequality holds since \( f_t(x + \delta u) \) is \( L \)-Lipschitz. This completes the proof. \( \square \)

To be self-contained, we restate the LAZOb part of Lemma 4 by the following lemma and prove it.

Restatement of Lemma 4: (LAZOb estimator). The second moment bound of the gradient estimator \( \tilde{g}_t^b(x_t) \) satisfies that there exists some constant \( c \) such that

\[
\mathbb{E} \left[ \| \tilde{g}_t^b(x_t) \|^2 \right] \leq \frac{D^2 \eta^2 \delta^2 L^2}{\delta^2} + cdL^2. \tag{36}
\]

Proof: Using the definition of LAZOb estimator in (12), we have

\[
\mathbb{E} \left[ \| \tilde{g}_t^b(x_t) \|^2 \right] = \mathbb{E} \left[ \frac{d^2}{d^2} (f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1}))^2 \mathbf{1}_A, x_t \right] + \mathbb{E} \left[ \frac{d^2}{d^2} (f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t))^2 \mathbf{1}_A, x_t \right] \leq \mathbb{E} \left[ \frac{d^2}{d^2} (f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1}))^2 \mathbf{1}_A, x_t \right] + \mathbb{E} \left[ \frac{d^2}{d^2} (f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t))^2 \mathbf{1}_A, x_t \right] \leq \frac{D^2 \eta^2 \delta^2 L^2}{\delta^2} + cdL^2 \tag{37}
\]

where the first inequality is due to \( \| u_t \| = 1 \) and \( \mathbf{1}_A \times \mathbf{1}_A = 0 \); the second term in the first inequality is due to \( \mathbf{1}_A; \leq 1 \); and the last term is derived from Lemma 6 and \( \mathbb{P}(A) \leq 1 \). \( \square \)

3) The Regret Bound for LAZOb: To be self-contained, we restate the LAZOb part of Theorem 1 as follows.

Restatement of Theorem 1: (LAZOb estimator). Under Assumptions 1–4, we run LAZOb for \( T \) iterations with \( \eta = \frac{R}{L \sqrt{dT}} \), and \( \delta = R \sqrt{\frac{1}{T}} \). If we use \( D = O(\sqrt{dL}) \) for LAZOb, then the regret for LAZOb satisfy

\[
\mathbb{E}[R_T(\text{LAZOb})] = O(\sqrt{dT}). \tag{38}
\]

Proof: Using \( \mathbb{E} [\tilde{g}_t^b(x_t)|x_t] = \nabla f_{\delta,t}(x_t) + \mathbb{E} [b_t|x_t] \) in (3) and taking expectation of both sides of (22), we get for all \( t \in \mathcal{N} \),

\[
\mathbb{E} [f_{\delta,t}(x_t) - f_{\delta,t}(x)] \leq \mathbb{E} \left[ \mathbb{E} \left[ \tilde{g}^b_t(x_t) - b_t[x_t], x_t - x \right] \right] = \mathbb{E} \left[ \mathbb{E} \left[ \tilde{g}^b_t(x_t) - b_t[x_t], x_t - x \right] \right] = \mathbb{E} \left[ \mathbb{E} \left[ \tilde{g}^b_t(x_t), x_t - x \right] \right] + \mathbb{E} \left[ \mathbb{E} \left[ b_t[x_t], x_t - x \right] \right] \leq \mathbb{E} \left[ \mathbb{E} \left[ \tilde{g}^b_t(x_t), x_t - x \right] \right] + 4 \left( \frac{D \eta L}{\delta} + dL \right) \mathbb{E} \left[ \mathbb{P}(A_t | A_t^*) \right]. \tag{39}
\]

Taking expectations on both sides of inequality (25) with respect to \( u_t \), substituting the resulting bound into (38) and summing from \( t = 0 \) to \( T \), we obtain that

\[
\mathbb{E} \left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) \right] \leq \frac{1}{2\eta} \| x_0 - x \|^2 + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=0}^{T} \| \tilde{g}^b_t(x_t) \|^2 \right] + 4 \left( \frac{D \eta L}{\delta} + dL \right) \mathbb{E} \left[ \mathbb{P}(A_t | A_t^*) \right]. \tag{39}
\]

