Federated Online and Bandit Convex Optimization

Kumar Kshitij Patel[1] **Lingxiao Wang**[1] **Aadirupa Saha**[2] **Nati Srebro**[1]

Abstract

We study the problems of *distributed online and bandit convex optimization* against an adaptive adversary. We aim to minimize the average regret on M machines working in parallel over T rounds with R intermittent communications. Assuming the underlying cost functions are convex and can be generated adaptively, our results show that *collaboration is not beneficial when the machines have access to the first-order gradient information at the queried points*. This is in contrast to the case for stochastic functions, where each machine samples the cost functions from a fixed distribution. Furthermore, we delve into the more challenging setting of *federated online optimization with bandit (zeroth-order) feedback*, where the machines can only access values of the cost functions at the queried points. The key finding here is *identifying the high-dimensional regime where collaboration is beneficial and may even lead to a linear speedup in the number of machines*. We further illustrate our findings through federated adversarial linear bandits by developing novel distributed single and two-point feedback algorithms. Our work is the first attempt towards a systematic understanding of federated online optimization with limited feedback, and it attains tight regret bounds in the intermittent communication setting for both first and zeroth-order feedback. Our results thus bridge the gap between stochastic and adaptive settings in federated online optimization.

1. Introduction

We consider the following distributed regret minimization problem on M machines over a horizon of length T:

$$\sum_{m\in[M],t\in[T]} f_t^m(x_t^m) - \min_{\|x^\star\|_2\leq B} \sum_{m\in[M],t\in[T]} f_t^m(x^\star), \quad (1)$$

where f_t^m is an arbitrary convex cost function observed by machine m at time t, x_t^m is the model the machine plays based on available history, and the comparator $x^\star$ is shared across the machines. This formulation is a natural extension of the classical federated optimization (McMahan et al., 2016; Kairouz et al., 2019). It captures many real-world applications, such as mobile keyboard prediction (Hard et al., 2018; Chen et al., 2019; Hartmann, 2021), self-driving vehicles (Elbir et al., 2020; Nguyen et al., 2022), and recommendation systems (Shi et al., 2021; Liang et al., 2021; Khan et al., 2021). These applications involve sequential decision-making across multiple machines where data is generated in real time and might not be stored. This regret minimization problem can be solved in a federated manner, i.e., by storing and analyzing the data locally at each machine while only communicating the models intermittently, reducing communication load and mitigating privacy concerns. However, *sequential decision-making and potentially adversarial cost functions bring new challenges to this problem*. In particular, as opposed to usual federated optimization, where we want to come up with one good final model, the goal here is to develop a sequence of instantaneously good models. This is a natural requirement in many applications, such as a voice assistant (Hao, 2020; Google, 2023), where the service needs to improve as it is used. These challenges require designing new methods to benefit from collaboration while attaining small regret.

Furthermore, we want to solve problem (1) in the *intermittent communication* (IC) setting (Woodworth et al., 2018; 2021) where the machines work in parallel and are allowed to communicate R times with K time steps in between communication rounds. The IC setting has been widely studied over the past decade (Zinkevich et al., 2010; Cotter et al., 2011; Dekel et al., 2012; Zhang et al., 2013; 2015; Shamir et al., 2014; Stich, 2018; Dieuleveut & Patel, 2019; Woodworth et al., 2020a; Bullins et al., 2021; Patel et al., 2022), and it captures the expensive nature of communication in collaborative learning, such as in cross-device federated learning (McMahan et al., 2016; Kairouz et al., 2019).

Limitations of Recent Attempts. Although many recent attempts (Wang et al., 2020; Dubey & Pentland, 2020; Huang et al., 2021; Li & Wang, 2022; He et al., 2022; Gauthier et al., 2022; Gogineni et al., 2022; Dai & Meng, 2022; Kuh, 2021; Mitra et al., 2021) have been made towards tack-

This work was done while AS was visiting TTIC. [1]TTIC [2]Apple. Correspondence to: Kumar Kshitij Patel <kkpatel@ttic.edu>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

ling problem (1), most existing works (Wang et al., 2020; Dubey & Pentland, 2020; Huang et al., 2021; Li & Wang, 2022) study problem (1) by focusing on the ***"stochastic"*** setting, where $\{f_t^m\}$'s are sampled from distributions specified in advance. However, real-world applications may have distribution shifts, un-modeled perturbations, or even an adversarial sequence of cost functions, all of which violate the fixed distribution assumption, and thus the above problem setups fail. To alleviate this issue, in this paper, we extend our understanding of distributed online and bandit convex optimization, i.e., problem (1), to ***"adaptive"*** adversaries that could potentially generate a worst-case sequence of cost functions. Although some recent works have underlined the importance of the adaptive setting (Gauthier et al., 2022; Gogineni et al., 2022; Dai & Meng, 2022; Kuh, 2021; Mitra et al., 2021; He et al., 2022), our understanding of the optimal regret guarantees for problem (1) with different classes of algorithms is still lacking.

Our Contributions. Our work is the first attempt towards a *systematic understanding of federated online optimization with limited feedback and attains tight regret bounds (in at least some regime) in the intermittent communication setting for both first and zeroth-order feedback, as well as bridges the gap between stochastic and adaptive settings in federated online optimization*. The following summarizes the main contributions of our work:

- **Federated Online Optimization with First-Order (Gradient) Feedback.** We first show in Section 3 that, under usual assumptions, there is no benefit of collaboration if all the machines have access to the gradients, a.k.a. first-order feedback for their cost functions. Specifically, in this setting, running online gradient descent on each device without any communication is min-max optimal for problem (1) (c.f., Theorems 3.1, 3.2). This motivates us to study weaker forms of feedback/oracles accesses where collaboration can be provably useful.
- **Federated Online Optimization with Zeroth-Order (Bandit) Feedback.** We then study the problem of federated adversarial linear bandits in Section 4, which is an important application of problem (1) and has not been studied in federated learning literature. We propose a novel federated projected online stochastic gradient descent algorithm to solve this problem (c.f., Algorithm 2). Our algorithm only requires single bandit feedback for every function. We show that when the problem dimension d is large, our algorithm can outperform the non-collaborative baseline, where each agent solves its problem independently. Additionally, **when d is larger than $\mathcal{O}\big(M\sqrt{T/R}\big)$, we prove that the proposed algorithm can achieve $\mathcal{O}\big(d/\sqrt{MT}\big)$ average regret**, where M is the number of agents and R is the communication round (c.f., Theorem 4.1). These results suggest that one can benefit from collaborations in the more challenging setting of using bandit feedback when the problem dimension is large.
- **Bandit Feedback Beyond Linearity.** We next consider the general (non-linear) problem (1) with bandit feedback in Section 5, and study a natural variant of FedAvg equipped with a stochastic gradient estimator using two-point feedback (Shamir, 2017). We show that collaboration reduces the variance of the stochastic gradient estimator and is thus beneficial for problems of high enough dimension (c.f., Theorems 5.1 and 5.2). We prove a linear speedup in the number of machines for high-dimensional problems, which mimics the stochastic setting with first-order gradient information (Woodworth et al., 2021; Woodworth & Srebro, 2021).
- **Tighter Rates with Two-Point Bandit Feedback.** Additionally, we show that (c.f., Corollary 5.3) one can achieve better regret for federated adversarial linear bandits using two-point feedback, indicating that multi-point feedback can be beneficial for federated adversarial linear bandits.
- **Characterizing the General Class of Problem.** Finally, we characterize the full space of related min-max problems in distributed online optimization, thus connecting the adversarial and stochastic versions of problem (1) (see Appendix A). This underlines how we understand only a small space of problems in federated online optimization, despite more than a decade of work. This also identifies problems at the intersection of sequential and collaborative decision-making for future research.

2. Problem Setting

This section introduces definitions and assumptions used in our analysis. We formalize our adversary class and algorithm class. We also specify the notion of min-max regret.

Notation. We denote the horizon by $T = KR$. $\succeq, \preceq, \cong$ denote inequalities up to numerical constants. We denote the average function by $f_t(\cdot) := \sum_{m\in[M]} f_t^m(\cdot)/M$ for all $t \in [T]$. We use $\mathbb{1}_A$ to denote the indicator function for the event A. $\mathbb{B}_2(B) \subset \mathbb{R}^d$ denotes the L_2 ball of radius B centered at zero. We suppress the indexing in the summation $\sum_{t\in[T],m\in[M]}$ to $\sum_{t,m}$ wherever the usage is clear.

Function Classes. We consider several common function classes (Saha & Tewari, 2011; Hazan, 2016) in this paper:

1. $\boldsymbol{\mathcal{F}^{G,B}}$, the class of convex, differentiable, and G-Lipschitz functions, i.e.,

$$\forall x, y \in \mathbb{R}^d,\ |f(x) - f(y)| \leq G\left\|x - y\right\|_2$$

with bounded optima $x^\star \in \mathbb{B}_2(B)$.