Then similar to (27), we obtain that

\[
\mathbb{E} \left[ \sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x) \right] = \mathbb{E} \left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) + \sum_{t=0}^{T} (f_t(x_t) - f_{\delta,t}(x_t)) \right] - \sum_{t=0}^{T} (f_t(x) - f_{\delta,t}(x)) \leq \frac{1}{2\eta} \| x_0 - x \|^2 + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=0}^{T} \| \tilde{g}^b_t(x_t) \|^2 \right] + 4 \left( \frac{D \eta L}{\delta} + dL \right) \mathbb{E} \left[ \mathbb{P}(A_t | A_t^*) + 2L \delta T. \tag{40}
\]
Plugging $\eta = \frac{R}{L \sqrt{dT}}$, $\delta = R \sqrt{\frac{d}{T}}$ to the second moment bound in Lemma 4, then we can get that
\[
E \left[ \| \hat{g}_b^a(x_t) \|^2 \right] \leq D^2 + cdL^2. \tag{41}
\]
Then plugging $x^* = \arg \min_{x \in \mathcal{X}} \sum_{t=0}^{T} f_t(x)$ and the second moment bound for the $\hat{g}_b^a(x_t)$ into (50), and according to the definition of regret in (1), we can reach the conclusion that
\[
E[R_T(LAZOb)] \leq \frac{R^2}{2\eta} + \frac{\eta}{2}(D^2 + cdL^2)T + 4(D + dL) \times R \sum_{i=0}^{T} \mathbb{P}(A_i \setminus A_i^*) + 2L\delta T. \tag{42}
\]
Then plugging $\eta = \frac{R}{L \sqrt{dT}}$, $\delta = R \sqrt{\frac{d}{T}}$, $D = O(\sqrt{dL})$ and $\sum_{i=0}^{T} \mathbb{P}(A_i \setminus A_i^*) = O(\sqrt{T \sqrt{d} / \sqrt{\delta}})$ into (42), we can get that
\[
E[R_T(LAZOb)] = O(R\sqrt{dT}). \tag{43}
\]
This completes the proof. \qed

E. LAZOb Estimator

Similar to the derivation for LAZOb, we can denote $\hat{A}_i$ in the subsection to be the region triggering one-point in LAZOb, that is $\hat{A}_i \triangleq \{ u_t | D^a_p(u_t, w_{t-1}) \leq D \}$.

First, we prove that under some conditions, the LAZOb gradient estimator is bounded.

**Lemma 7:** Under Assumptions 1–2, we run (2) with $\hat{g}_t(x_t)$ and set $\eta = \frac{R}{L \sqrt{dT}}$, $\delta = R \sqrt{\frac{d}{T}}$. If $D < \frac{L}{\sqrt{\delta}}$, then for all $t$, $\| \hat{g}_t(x_t) \| \leq d^2L^2$. 

**Proof:** Since the first step of LAZOb is using the two-point estimator, then $\| \hat{g}_0^a(x_0) \| \leq d^2L^2$. With the choice of $\eta$ and $\delta$, we have that $\frac{R}{\sqrt{\delta}} = \frac{L}{\sqrt{T}}$. Then using (9), and assuming $\| \hat{g}_t(x_t) \| \leq d^2L^2$ holds for any $i < m$, then
\[
\| \hat{g}_m^a(x_m) \| = \| \hat{g}_m^{(1)}(x_m)1_{\hat{A}_i} + \hat{g}_m^{(2)}(x_m)1_{\hat{A}_i} \| \\
= \| \hat{g}_m^{(1)}(x_m) \|^2 1_{\hat{A}_i} + \| \hat{g}_m^{(2)}(x_m) \|^2 1_{\hat{A}_i} \\
\leq 2d^2D^2 \left( 4 + \frac{\eta^2}{\delta^2} \| g_{m-1}(x_{m-1}) \|^2 \right) 1_{\hat{A}_i} + d^2L^2 \| 1_{\hat{A}_i} \| \\
< 2d^2L^2 \left( 4 + \frac{1}{dL^2} \| g_{m-1}(x_{m-1}) \|^2 \right) 1_{\hat{A}_i} + d^2L^2 1_{\hat{A}_i} \\
= d^2L^2 1_{\hat{A}_i} + d^2L^2 1_{\hat{A}_i} = d^2L^2. \tag{44}
\]

Thus we can arrive at the conclusion using induction. \qed

1) Bias for LAZOb: To be self-contained, we restate the LAZOb part of Lemma 3 as follows.

**Restatement of Lemma 3:** (LAZOb estimator). If $\eta = \frac{R}{L \sqrt{dT}}$, $\delta = R \sqrt{\frac{d}{T}}$, and $D = O(\frac{L}{\sqrt{\delta}}) < \frac{L}{\sqrt{10}}$, then we have
\[
\sum_{t=0}^{T} E \left[ \| \hat{b}_t \| \right] = O(\sqrt{dT}). \tag{45}
\]

**Proof:** It is easy to see that (32) holds for $\hat{g}_b^a(x_t)$ if changing $A_i \setminus A_i^*$ to $\hat{A}_i \setminus \hat{A}_i^*$. Then according to the definition of $\hat{A}_i$, we can get that
\[
E \left[ \| \hat{b}_t \| \right] \leq \frac{d}{\delta^2} E \left[ \| 2(f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})) \| \right] \\
+ \frac{d}{\delta^2} E \left[ \| f_t(x_t - \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})1_{A_i \setminus A_i^*} \| \right] \\
\leq \frac{d}{\delta^2} E \left[ \| f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})1_{A_i \setminus A_i^*} \| \right] \\
+ \frac{d}{\delta^2} E \left[ \| f_t(x_t - \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})1_{A_i \setminus A_i^*} \| \right] \\
\leq dL \mathbb{P}(A_i \setminus A_i^*) \\
+ \frac{d}{\delta^2} E \left[ \| f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})1_{A_i \setminus A_i^*} \| \right] \\
+ \frac{d}{\delta^2} E \left[ \| f_t(x_t - \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})1_{A_i \setminus A_i^*} \| \right] \\
\leq dL \mathbb{P}(A_i \setminus A_i^*) + \frac{d^2L^2}{\delta} \mathbb{P}(A_i \setminus A_i^*) \tag{46}
\]

where the first inequality is due to $E \left[ \| X + Y \| \right] \leq E \left[ \| X \| \right] + E \left[ \| Y \| \right]$; the first term in the third inequality is because the Lipschitz condition; the fourth inequality is derived from the fact that if $u \in A_i \setminus A_i^* \subseteq A_i$, then $| f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1}) | \leq D | x_t - x_{t-1} + \delta(u_t - u_{t-1}) |$ and the second term in the third inequality is due to the Lipschitz condition; the second term in the fifth inequality is due to $E \left[ \| X + Y \| \right] \leq E \left[ \| X \| \right] + E \left[ \| Y \| \right]$, $| u_t - u_{t-1} | \leq 4$ and $| x_t - x_{t-1} | \leq \eta \| g_{t-1}(x_{t-1}) \|$; the last inequality is according to Lemma 7.

Thus, plugging $\eta = \frac{R}{L \sqrt{dT}}$, $\delta = R \sqrt{\frac{d}{T}}$, $D = O(\frac{L}{\sqrt{\delta}}) \leq \frac{L}{\sqrt{10}}$, and $\sum_{t=0}^{T} \mathbb{P}(A_i \setminus A_i^*) = O(\sqrt{dT})$ to (46), we can get that

2) The Second Moment Bound for LAZOb: To be self-contained, we restate the LAZOb part of Lemma 4 as follows.