2. $\boldsymbol{\mathcal{F}^{H,B}}$, the class of convex, differentiable, and H-smooth functions, i.e.,

$$\forall x, y \in \mathbb{R}^d,\ \left\|\nabla f(x) - \nabla f(y)\right\|_2 \leq H\left\|x - y\right\|_2$$

with bounded optima $x^\star \in \mathbb{B}_2(B)$. We also define $\boldsymbol{\mathcal{F}^{G,H,B}} := \boldsymbol{\mathcal{F}^{G,B}} \cap \boldsymbol{\mathcal{F}^{H,B}}$, i.e., the class of Lipschitz and smooth functions.

3. $\boldsymbol{\mathcal{F}_{lin}^{G,B}} \subset \boldsymbol{\mathcal{F}^{G,B}}$, which includes linear cost functions with gradients bounded by G. Note that this includes the cost functions in federated adversarial linear bandits, discussed in Section 4.

Adversary Model. Note that in the most general setting, each machine will encounter arbitrary functions from a class $\mathcal{F}$ at each time step. Our algorithmic results are for this general model, which is usually referred to as an ***"adaptive"*** adversary. Specifically, we allow the adversary's functions to depend on the past sequence of models played by the machines but not on the randomness used by the machines, i.e., $\{f_t^m\}^{m\in[M]}$ for all t can depend on $\left(\{x_i^m\}_{i\in[t-1]}^{m\in[M]}, A\right)$, where A is the algorithm being used by the machines.

We also consider a weaker ***"stochastic"*** adversary model to explain some baseline results. More specifically, the adversary cannot adapt to the sequence of the models used by each machine but must fix a distribution in advance for each machine, i.e., $\forall m \in [M],\ \mathcal{D}_m \in \Delta(\mathcal{F})$ such that at each time $t \in [T]$, $f_t^m \sim \mathcal{D}_m$. An example of this less challenging model is *distributed stochastic optimization* where $f_t^m(\cdot) := f(\cdot; z_t^m \sim \mathcal{D}_m)$ for $f(\cdot;\cdot) \in \mathcal{F}$ and z_t^m is a data-point sampled from the distribution $\mathcal{D}_m$. For the stochastic adversaries throughout we denote

$$F_m(x) := \mathbb{E}_{f\sim\mathcal{D}_m}[f(x)],\ F(x) := \frac{1}{M}\sum_{m\in[M]} F_m(x),$$

where with a slight abuse of notation we suppress $z \sim \mathcal{D}_m$. The distinction between different adversary models is discussed further in Appendix A.

Oracle Model. We consider three kinds of access to the cost functions in this paper. Each machine $m \in [M]$ for all time steps $t \in [T]$ has access to one of the following:

1. **Gradient** of f_t^m at a single point, a.k.a., first-order feedback.
2. **Function value** of f_t^m at a single point, a.k.a., one-point bandit feedback.
3. **Function values** of f_t^m at two different points, a.k.a., two-point bandit feedback.

Remark 2.1. Note that we always look at the regret at the queried points. Thus in the two-point feedback setting, if the machine m queries at points $x_t^{m,1}$ and $x_t^{m,2}$ at time t, it incurs the cost $f_t^m(x_t^{m,1}) + f_t^m(x_t^{m,2})$ (c.f., Theorem 5.1).

Algorithm Class. We assume the algorithms on each machine can depend on all the history it has seen. Since we are in the IC setting, at time t on machine m, the algorithm A's output can only depend on

$$\left(f_1^{1:M}, \dots, f_{\tau(t)}^{1:M}, f_{\tau(t)+1}^m, \dots, f_{t-1}^m\right),$$

where $\tau(t)$ is the last time step smaller than or equal to t where communication happened. We assume $T = KR$ so that $\tau(t) := t \mod K$. In other words, the algorithms' output on a machine can depend on all the information (gradients or function values) it has seen on the machine or other machines' information communicated to it. We denote this class of algorithms by $\mathcal{A}_{IC}$ and add super-scripts 1, 0, $(0,2)$ to denote first-order, one-point bandit, and two-point bandit feedback. Thus, we consider three algorithm classes $\mathcal{A}_{IC}^1,\ \mathcal{A}_{IC}^0,\ \mathcal{A}_{IC}^{0,2}$ in this paper.

Finally, we consider two more assumptions controlling how similar the cost functions look across machines and the average regret at the comparator (Srebro et al., 2010):

Assumption 1 (Bounded First-Order Heterogeneity)**.** For all $t \in [T],\ x \in \mathbb{R}^d$,

$$\frac{1}{M}\sum_{m\in[M]} \|\nabla f_t^m(x) - \nabla f_t(x)\|_2^2 \le \zeta^2 \le 4G^2.$$

Remark 2.2 (Bounded First-Order Heterogeneity for Stochastic Setting). In the stochastic setting, Woodworth et al. (2020b) consider a related but more relaxed assumption, i.e., $\forall\, x \in \mathbb{R}^d$,

$$\frac{1}{M}\sum_{m\in[M]} \|\nabla F_m(x) - \nabla F(x)\|_2^2 \le \zeta^2 \le 4G^2.$$

This assumption does not require the gradients at each time step to be point-wise close but requires them to be close only on average over the draws from the machines' distributions.

Assumption 2 (Bounded Optimality at Optima)**.** For all $x^\star \in \arg\min_{x\in\mathbb{B}_2(B)} \sum_{t\in[T]} f_t(x)$, $\sum_{t\in[T]} f_t(x^\star)/T \le F_\star$. For non-negative functions in $\mathcal{F}^{H,B}$, this implies $\sum_{t\in[T]} \|\nabla f_t(x^\star)\|_2^2/T \le HF_\star$ (c.f., Lemma 4.1 in Srebro et al. (2010)).

Remark 2.3 (Bounded Optimiality for Stochastic Setting). When the functions are generated stochastically, denote $F_\star := \min_{x\in\mathbb{B}_2(B)} F(x)$ as the minimum realizable function value. Then assuming $\mathcal{D}_m$ is supported on functions in $\mathcal{F}^{H,B}$, implies for all $t \in [T]$,

$$\mathbb{E}_{\{f_t^m\sim\mathcal{D}_m\}_{m\in[M]}}\left\|\frac{1}{M}\sum_{m\in[M]} \nabla f_t^m(x^\star)\right\|_2^2 \le HF_\star$$

(c.f., Lemma 3 in Woodworth & Srebro (2021)).

Min-Max Regret. We now define our problem class. We use $\mathcal{P}_{M,K,R}(\mathcal{F}) := \mathcal{F}^{\otimes MKR}$ to denote all selections of MKR functions from a class $\mathcal{F}$. We use the argument ζ, $F_\star$ to further restrict this to selections that satisfy Assumptions 1 and 2 respectively. In this paper, we consider four problem classes: $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$, $\mathcal{P}_{M,K,R}(\mathcal{F}^{H,B},\zeta,F_\star)$, $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,H,B},\zeta,F_\star)$, and $\mathcal{P}_{M,K,R}(\mathcal{F}_{lin}^{G,B},\zeta)$. We are interested in characterizing the min-max regret for these classes. In particular, for a problem class $\mathcal{P}$ and algorithm class $\mathcal{A}$, we want to characterize up to numerical constants the following quantity:

$$\mathcal{R}(\mathcal{P}) := \min_{A\in\mathcal{A}} \max_{P\in\mathcal{P}} \mathbb{E}_A \frac{1}{MT}\left(\sum_{\substack{t\in[T]\\ m\in[M]}} f_t^m(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)} \sum_{\substack{t\in[T]\\ m\in[M]}} f_t^m(x^\star)\right), \tag{2}$$

where A is a randomized algorithm producing models x_t^m's. We add super-scripts 0, $(0,2)$, 1 to $\mathcal{R}$ to denote the oracle access of the min-player's algorithm.

The second player, i.e., the adversary selecting the functions, does not benefit from randomization for problem (2), which is why ignore that here[1]. It is also important to note that the max player does not have access to the randomness of the min player (c.f., problem (P1) in appendix A). Finally, note again that the max player can adapt to the sequence of models played by A, which is why we call this regret minimization against an adaptive adversary. We want to characterize the min-max value of this game, i.e., $\mathcal{R}(\mathcal{P})$ in this paper.