**Restatement of Lemma 4:** (LAZOb estimator). Under Assumptions 1–3, the second moment bound of the gradient estimator $\hat{g}_b^a(x_t)$ satisfy that there exists a constant $c = O(1)$ such that
\[
E \left[ \| \hat{g}_b^a(x_t) \|^2 \| x_t \| \right] \leq cdL^2 + \frac{2d^2D^2\eta^2}{\delta^2} \| \hat{g}_{t-1}^a(x_{t-1}) \|^2 + 8d^2D^2. \tag{47}
\]

**Proof:** Using the definition of $\hat{g}_b^a(x_t)$, we have
\[
E \left[ \| \hat{g}_b^a(x_t) \|^2 \| x_t \| \right] = E \left[ \frac{d^2}{\delta^2} (f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1}))^2 1_{A_i \setminus A_i^*} \| x_t \| \right] \\
+ E \left[ \frac{d^2}{\delta^2} (f_t(x_t - \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1}))^2 1_{A_i \setminus A_i^*} \| x_t \| \right]
\]
≤ \mathbb{E}\left[ \frac{d^2}{2\sigma^2} \left(f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})\right)^2 \mathbf{1}_{\tilde{A}_t} | x_t \right] \\
\quad + \mathbb{E}\left[ \frac{d^2}{4\sigma^2} \left(f_t(x_t + \delta u_t) - f_t(x_t - \delta u_t)\right)^2 \mathbf{1}_{\tilde{A}_t} | x_t \right] \\
\leq cdL^2\sqrt{\mathbb{P}(\tilde{A}_t^c | x_t)} \\
\quad + \mathbb{E}\left[ \frac{d^2}{\sigma^2} \left(f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})\right)^2 \mathbf{1}_{\tilde{A}_t} | x_t \right] \\
\leq cdL^2 + \frac{d^2D^2}{\sigma^2} \mathbb{E}\left[ \left\| x_t - x_{t-1} + \delta(u_t - u_{t-1}) \right\| \mathbf{1}_{\tilde{A}_t} | x_t \right] \\
\leq cdL^2 + 2\frac{d^2D^2}{\sigma^2} \mathbb{E}\left[ \left\| x_t - x_{t-1} \right\|^2 | x_t \right] + 8d^2D^2 \\
\leq cdL^2 + 2\frac{d^2D^2}{\sigma^2} \mathbb{E}\left[ \left\| x_t - x_{t-1} \right\|^2 | x_t \right] + 8d^2D^2 \tag{48}

where the first inequality is due to \( \|u_t\| = 1 \) and \( \mathbf{1}_{\tilde{A}_t} \times \mathbf{1}_{\tilde{A}_t} = 0 \); the second inequality is due to \( \tilde{A}_t \subset \tilde{A}_t^c \); the second inequality is due to Lemma 5; the third inequality is derived from the fact that if \( u_t \in \tilde{A}_t \setminus \tilde{A}_t^c \), then \( |f_t(x_t + \delta u_t) - f_{t-1}(x_{t-1} + \delta u_{t-1})| \leq D|\|x_t - x_{t-1} + \delta(u_t - u_{t-1})|\| \); the fourth inequality is due to \( \mathbb{E}[\|X + Y\|] \leq \mathbb{E}[\|X\|] + \mathbb{E}[\|Y\|] \), \( \|u_t - u_{t-1}\| \leq \delta \); the last inequality is due to \( \|x_t - x_{t-1}\| \leq \epsilon \) and \( \mathbb{E}[\left\| \tilde{g}_t^a(x_t) \right\|^2] \).

3) Regret for LAZOa: To be self-contained, we restate the LAZOa part of Theorem 1 as follows.

**Restatement of Theorem 1:** (LAZOa estimator). Under Assumptions 1–4 (replacing \( A_t \) with \( \tilde{A}_t \)), we run LAZOa for \( T \) iterations with \( \eta = \frac{R}{L \sqrt{dT}} \), \( \beta = R \sqrt{\frac{d}{T}} \). If we use \( D = O(\frac{L}{\sqrt{d}}) \) for LAZOa, then the regret satisfies

\[
\mathbb{E}[R_T(LAZOa)] = O(\sqrt{dT}).
\]

**Proof:** First, plugging \( \eta = \frac{R}{L \sqrt{dT}} \), \( \beta = R \sqrt{\frac{d}{T}} \) to the (46), we can simplify it to

\[
\mathbb{E}\left[ \left\| \tilde{b}_t \right\| \right] \leq d(L + 3D)\mathbb{P}(\tilde{A}_t \setminus \tilde{A}_t^c).
\]

Similar to the derivation for (38), we can get that

\[
\mathbb{E}\left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) \right] \leq \frac{1}{2\eta} \left\| x_0 - x \right\|^2 + \frac{\eta}{2} \mathbb{E}\left[ \sum_{t=0}^{T} \left\| \tilde{g}_t^a(x_t) \right\|^2 \right] \\
+ d(L + 3D)R \sum_{t=0}^{T} \mathbb{P}(\tilde{A}_t \setminus \tilde{A}_t^c). \tag{49}
\]

Then similar to (27), we obtain that

\[
\mathbb{E}\left[ \sum_{t=0}^{T} f_t(x_t) - \sum_{t=0}^{T} f_t(x) \right] \\
= \mathbb{E}\left[ \sum_{t=0}^{T} f_{\delta,t}(x_t) - \sum_{t=0}^{T} f_{\delta,t}(x) + \sum_{t=0}^{T} (f_t(x_t) - f_{\delta,t}(x_t)) \right] \\
- \sum_{t=0}^{T} (f_t(x) - f_{\delta,t}(x))
\]

\[
\leq \frac{1}{2\eta} \left\| x_0 - x \right\|^2 + \frac{\eta}{2} \mathbb{E}\left[ \sum_{t=0}^{T} \left\| \tilde{g}_t^a(x_t) \right\|^2 \right] \\
+ d(L + 3D)R \sum_{t=0}^{T} \mathbb{P}(\tilde{A}_t \setminus \tilde{A}_t^c) + 2\delta L T. \tag{50}
\]

Second, plugging \( \eta = \frac{R}{L \sqrt{dT}} \), \( \beta = R \sqrt{\frac{d}{T}} \) to (48), we simplify the second moment bound of LAZOa estimator to

\[
\mathbb{E}\left[ \left\| \tilde{g}_t^a(x_t) \right\|^2 \right] \leq 10d^2D^2 + cdL^2. \tag{51}
\]

Then plugging \( x^* = \arg \min_{x \in X} \sum_{t=0}^{T} f_t(x) \) and (51) into (50), and according to the definition of regret in (1), we can reach the conclusion that

\[
\mathbb{E}[R_T(LAZOa)] \leq \frac{R^2}{2\eta} + \frac{\eta}{2}(10d^2D^2 + cdL^2)T \\
+ d(L + 3D)R \sum_{t=0}^{T} \mathbb{P}(\tilde{A}_t \setminus \tilde{A}_t^c) + 2\delta L T. \tag{52}
\]

Finally, plugging \( \eta = \frac{R}{L \sqrt{dT}} \) and \( \delta = R \sqrt{\frac{d}{T}} \) and \( D = O(\frac{L}{\sqrt{d}}) \) and \( \sum_{t=0}^{T} \mathbb{P}(\tilde{A}_t \setminus \tilde{A}_t^c) = O(\sqrt{T}/\sqrt{d}) \) into (52), we can get that

\[
\mathbb{E}[R_T(LAZOa)] = O(\sqrt{dT}). \tag{53}
\]

This completes the proof. \( \square \)

**VII. CONCLUSIONS**

This article proposes a novel ZO gradient estimation method based on a lazy query condition. Different from the classic ZO methods, LAZO monitors the informativeness of old queries, and then adaptively reuses them to construct the low-variance stochastic gradient estimates. We rigorously establish that through judiciously reusing old queries, LAZO not only saves queries per iteration but also achieves the regrets of the symmetric two-point ZO methods [17].

**REFERENCES**

[1] T. Salimans, J. Ho, X. Chen, S. Sidor, and I. Sutskever, “Evolution strategies as a scalable alternative to reinforcement learning,” Sep. 2017, arXiv:1703.03864.

[2] P.-Y. Chen, H. Zhang, Y. Sharma, J. Yi, and C.-J. Hsieh, “ZOO: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models,” in Proc. ACM Workshop Artif. Intell. Secur., Dallas, TX, USA, Nov. 2017, pp. 15–26.