Most existing work (Khaled et al., 2020; Woodworth et al., 2020b; Patel et al., 2022; Wang et al., 2022) in federated learning has instead focused on characterizing the min-max regret for the following simpler problem[2]:

$$\begin{aligned}
&\mathcal{R}_{stoc.}(\mathcal{P})\\
&:= \min_{A\in\mathcal{A}} \max_{\{\mathcal{D}_m\in\Delta(\mathcal{F})\}} \mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left(\frac{1}{MT}\sum_{t,m} f_t^m(x_t^m)\right.\\
&\qquad\left. - \min_{x^\star\in\mathbb{B}_2(B)} \frac{1}{MT}\sum_{t,m}\mathbb{E}_{f_t^m\sim\mathcal{D}_m}[f_t^m(x^\star)]\right)\\
&= \min_{A\in\mathcal{A}} \max_{\{\mathcal{D}_m\in\Delta(\mathcal{F})\}} \mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left(\frac{1}{MT}\sum_{t,m} F_m(x_t^m)\right.\\
&\qquad\left. - \min_{x^\star\in\mathbb{B}_2(B)} \frac{1}{M}\sum_{m} F_m(x^\star)\right),
\end{aligned} \tag{3}$$

where $F_m(\cdot) := \mathbb{E}_{f\sim\mathcal{D}_m}[f(\cdot)]$. This problem has a "stochastic" adversary. The min-max complexity of this easier problem is understood in the "homogeneous" setting when the machines have the same distribution (Woodworth et al., 2021; Woodworth, 2021). However, it is not yet fully understood in the "heterogeneous" setting where these distributions are allowed to differ across the machines (Woodworth et al., 2020b; Wang et al., 2022; Glasgow et al., 2022). In general, note that for any problem class $\mathcal{P}$, $\mathcal{R}_{stoc.}(\mathcal{P}) \le \mathcal{R}(\mathcal{P})$, and we are interested in understanding the higher $\mathcal{R}(\mathcal{P})$. We further discuss some related distributed problems and their potential applications in Appendix A.

3. Collaboration Does Not Help with First-Order Feedback

In this section, we will explain why collaboration does not improve the min-max regret $\mathcal{R}(\mathcal{P})$ for the adaptive problem while using first-order oracles, even though it does improve $\mathcal{R}_{stoc.}(\mathcal{P})$ for the stochastic problem.

We first consider the class of algorithms in $\mathcal{A}_{IC}^1$, i.e., algorithms that have access to one gradient per cost function on each machine at each time step. This class is very strong in the serial setting, i.e., when $M=1$. For instance, online gradient descent (OGD) (Zinkevich et al., 2010; Hazan, 2016) attains the min-max regret for both the problem classes $\mathcal{P}_{1,K,R}(\mathcal{F}^{G,B})$ and $\mathcal{P}_{1,K,R}(\mathcal{F}^{H,B}, F_\star)$ (Woodworth & Srebro, 2021). This raises the question of whether a distributed version of OGD can also be shown to be min-max optimal when $M>1$. The answer is, unfortunately, no! To state this more formally, we first introduce a trivial non-collaborative algorithm in the distributed setting: we run online gradient descent independently on each machine without any communication (see Algorithm 1). We prove the following result for this algorithm (lies in the class $\mathcal{A}_{IC}^1$), which essentially shows the non-collaborative baseline is min-max optimal.

Algorithm 1: Non-collaborative OGD (η)

Initialize $x_0^m = 0$ on all machines $m\in[M]$
for $t\in\{0,\ldots,KR-1\}$ **do**
  for $m\in[M]$ ***in parallel*** **do**
    Play model x_t^m and see function $f_t^m(\cdot)$
    Incur loss $f_t^m(x_t^m)$
    Compute gradient at point x_t^m as $\nabla f_t^m(x_t^m)$
    $x_{t+1}^m \leftarrow x_t^m - \eta\cdot\nabla f_t^m(x_t^m)$

Theorem 3.1 (Optimal Bounds with Non-collaborative OGD for Lipschitz Functions)**.** *Algorithm 1 incurs an average regret of $\mathcal{O}\left(GB/\sqrt{T}\right)$ for problem class $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$. This regret is optimal for $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$ among algorithms in $\mathcal{A}_{IC}^1$, where $0\le\zeta\le 2G$.*

Proof. We first prove the upper bound on the average regret

[1] It makes sense to make the randomization on the second player explicit when comparing to a weaker comparator, c.f., problem (P3) in Appendix A.

[2] These works aim to upper bound the function sub-optimality at a single point, but most of their analyses actually provide regret guarantees that are converted to function sub-optimality bounds at the averaged iterate by applying Jensen's inequality.

of non-collaborative OGD and then show that it is optimal, i.e., equals $\mathcal{R}^1\left(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)\right)$. Note the following bound is always true for any stream of functions and sequence of models, we are just changing the comparator:

$$\frac{1}{M}\sum_{m\in[M]}\left(\sum_{t\in[KR]} f_t^m(x_t^m) - \min_{x^{m,\star}\in\mathbb{B}_2(B)}\sum_{t\in[KR]} f_t^m(x^m)\right)$$
$$\geq \frac{1}{M}\sum_{t\in[KR],m\in[M]} f_t^m(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)}\sum_{t\in[KR]} f_t(x).$$

This means we can upper bound $\mathcal{R}^1(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta))$ by running online gradient descent (OGD) independently on each machine without collaboration. In other words, for functions from $\mathcal{F}^{G,B}$, we have:

$$\mathcal{R}^1\left(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)\right) \preceq \mathcal{R}^1\left(\mathcal{P}_{1,K,R}(\mathcal{F}^{G,B})\right) \cong \frac{GB}{\sqrt{T}}. \tag{4}$$

The min-max rate for a single machine follows classical results using vanilla OGD (c.f., Theorem 3.1 in Hazan (2016)). Now we prove that this average regret is optimal by noting the following:

$$\mathcal{R}\left(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)\right)\cdot MT \geq$$
$$\min_{A\in\mathcal{A}_{IC}^1}\max_{\{f_t\in\mathcal{F}\}_t}\mathbb{E}_A\left(\sum_{m,t} f_t(x_t^m) - M\min_{x^\star\in\mathbb{B}_2(B)}\sum_t f_t(x^\star)\right) \tag{5}$$
$$\geq \min_{A\in\mathcal{A}_{IC}^1}\ \mathbb{E}_{\substack{A\\ \{f_t\sim\mathcal{D}\}_t}}\left(\sum_{m,t} f_t(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)} M\sum_t f_t(x^\star)\right),$$
$$\succeq \frac{GB}{\sqrt{T}}, \tag{6}$$

where $\mathcal{F}$ denotes $\mathcal{F}^{G,B}$ in the first inequality, and this inequality holds because we can choose a weaker adversary that can pick functions that do not vary across the machines but may change over time (which implies that $\zeta = 0$). The second inequality holds because we choose an even weaker adversary that fixes the distribution for picking these functions in advance. Furthermore, the last inequality follows from choosing $\mathcal{D}$ as the usual lower bound construction for serial first-order online convex optimization (Hazan, 2016). We include the construction in Appendix B. Combining equations (4) and (6) proves the claim, showing that non-collaborative OGD attains the min-max regret. □

As the lower bound instance is linear, it is easy to show that the above theorem holds for the linear functions $\mathcal{F}_{lin}^{G,B}$.

We can use the same proof strategy to prove a similar result for smooth functions. The upper bound in the serial setting follows from a classical work on optimistic rates (c.f, Theorem 3 (Srebro et al., 2010)) while the lower bound follows from a more recent paper (c.f., Theorem 4 (Woodworth & Srebro, 2021) and Appendix B).

Theorem 3.2 (Optimal Bounds with Non-collaborative OGD for Smooth Functions)**.** *Algorithm 1 incurs the optimal regret of* $\Theta\left(HB^2/T + \sqrt{HF_\star}B/\sqrt{T}\right)$ *for problem class* $\mathcal{P}_{M,K,R}(\mathcal{F}^{H,B},\zeta,F_\star)$.

Implications of Theorems 3.1 and 3.2: The above theorems imply that in the worst case, there is no benefit of collaboration if the machines already have access to gradient information! This is counter-intuitive at first because several works have shown in the stochastic setting, $\mathcal{R}_{stoc}(\mathcal{P})$ indeed improves with collaboration in several regimes (Woodworth et al., 2020b; Koloskova et al., 2020).

How do we reconcile these results?

Note that while proving Theorem 3.1, we crucially rely on equation (5), i.e., lower bounding by the min-max regret with the adversary that can put the same function on each machine every time-step. Let's compare it against the expression for stochastic min-max regret. One key difference is that in equation (5), $x^\star$ is allowed to depend on the random draws of f_t from $\mathcal{D}$, which makes the comparator smaller and regret minimization harder. In particular, for any problem class $\mathcal{P}$, we have

$$\mathcal{R}_{stoc.}(\mathcal{P})\cdot MT \leq$$
$$\min_{A\in\mathcal{A}_{IC}^1}\ \mathbb{E}_{\substack{A\\ \{f_t\sim\mathcal{D}\}_t}}\left(\sum_{m,t} f_t(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)} M\sum_t f_t(x^\star)\right).$$

This is the key reason why there is a benefit of collaboration in stochastic federated optimization, i.e., the problem captured by $\mathcal{R}^{stoc.}(\mathcal{P})$. The lower bound in equation (5) does not apply to the problem. We discuss this issue in more detail in Appendix A while discussing the space of federated online optimization problems, where we explicitly show (c.f., Figure 1 in Appendix A) how there is no contradiction between the theorems in this section and most existing results, which do show a benefit of collaboration with first-order oracles (Woodworth et al., 2020a;b).