[3] S. Karyampaa, A. Prakash, and M. K. Qureshi, “Maze: Data-free model stealing attack using zeroth-order gradient estimation,” in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2021, pp. 13814–13823.

[4] Y. Ruan, Y. Xiong, S. Reddi, S. Kumar, and C.-J. Hsieh, “Learning to learn by zeroth-order oracle,” in Proc. Int. Conf. Learn. Representations, 2019, p. 2.

[5] X. Song, W. Gao, Y. Yang, K. Choromanski, A. Pacchiano, and Y. Tang, “ES-MAML: Simple Hessian-free meta learning,” in Proc. Int. Conf. Learn. Representations, 2019, p. 4.

[6] D. Sow, K. Ji, and Y. Liang, “On the convergence theory for Hessian-free bilevel algorithms,” in Proc. Adv. Neural Inf. Process. Syst., New Orleans, LA, USA, 2022, p. 4.

[7] L. Xie, K. Huang, F. Xu, and Q. Shi, “ZO-DARTS: Differentiable architecture search with zeroth-order approximation,” in Proc. IEEE Int. Conf. Acount., Speech Signal Process., Rhodes Island, Greece, Jun. 2023, pp. 1–5.

[8] T. Chen, S. Barbarossa, X. Wang, G. B. Giannakis, and Z.-L. Zhang, “Learning and management for Internet of Things: Accounting for
XIAO et al.: LAZY QUERIES CAN REDUCE VARIANCE IN ZEROTH-ORDER OPTIMIZATION

[57] X. Chen, S. Z. Wu, and M. Hong, “Understanding gradient clipping in private SGD: A geometric perspective,” in Proc. Adv. Neural Inf. Process. Syst., Dec. 2020, p. 4.

[58] J. Bernstein, J. Zhao, K. Azizzadenesheli, and A. Anandkumar, “SignSGD with majority vote is communication efficient and fault tolerant,” in Proc. Int. Conf. Learn. Representations, 2018, pp. 4–6.

[59] Q. Xiao, Q. Ling, and T. Chen, “Lazy queries can reduce variance in zeroth-order optimization,” 2022, arXiv:2206.07126.

[60] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, “Global convergence of policy gradient methods for the linear quadratic regulator,” in Proc. Int. Conf. Mach. Learn., Vienna, Austria, Jul. 2018, pp. 1467–1476.

Quan Xiao received the B.S. degree in statistics from the University of Science and Technology of China, Hefei, China, in 2020. She was a Research Assistant with the School of Computer Science and Engineering at Sun Yat-sen University, Guangzhou, China, in 2021. She is now working toward the Ph.D. degree with the Department of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute, Troy, NY, USA. Her primary research areas are bilevel and distributed optimization, and their applications for signal processing and machine learning.

Qing Ling (Senior Member, IEEE) received the B.E. degree in automation and the Ph.D. degree in control theory and control engineering from the University of Science and Technology of China, Hefei, China, in 2001 and 2006, respectively. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, MI, USA, from 2006 to 2009, and an Associate Professor with the Department of Automation, University of Science and Technology of China, from 2009 to 2017. He is currently a Professor with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China. His current research interest includes distributed and decentralized optimization and its application in machine learning. He received the 2017 IEEE Signal Processing Society Young Author Best Paper Award as a Supervisor. He was a Senior Area Editor of IEEE SIGNAL PROCESSING LETTERS.

Tianyi Chen received the B.E. degree in communication science and engineering from Fudan University, Shanghai, China, in 2014, and the M.Sc. and Ph.D. degrees in electrical and computer engineering (ECE) from the University of Minnesota (UMN), Minneapolis, MN, USA, in 2016 and 2019, respectively. Since August 2019, he has been an Assistant Professor with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA. His research focuses on the theory and application of distributed and bilevel optimization to problems emerging in signal processing and machine learning. He was the inaugural recipient of the IEEE Signal Processing Society Best Ph.D. Dissertation Award in 2020, a recipient of the NSF CAREER Award in 2021, and a recipient of the Amazon Research Award in 2022. He is also a co-author of the Best Student Paper Award at the NeurIPS Federated Learning Workshop in 2020 and at ICASSP in 2021.