4. Collaboration Helps with Bandit Feedback

In this section, we move our attention to the more challenging setting of using bandit (zeroth-order) feedback at each machine. We begin by studying one important application of problem (1), i.e., *federated adversarial linear bandits*. We then investigate the *more general problem of federated bandit convex optimization with two-point feedback* in the next section.

Algorithm 2: FEDPOSGD (η, δ) with one-point bandit feedback

1 Initialize $x_0^m = 0$ on all machines $m \in [M]$
2 **for** $t \in \{0, \ldots, KR-1\}$ **do**
3   **for** $m \in [M]$ ***in parallel*** **do**
4     $w_t^m = \mathbf{Proj}(x_t^m)$
5   Sample $u_t^m \sim Unif(\mathbb{S}_{d-1})$, i.e., a random unit vector
6   Query function f_t^m at point $w_t^{m,1} = w_t^m + \delta u_t^m$
7   **Incur loss** $f_t^m(w_t^m + \delta u_t^m)$
8   Compute stochastic gradient at point w_t^m as $g_t^m = df(w_t^m + \delta u_t^m)u_t^m/\delta$
9   **if** $(t+1) \mod K = 0$ **then**
10   **Communicate to server:** $(x_t^m - \eta \cdot g_t^m)$
11   On server $x_{t+1} \leftarrow \frac{1}{M}\sum_{m\in[M]}(x_t^m - \eta \cdot g_t^m)$
12   **Communicate to machine:** $x_{t+1}^m \leftarrow x_{t+1}$
13   **else**
14       $x_{t+1}^m \leftarrow x_t^m - \eta \cdot g_t^m$

Federated Adversarial Linear Bandits. One important application of problem (1) is federated linear bandits, which has received increasing attention in recent years. However, most existing works (Wang et al., 2020; Huang et al., 2021; Li & Wang, 2022; He et al., 2022) do not consider the more challenging adaptive adversaries, leaving it unclear whether collaboration can be beneficial in this scenario. Therefore, we propose to study federated adversarial linear bandits, a natural extension of single-agent adversarial linear bandits (Bubeck et al., 2012) to the federated optimization setting. Specifically, at each time $t \in [T]$, an agent $m \in [M]$ chooses an action $x_t^m \in \mathbb{R}^d$ while simultaneously environment picks $\beta_t^m \in \mathbb{B}_2(G) \subset \mathbb{R}^d$. Agent m then suffers the loss $f_t^m(x_t^m) = \langle \beta_t^m, x_t^m \rangle$. Our goal is to output a sequence of models $\{x_t^m\}_{t\in[T]}^{m\in[M]}$ that minimize the following expected regret

$$\mathbb{E}\left[\sum_{m,t}\langle \beta_t^m, x_t^m\rangle - \min_{\|x^\star\|_2 \le B}\sum_{m,t}\langle \beta_t^m, x^\star\rangle\right], \tag{7}$$

where the expectation is w.r.t. the randomness of the model selection. To solve this federated adversarial linear bandits problem, we propose a novel projected online stochastic gradient descent with one-point feedback algorithm, which we call FEDPOSGD, as illustrated in Algorithm 2.

At the core of Algorithm 2 is the gradient estimator g_t^m constructed using one-point feedback (see line 8). This approach is motivated by Flaxman et al. (2004), with the key distinction that our gradient estimator is calculated at the projected point $w_t^m = \mathbf{Proj}(x_t^m) = \arg\min_{\|w\|_2 \le B}\|w - x_t^m\|_2$ (see line 4). For the linear cost function $f_t^m(x) = \langle \beta_t^m, x\rangle$, we can show that g_t^m is an unbiased gradient estimator $\mathbb{E}_{u_t^m}[g_t^m] = \nabla f_t^m(x)$ with the variance (Hazan, 2016):

$$\mathbb{E}_{u_t^m}\left[\|g_t^m - \nabla f_t^m(x)\|_2^2\right] \preceq \left(d\|\beta_t^m\|_2 \cdot (\|x\|_2 + \delta)/\delta\right)^2.$$

Therefore, the projection step in Algorithm 2 can ensure the variance of our gradient estimator is bounded. However, the extra projection step can make it difficult to benefit from collaboration when we have gradient estimators from multiple agents. To address this issue, we propose to perform the gradient update in the unprojected space, i.e., $x_t^m - \eta \cdot g_t^m$ (see line 14), instead of the projected space, i.e., $\mathbf{Proj}(w_t^m - \eta \cdot g_t^m)$, which is motivated by the lazy mirror descent based methods (Nesterov, 2009; Bubeck et al., 2015; Yuan et al., 2021). We obtain the following guarantee for Algorithm 2 for federated adversarial linear bandits.

Theorem 4.1 (Regret Guarantee of Algorithm 2 for Federated Adversarial Linear Bandits)**.** *For the problem class $\mathcal{P}_{M,K,R}(\mathcal{F}_{lin}^{G,B}, \zeta)$, if we choose $\eta = \frac{B}{G\sqrt{T}} \cdot \min\left\{1, \frac{\sqrt{M}}{dB}, \frac{1}{\mathbb{1}_{K>1}\sqrt{dB}K^{1/4}}, \frac{\sqrt{G}}{\mathbb{1}_{K>1}\sqrt{\zeta K}}\right\}$ and $\delta = B$, the queried points $\{w_t^{m,1}\}_{t,m=1}^{T,M}$ of Algorithm 2 satisfy:*

$$\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\mathbb{E}\left[f_t^m(w_t^{m,1}) - f_t^m(x^\star)\right]$$
$$\preceq \frac{GB}{\sqrt{KR}} + \frac{GBd}{\sqrt{MKR}} + \mathbb{1}_{K>1}\cdot\left(\frac{GB\sqrt{d}}{K^{1/4}\sqrt{R}} + \frac{\sqrt{G\zeta}B}{\sqrt{R}}\right),$$

where $x^\star \in \arg\min_{x\in\mathbb{B}_2(B)}\sum_{t\in[KR]} f_t(x)$, and the expectation is w.r.t. the choice of function queries.

Implication of Theorem 4.1: When degenerated to the single-agent adversarial linear bandits, FEDPOSGD with $K = 1$, $M = 1$ achieves $\mathcal{O}\left(GBd/\sqrt{T}\right)$ average regret, which matches the optimal average regret (Dani et al., 2007; Hazan, 2016) using one-point feedback when we have unconstrained action space. When $K > 1$ and $M > 1$, we would like to compare our results to the non-collaborative baseline as we did in Section 3. More specifically, the non-collaborative baseline [3] can obtain the average regret of $\mathcal{O}\left(GBd/\sqrt{KR}\right)$. If we have $d \succeq \sqrt{K}$, FEDPOSGD outperforms the non-collaborative baseline. Furthermore, if we have $d \succeq \sqrt{KM}$, then the average regret of FEDPOSGD will be dominated by $\mathcal{O}\left(GBd/\sqrt{MKR}\right)$. As a result, FEDPOSGD achieves a *"linear speed-up"* in terms of the number of machines compared to the non-collaborative baseline. Although a smaller ζ will lead to a smaller regret bound, we cannot benefit from a small ζ. This is because the ζ dependent term $\sqrt{G\zeta}B/\sqrt{R}$ in the average regret bound will be

[3]For the non-collaborative baseline, we can run SCRIBLE (Hazan, 2016) independently on each machine or run Algorithm 1 using the gradient estimator proposed by Flaxman et al. (2004) with an extra projection step, i.e., $x_{t+1}^m = \mathbf{Proj}(x_{t+1}^m)$.

dominated by other terms when the problem dimension is large, i.e., $d \succeq \sqrt{K}$.

Limitations of Algorithm 2. Based on the average regret, we can see that: (1) the proposed one-point feedback algorithm, i.e., FEDPOSGD, requires *an extra projection step*; (2) the averaged regret bound of FEDPOSGD *increases linearly with the problem dimension* d; and (3) the average regret of FEDPOSGD *cannot benefit from the small heterogeneity* in the regime where FEDPOSGD outperforms the non-collaborative baseline. To address these issues, we propose to use a two-point feedback algorithm in the next section.

5. Better Rates with Two-Point Bandit Feedback

In this section, we study distributed bandit convex optimization with two-point feedback, i.e., at each time step, the machines can query the value (and not the full gradient) of their cost functions at two different points. We show an improved regret guarantee for general Lipschitz smooth functions and then specify the improvements for federated adversarial linear bandits.

The two-point feedback structure is useful for single-agent bandit convex optimization, as it can help attain the optimal horizon dependence in the regret (Duchi et al., 2015; Shamir, 2017) using simple algorithms. We consider a general convex cost function rather than the linear cost function discussed in the last section. We analyze the online variant of the FEDAVG or LOCAL-SGD algorithm, which is commonly used in the stochastic setting. We refer to this algorithm as FEDOSGD and describe it in Algorithm 3.

The key idea in FEDOSGD is using the gradient estimator constructed with two-point feedback (see line 7), originally proposed by Shamir (2017) and based on a similar estimator by Duchi et al. (2015). For a smoothed version of the function f_t^m, i.e., $\hat{f}_t^m(x) := \mathbb{E}_{u_t^m}[f_t^m(x+\delta u_t^m)]$, the gradient estimator g_t^m is an unbiased estimator $\mathbb{E}_{u_t^m}[g_t^m] = \nabla \hat{f}_t^m(x)$ with variance (c.f., Lemmas 3 and 5 in Shamir (2017)):

$$\mathbb{E}_{u_t^m}\left[\left\|g_t^m - \nabla \hat{f}_t^m(x)\right\|_2^2\right] \preceq dG^2,$$

where G is the Lipschitz parameter for f_t^m.

Equipped with this gradient estimator, we can prove the following guarantee for $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$ using FEDOSGD.

Theorem 5.1 (Better Bounds with Two-Point Feedback for Lipschitz Functions)**.** *For the problem class* $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$. *If we choose* $\eta = \frac{B}{G\sqrt{T}} \cdot \min\left\{1, \frac{\sqrt{M}}{\sqrt{d}}, \frac{1}{\mathbb{1}_{K>1}\sqrt{K}d^{1/4}}\right\}$, *and* $\delta = \frac{Bd^{1/4}}{\sqrt{R}}\left(1+\frac{d^{1/4}}{\sqrt{MK}}\right)$, *the queried points* $\{x_t^{m,j}\}_{t,m,j=1}^{T,M,2}$ *of Algorithm 3 satisfy:*

$$\frac{1}{2MKR}\sum_{t\in[KR],m\in[M],j\in[2]} \mathbb{E}\left[f_t^m(x_t^{m,j}) - f_t^m(x^\star)\right] \preceq \frac{GB}{\sqrt{KR}} + \frac{GB\sqrt{d}}{\sqrt{MKR}} + \mathbb{1}_{K>1}\cdot\frac{GBd^{1/4}}{\sqrt{R}},$$

where $x^\star \in \arg\min_{x\in\mathbb{B}_2(B)}\sum_{t\in[KR]} f_t(x)$*, and the expectation is w.r.t. the choice of function queries.*

Algorithm 3: FEDOSGD (η,δ) with two-point bandit feedback

1 Initialize $x_0^m = 0$ on all machines $m \in [M]$
2 **for** $t \in \{0,\ldots,KR-1\}$ **do**
3   **for** $m \in [M]$ ***in parallel*** **do**
4   Sample $u_t^m \sim Unif(\mathbb{S}_{d-1})$, i.e., a random unit vector
5   Query function f_t^m at points $(x_t^{m,1}, x_t^{m,2}) := (x_t^m + \delta u_t^m, x_t^m - \delta u_t^m)$
6   **Incur loss** $(f_t^m(x_t^m + \delta u_t^m) + f_t^m(x_t^m - \delta u_t^m))$
7   Compute stochastic gradient at point x_t^m as $g_t^m = \frac{d(f(x_t^m+\delta u_t^m) - f(x_t^m - \delta u_t^m))u_t^m}{2\delta}$
8   **if** $(t+1) \mod K = 0$ **then**
9   **Communicate to server:** $(x_t^m - \eta\cdot g_t^m)$
10   On server $x_{t+1} \leftarrow \frac{1}{M}\sum_{m\in[M]}(x_t^m - \eta\cdot g_t^m)$
11   **Communicate to machine:** $x_{t+1}^m \leftarrow x_{t+1}$
12   **else**
13            $x_{t+1}^m \leftarrow x_t^m - \eta\cdot g_t^m$

Implication of Theorem 5.1: When $K = 1$, the above average regret reduces to the first two terms, which are known to be tight for two-point bandit feedback (Duchi et al., 2015; Hazan, 2016) (see Appendix D), making FEDOSGD optimal. When $K > 1$, we would like to compare our results to the non-collaborative baseline as we did in Section 3. The non-collaborative baseline [4] attains the average regret of $\mathcal{O}\left(GB\sqrt{d}/\sqrt{KR}\right)$. Thus, as long as $d \succeq K^2$, FEDOSGD performs better than the non-collaborative baseline. Furthermore, if $d \succeq K^2M^2$, then the average regret of FEDOSGD is dominated by $\mathcal{O}\left(GB\sqrt{d}/\sqrt{MKR}\right)$. Therefore, FEDOSGD achieves a *"linear speed-up"* in the number of machines compared to the non-collaborative baseline. Unfortunately, the bound doesn't improve with smaller ζ.

Note that the Lipschitzness assumption is crucial for the two-point gradient estimator in Algorithm 3 to control the variance of the proposed gradient estimator. While there

[4] For the non-collaborative baseline, we run Algorithm 1 using the gradient estimator proposed by Shamir (2017).

are gradient estimators that do not require Lipschitzness or bounded gradients (Flaxman et al., 2004), they actually impose stronger assumptions such as bounded function values or necessitate extra projection steps (see the gradient estimator in Algorithm 2). To avoid making these assumptions, we instead focus on problems in $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,H,B}, \zeta, F_\star)$ rather than problems in $\mathcal{P}_{M,K,R}(\mathcal{F}^{H,B}, \zeta, F_\star)$.

Theorem 5.2 (Better Bounds with Two-Point Feedback for Smooth Functions)**.** *Consider the problem class $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,H,B}, \zeta, F_\star)$. If we choose appropriate η, δ (c.f., Lemma E.1 in Appendix E), the queried points $\{x_t^{m,j}\}_{t,m,j=1}^{T,M,2}$ of Algorithm 3 satisfy:*

$$\frac{1}{2MKR} \sum_{t\in[KR], m\in[M], j\in[2]} \mathbb{E}\left[f_t^m(x_t^{m,j}) - f_t^m(x^\star)\right]$$
$$\preceq \frac{HB^2}{KR} + \frac{\sqrt{HF_\star}B}{\sqrt{KR}} + \frac{GB}{\sqrt{KR}} + \frac{GB\sqrt{d}}{\sqrt{MKR}}$$
$$+ \mathbb{1}_{K>1} \cdot \min\left\{\frac{H^{1/3}B^{4/3}G^{2/3}d^{1/3}}{K^{1/3}R^{2/3}} + \frac{H^{1/3}B^{4/3}\zeta^{2/3}}{R^{2/3}}\right.$$
$$\left. + \frac{\sqrt{\zeta G}Bd^{1/4}}{K^{1/4}\sqrt{R}} + \frac{\zeta B}{\sqrt{R}}, \frac{GBd^{1/4}}{K^{1/4}\sqrt{R}} + \frac{\sqrt{G\zeta}B}{\sqrt{R}}\right\},$$

where $x^\star \in \arg\min_{x\in\mathbb{B}_2(B)} \sum_{t\in[KR]} f_t(x)$, and the expectation is w.r.t. the choice of function queries. Furthermore, we can obtain the same average regret as in Theorem 5.1 with the corresponding step size.

The above result is a bit technical, so to interpret it, we consider the simpler class $\mathcal{F}_{lin}^{G,B}$ of linear functions with bounded gradients. Linear functions are the "smoothest" Lipschitz functions as their smoothness constant $H = 0$. We can get the following guarantee (see Appendix G.1):

Corollary 5.3 (Better Bounds with Two-Point Feedback for Linear Functions)**.** *Consider the problem class $\mathcal{P}_{M,K,R}(\mathcal{F}_{lin}^{G,B}, \zeta, F_\star)$. If we choose the same η and δ as in Theorem 5.2, the queried points $\{x_t^{m,j}\}_{t,m,j=1}^{T,M,2}$ of Algorithm 3 satisfy:*

$$\frac{1}{2MKR} \sum_{t\in[KR], m\in[M], j\in[2]} \mathbb{E}\left[f_t^m(x_t^{m,j}) - f_t^m(x^\star)\right]$$
$$\preceq \frac{GB}{\sqrt{KR}} + \frac{GB\sqrt{d}}{\sqrt{MKR}} + \mathbb{1}_{K>1} \cdot \left(\frac{\sqrt{\zeta G}Bd^{1/4}}{K^{1/4}\sqrt{R}} + \frac{\zeta B}{\sqrt{R}}\right),$$

where $x^\star \in \arg\min_{x\in\mathbb{B}_2(B)} \sum_{t\in[KR]} f_t(x)$, and the expectation is w.r.t. the choice of function queries.

Implications of Theorem 5.2 and Corollary 5.3: Unlike general Lipschitz functions, the last two terms in the average regret bound for linear functions are zero when $\zeta = 0$, and the upper bound is smaller for smaller ζ. In fact, when $K = 1$ or $\zeta = 0$, the upper bound reduces to $\mathcal{O}\big(GB/\sqrt{KR} + GB\sqrt{d}/\sqrt{MKR}\big)$, which is optimal (c.f., Appendix B). These results show that FEDOSGD can benefit from small heterogeneity. More generally, when $K \leq \max\left\{1, G^2\zeta^2 d, G^2 d/(\zeta^2 M^2)\right\}$, then FEDOSGD again achieves the optimal average regret of $\mathcal{O}\big(GB/\sqrt{KR} + GB\sqrt{d}/\sqrt{MKR}\big)$. Recall that in this setting, the non-collaborative baseline [5] obtains an average regret of $\mathcal{O}(GB\sqrt{d}/\sqrt{KR})$. Thus, the benefit of collaboration through FEDOSGD again appears in high-dimensional problems in $\mathcal{P}_{M,K,R}(\mathcal{F}_{lin}^{G,B}, \zeta, F_\star)$ similar to what we discussed for $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B}, \zeta)$.

Single v/s Two-Point Feedback. For federated adversarial linear bandits, we can directly apply Algorithm 3, i.e., FEDOSGD, to solve it and achieve the average regret bound in Corollary 5.3 as follows:

$$\frac{GB}{\sqrt{KR}} + \frac{GB\sqrt{d}}{\sqrt{MKR}} + \mathbb{1}_{K>1} \cdot \left(\frac{\sqrt{\zeta G}Bd^{1/4}}{K^{1/4}\sqrt{R}} + \frac{\zeta B}{\sqrt{R}}\right).$$

Recall that we can also use the one-point feedback based Algorithm 2, i.e., FEDPOSGD, to get the following average regret bound for federated adversarial linear bandits:

$$\frac{GB}{\sqrt{KR}} + \frac{GBd}{\sqrt{MKR}} + \mathbb{1}_{K>1} \cdot \left(\frac{GB\sqrt{d}}{K^{1/4}\sqrt{R}} + \frac{\sqrt{G\zeta}B}{\sqrt{R}}\right).$$

According to these results, FEDOSGD significantly improves the average regret bound of FEDPOSGD in terms of the dependence on d and ζ. In addition, FEDOSGD does not require the extra projection step and can benefit from the small heterogeneity compared to FEDPOSGD. These results also imply that multi-point feedback can be beneficial in federated adversarial linear bandits.

6. Conclusion and Future Work

In this paper, we show that, in the setting of distributed bandit convex optimization against an adaptive adversary, the benefit of collaboration is very similar to the stochastic setting with first-order gradient information, where the collaboration is useful when: (i) there is stochasticity in the problem and (ii) the variance of the gradient estimator is *"high"* (Woodworth et al., 2021) and reduces with collaboration. There are several open questions and future research directions:

1. Does collaboration provably not help for the smaller class $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,H,B}, \zeta, F_\star)$ using algorithms with first-order information?
2. Is the final term tight in Theorems 5.1 and 5.2? We do not know any lower bounds in the intermittent communication setting against an adaptive adversary.
3. When K is large, but R is a fixed constant, the average regret of the non-collaborative baseline goes to zero, but

[5] The same baseline as for problems in $\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B}, \zeta)$.

our upper bounds for FEDOSGD do not. It is unclear if our analysis is loose or if we need another algorithm.

4. What structures in the problem can we further exploit to reduce communication (c.f., federated stochastic linear bandits (Huang et al., 2021))?

Acknowledgments

We thank the anonymous reviewers who helped us improve the writing of the paper. We would also like to thank Ayush Sekhari, John Duchi, and Ohad Shamir for several useful discussions. This research was partly supported by the NSF-BSF award 1718970, the NSF TRIPOD IDEAL award, and the NSF-Simons funded Collaboration on the Theoretical Foundations of Deep Learning.

References

Bubeck, S., Cesa-Bianchi, N., et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5 (1):1–122, 2012.

Bubeck, S. et al. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.

Bullins, B., Patel, K., Shamir, O., Srebro, N., and Woodworth, B. E. A stochastic newton algorithm for distributed convex optimization. *Advances in Neural Information Processing Systems*, 34, 2021.

Chen, M., Mathews, R., Ouyang, T., and Beaufays, F. Federated learning of out-of-vocabulary words. *arXiv preprint arXiv:1903.10635*, 2019.

Cotter, A., Shamir, O., Srebro, N., and Sridharan, K. Better mini-batch algorithms via accelerated gradient methods. *Advances in neural information processing systems*, 24, 2011.

Dai, S. and Meng, F. Addressing modern and practical challenges in machine learning: A survey of online federated and transfer learning. *arXiv preprint arXiv:2202.03070*, 2022.

Dani, V., Kakade, S. M., and Hayes, T. The price of bandit information for online optimization. *Advances in Neural Information Processing Systems*, 20, 2007.

Dekel, O., Gilad-Bachrach, R., Shamir, O., and Xiao, L. Optimal distributed online prediction using mini-batches. *Journal of Machine Learning Research*, 13(1), 2012.

Dieuleveut, A. and Patel, K. K. Communication trade-offs for local-sgd with large step size. *Advances in Neural Information Processing Systems*, 32, 2019.

Dubey, A. and Pentland, A. Differentially-private federated linear bandits. *Advances in Neural Information Processing Systems*, 33:6003–6014, 2020.

Duchi, J. C., Jordan, M. I., Wainwright, M. J., and Wibisono, A. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, 61(5):2788–2806, 2015.

Elbir, A. M., Soner, B., and Coleri, S. Federated learning in vehicular networks. *arXiv preprint arXiv:2006.01412*, 2020.

Flaxman, A. D., Kalai, A. T., and McMahan, H. B. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.

Gauthier, F., Gogineni, V. C., Werner, S., Huang, Y.-F., and Kuh, A. Resource-aware asynchronous online federated learning for nonlinear regression. In *ICC 2022-IEEE International Conference on Communications*, pp. 2828–2833. IEEE, 2022.

Glasgow, M. R., Yuan, H., and Ma, T. Sharp bounds for federated averaging (local sgd) and continuous perspective. In *International Conference on Artificial Intelligence and Statistics*, pp. 9050–9090. PMLR, 2022.

Gogineni, V. C., Werner, S., Huang, Y.-F., and Kuh, A. Communication-efficient online federated learning framework for nonlinear regression. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5228–5232. IEEE, 2022.

Google. Your voice amp; audio data stays private while google assistant improves, 2023. URL `https://support.google.com/assistant/answer/10176224?hl=en`.

Hao, K. How apple personalizes siri without hoovering up your data. *Technology Review*, 2020.

Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., and Ramage, D. Federated learning for mobile keyboard prediction. *arXiv preprint arXiv:1811.03604*, 2018.

Hartmann, F. Predicting text selections with federated learning, Nov 2021. URL `https://ai.googleblog.com/2021/11/predicting-[]text-[]selections-[]with.html`.

Hazan, E. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

He, J., Wang, T., Min, Y., and Gu, Q. A simple and provably efficient algorithm for asynchronous federated contextual linear bandits. In *Advances in Neural Information Processing Systems*, 2022.

Huang, R., Wu, W., Yang, J., and Shen, C. Federated linear contextual bandits. *Advances in Neural Information Processing Systems*, 34:27057–27068, 2021.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. Advances and open problems in federated learning. corr. *arXiv preprint arXiv:1912.04977*, 2019.

Khaled, A., Mishchenko, K., and Richtárik, P. Tighter theory for local sgd on identical and heterogeneous data. In *International Conference on Artificial Intelligence and Statistics*, pp. 4519–4529. PMLR, 2020.

Khan, F. K., Flanagan, A., Tan, K. E., Alamgir, Z., and Ammad-Ud-Din, M. A payload optimization method for federated recommender systems. In *Fifteenth ACM Conference on Recommender Systems*, pp. 432–442, 2021.

Koloskova, A., Loizou, N., Boreiri, S., Jaggi, M., and Stich, S. A unified theory of decentralized sgd with changing topology and local updates. In *International Conference on Machine Learning*, pp. 5381–5393. PMLR, 2020.

Kuh, A. Real time kernel learning for sensor networks using principles of federated learning. In *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 2089–2093. IEEE, 2021.

Li, C. and Wang, H. Asynchronous upper confidence bound algorithms for federated linear bandits. In *International Conference on Artificial Intelligence and Statistics*, pp. 6529–6553. PMLR, 2022.

Liang, F., Pan, W., and Ming, Z. Fedrec++: Lossless federated recommendation with explicit feedback. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 4224–4231, 2021.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. Communication-efficient learning of deep networks from decentralized data (2016). *arXiv preprint arXiv:1602.05629*, 2016.

Mitra, A., Hassani, H., and Pappas, G. J. Online federated learning. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 4083–4090. IEEE, 2021.

Nemirovski, A. Efficient methods in convex programming. *Lecture notes*, 1994.

Nesterov, Y. Primal-dual subgradient methods for convex problems. *Mathematical programming*, 120(1):221–259, 2009.

Nguyen, A., Do, T., Tran, M., Nguyen, B. X., Duong, C., Phan, T., Tjiputra, E., and Tran, Q. D. Deep federated learning for autonomous driving. In *2022 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1824–1830. IEEE, 2022.

Patel, K. K., Wang, L., Woodworth, B., Bullins, B., and Srebro, N. Towards optimal communication complexity in distributed non-convex optimization. In *Advances in Neural Information Processing Systems*, 2022.

Saha, A. and Tewari, A. Improved regret guarantees for online smooth convex optimization with bandit feedback. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 636–642. JMLR Workshop and Conference Proceedings, 2011.

Shamir, O. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *The Journal of Machine Learning Research*, 18(1):1703–1713, 2017.

Shamir, O., Srebro, N., and Zhang, T. Communication-efficient distributed optimization using an approximate newton-type method. In *International conference on machine learning*, pp. 1000–1008. PMLR, 2014.

Shi, C., Shen, C., and Yang, J. Federated multi-armed bandits with personalization. In *International Conference on Artificial Intelligence and Statistics*, pp. 2917–2925. PMLR, 2021.

Srebro, N., Sridharan, K., and Tewari, A. Optimistic rates for learning with a smooth loss. *arXiv preprint arXiv:1009.3896*, 2010.

Stich, S. U. Local sgd converges fast and communicates little. *arXiv preprint arXiv:1805.09767*, 2018.

Wang, J., Das, R., Joshi, G., Kale, S., Xu, Z., and Zhang, T. On the unreasonable effectiveness of federated averaging with heterogeneous data. *arXiv preprint arXiv:2206.04723*, 2022.

Wang, Y., Hu, J., Chen, X., and Wang, L. Distributed bandit learning: Near-optimal regret with efficient communication. In *International Conference on Learning Representations*, 2020. URL `https://openreview.net/forum?id=SJxZnR4YvB`.

Woodworth, B. The minimax complexity of distributed optimization. *arXiv preprint arXiv:2109.00534*, 2021.

Woodworth, B., Patel, K. K., Stich, S., Dai, Z., Bullins, B., Mcmahan, B., Shamir, O., and Srebro, N. Is local sgd better than minibatch sgd? In *International Conference on Machine Learning*, pp. 10334–10343. PMLR, 2020a.

Woodworth, B. E. and Srebro, N. An even more optimal stochastic optimization algorithm: minibatching and interpolation learning. *Advances in Neural Information Processing Systems*, 34:7333–7345, 2021.

Woodworth, B. E., Wang, J., Smith, A., McMahan, B., and Srebro, N. Graph oracle models, lower bounds, and gaps for parallel stochastic optimization. *Advances in neural information processing systems*, 31, 2018.

Woodworth, B. E., Patel, K. K., and Srebro, N. Minibatch vs local sgd for heterogeneous distributed learning. *Advances in Neural Information Processing Systems*, 33: 6281–6292, 2020b.

Woodworth, B. E., Bullins, B., Shamir, O., and Srebro, N. The min-max complexity of distributed stochastic convex optimization with intermittent communication. In *Conference on Learning Theory*, pp. 4386–4437. PMLR, 2021.

Yuan, H., Zaheer, M., and Reddi, S. Federated composite optimization. In *International Conference on Machine Learning*, pp. 12253–12266. PMLR, 2021.

Zhang, Y., Duchi, J., Jordan, M. I., and Wainwright, M. J. Information-theoretic lower bounds for distributed statistical estimation with communication constraints. *Advances in Neural Information Processing Systems*, 26, 2013.

Zhang, Y., Duchi, J., and Wainwright, M. Divide and conquer kernel ridge regression: A distributed algorithm with minimax optimal rates. *The Journal of Machine Learning Research*, 16(1):3299–3340, 2015.

Zinkevich, M., Weimer, M., Li, L., and Smola, A. Parallelized stochastic gradient descent. *Advances in neural information processing systems*, 23, 2010.

A. Comparison of Related Problems

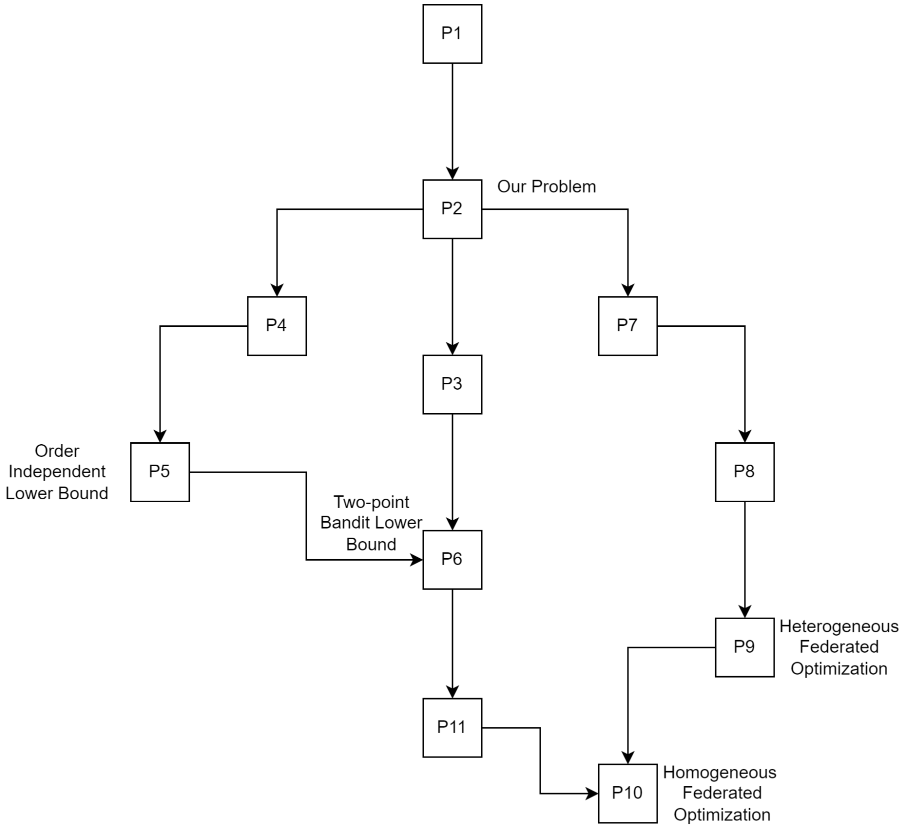

Figure 1. Summary of the problem space of federated online optimization. An arrow from the parent to child denotes that the child min-max problem is easier or has a lower min-max value than the parent problem. Note that to show that there is no benefit of collaboration for first-order algorithms for problem (P2), we use the lower bound construction for the problem (P5). The figure clarifies why this does not contradict the benefit of collaboration for problems (P9) and (P10), as they lie on a different path from the parent (P2).

First, we recall that the algorithms on each machine can depend on all the history it has seen. Since we are in the IC setting, at time t on machine m, the algorithm A's output can only depend on

$$\left(f_1^{1:M}, \dots, f_{\tau(t)}^{1:M}, f_{\tau(t)+1}^{m}, \dots, f_{t-1}^{m}\right),$$

where $\tau(t)$ is the last time step smaller than or equal to t where communication happened. We also assume that the algorithms can be randomized, i.e., at each time t on machine m, the output can also depend on a randomly drawn seed z_t^m. With that said, the hardest adversarial problem we can hope to solve for functions coming from some problem class $\mathcal{P} \subset \mathcal{F}^{\otimes MT}$ is

$$\min_{A\in\mathcal{A}} \mathbb{E}_A \left[\max_{P\in\mathcal{P}} \frac{1}{MT} \sum_{t\in[T],m\in[M]} f_t^m(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)} \frac{1}{MT} \sum_{t\in[T],m\in[M]} f_t^m(x^\star)\right]. \tag{\textbf{P1}}$$

For this problem, note that the max-player knows both the algorithm and the randomization of the min-player. Thus, the min-player doesn't gain from randomization at all. A simpler problem is the following, where the max-player knows the algorithm but not the random seeds

$$\min_{A\in\mathcal{A}} \max_{P\in\mathcal{P}} \mathbb{E}_A \left[\frac{1}{MT} \sum_{t\in[T],m\in[M]} f_t^m(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)} \frac{1}{MT} \sum_{t\in[T],m\in[M]} f_t^m(x^\star)\right]. \tag{\textbf{P2}}$$

Problem (P2) is usually considered when talking about randomized optimization algorithms. Note that randomization on the second player doesn't increase the min-max regret, so we can equivalently state (P2) with randomized max-players as follows

$$\min_{A\in\mathcal{A}}\max_{P\in\mathcal{P}}\mathbb{E}_{A,P}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x_t^m)-\min_{x^\star\in\mathbb{B}_2(B)}\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x^\star)\right]. \tag{P2}$$

However, making the randomization on the max-player explicit makes it easier to state the following easier version of the problem (P2) with a weaker comparator

$$\min_{A\in\mathcal{A}}\max_{P\in\mathcal{P}}\mathbb{E}_{A}\left(\mathbb{E}_P\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x_t^m)\right]-\min_{x^\star\in\mathbb{B}_2(B)}\mathbb{E}_P\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x^\star)\right]\right). \tag{P3}$$

The comparator in problem (P3) does not know the randomization of the max-player. This form of regret is common in multi-armed bandit literature and is often called *"pseudo-regret"*. In this paper, we upper bound (P2), thus also upper bounding (P3), which might be more relevant for the federated adversarial linear bandits problem. Note that we have only talked about the *"adaptive"* setting so far. We can also relax the problem (P2) by weakening the adversary. One way to do this is by requiring the functions to be the same across the machines, which leads to the following problem

$$\min_{A\in\mathcal{A}}\max_{\{f_t\in\mathcal{F}\}_{t\in[T]}^{\otimes M}\in\mathcal{P}}\mathbb{E}_{A}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t(x_t^m)-\min_{x^\star\in\mathbb{B}_2(B)}\frac{1}{T}\sum_{t\in[T]}f_t(x^\star)\right]. \tag{P4}$$

Note that if the functions across the machines are shared, then $\zeta=0$ in Assumption 1. Depending on the algorithm class, this problem may or may not be equivalent to a fully serial problem, as we showed in this paper. We can further simplify this problem by making the adversary stochastic

$$\min_{A\in\mathcal{A}}\max_{\mathcal{D}\in\Delta(\mathcal{F})}\mathbb{E}_{A,\{f_t\sim\mathcal{D}\}_{t\in[T]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t(x_t^m)-\min_{x^\star\in\mathbb{B}_2(B)}\frac{1}{T}\sum_{t\in[T]}f_t(x^\star)\right]. \tag{P5}$$

We can also simplify problem (P5) to have a weaker comparator or consider problem (P3) with a stochastic adversary

$$\min_{A\in\mathcal{A}}\max_{\mathcal{D}\in\Delta(\mathcal{F})}\mathbb{E}_{A,\{f_t\sim\mathcal{D}\}_{t\in[T]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}\mathbb{E}_{f_t\sim\mathcal{D}}\left[f_t(x_t^m)\right]\right]-\min_{x^\star\in\mathbb{B}_2(B)}\mathbb{E}_{f\sim\mathcal{D}}\left[f(x^\star)\right]. \tag{P6}$$

Recalling the definition of F_m, F this can be re-written as,

$$\min_{A\in\mathcal{A}}\max_{\mathcal{D}\in\Delta(\mathcal{F})}\mathbb{E}_{A,\{f_t\sim\mathcal{D}\}_{t\in[T]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}F(x_t^m)\right]-\min_{x^\star\in\mathbb{B}_2(B)}F(x^\star). \tag{P6}$$

Now let's relax (P2) directly to have stochastic adversaries, i.e., have fixed distributions on each machine. Note that this will require appropriately changing the problem classes' assumptions, as discussed in the remarks 2.2 and 2.3. To simplify the discussion, we assume that the problem class has no additional assumption and is just a selection of MKR functions from some class $\mathcal{F}$. With this simplification in mind, we can now relax $(P2)$ by picking the functions at machine m at time t from distribution $\mathcal{D}_m$. This leads to the following problem

$$\min_{A\in\mathcal{A}}\max_{\{\mathcal{D}_m\sim\Delta(\mathcal{F})\}_{m\in[M]}}\mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x_t^m)-\min_{x^\star\in\mathbb{B}_2(B)}\frac{1}{MT}\sum_{t\in[T],m\in[M]}f_t^m(x^\star)\right]. \tag{P7}$$

If we weaken the comparator for this problem and recall the definitions for $\{F_m := \mathbb{E}_{f\sim\mathcal{D}_m}[f]\}_{m\in[M]}$ and $F := \frac{1}{M}\sum_{m\in[M]}F_m$, we get the following problem

$$\min_{A\in\mathcal{A}}\max_{\{\mathcal{D}_m\sim\Delta(\mathcal{F})\}_{m\in[M]}}\mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]}F_m(x_t^m)\right]-\min_{x^\star\in\mathbb{B}_2(B)}F(x^\star). \tag{P8}$$

We note that problem (P8) is the regret minimization version of the usual heterogeneous federated optimization problem (McMahan et al., 2016; Woodworth et al., 2020b). To make the final connection to the usual federated optimization literature, we note that the problem becomes easier if the algorithm can look at all the functions before deciding which model to choose. In other words, it is harder to minimize regret in the online setting than to come up with one final retrospective model. This means we can simplify the problem (P8) to the following problem, where A outputs $\hat{x}$ after looking at all the functions

$$\min_{A\in\mathcal{A}} \max_{\{\mathcal{D}_m\sim\Delta(\mathcal{F})\}_{m\in[M]}} \mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left[\frac{1}{MT}\sum_{t\in[T],m\in[M]} F_m(\hat{x})\right] - \min_{x^\star\in\mathbb{B}_2(B)} F(x^\star). \tag{P9}$$

With some re-writing of the notation, this reduces the usual heterogeneous federated optimization problem (McMahan et al., 2016; Woodworth et al., 2020b)

$$\min_{A\in\mathcal{A}} \max_{\{\mathcal{D}_m\sim\Delta(\mathcal{F})\}_{m\in[M]}} \mathbb{E}_{A,\{f_t^m\sim\mathcal{D}_m\}_{t\in[T]}^{m\in[M]}}\left[F(\hat{x})\right] - \min_{x^\star\in\mathbb{B}_2(B)} F(x^\star). \tag{P9}$$

Assuming $\mathcal{D}_m = \mathcal{D}$ for all $m\in[M]$ this reduces to the usual homogeneous federated optimization problem (Woodworth et al., 2020a)

$$\min_{A\in\mathcal{A}} \max_{\mathcal{D}\in\Delta(\mathcal{F})} \mathbb{E}_{A,\{f_t^m\sim\mathcal{D}\}_{t\in[T]}^{m\in[M]}}\left[F(\hat{x})\right] - \min_{x^\star\in\mathbb{B}_2(B)} F(x^\star). \tag{P10}$$

Note that we can get a similar relaxation of the problem (P6) by converting regret minimization to find a final good solution. The problem will look as follows

$$\min_{A\in\mathcal{A}} \max_{\mathcal{D}\in\Delta(\mathcal{F})} \mathbb{E}_{A,\{f_t\sim\mathcal{D}\}_{t\in[T]}}\left[F(\hat{x})\right] - \min_{x^\star\in\mathbb{B}_2(B)} F(x^\star). \tag{P11}$$

The key difference between (P10) and (P11) is that $\hat{x}$ depends on MT v/s T samples respectively each case. This means (P10) is simpler than (P11). This concludes the discussion, and we summarize the comparisons between different min-max problems in Figure 1. Next, we discuss two relevant lower bounds that follow from this understanding of problem hierarchies.

B. Relevant Lower Bound Proofs

We would like to understand the lower bounds for the problem (P2) from the previous section, i.e., this paper's main quantity of interest for relevant problem and algorithm classes. Note that to lower bound (P2), we can lower bound any children problems in Figure 1. We first restate a well-known OCO sample complexity lower bound to show that

$$\mathcal{R}^1(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B},\zeta)) \cong \mathcal{R}^1(\mathcal{P}_{M,K,R}(\mathcal{F}^{G,B}_{lin},\zeta)) \cong \frac{GB}{\sqrt{KR}}. \tag{8}$$

To see this, we recall that $(P2) \succeq (P4) \succeq (P5)$ and then note that for the adversary in $(P5)$, $\zeta = 0$ by design as all the machines see the same function. Thus to prove a lower bound, it is sufficient to specify a distribution $\mathcal{D}\in\Delta(\mathcal{F}^{G,B}_{lin})\subset\Delta(\mathcal{F}^{G,B})$ such that for **any** sequence of models $\{x_t^m\}_{t\in[T]}^{m\in[M]}$,

$$\mathbb{E}_{\{f_t\}_{t\in[T]}}\left[\frac{1}{MT}\sum_{m,t} f_t(x_t^m) - \min_{x^\star\in\mathbb{B}_2(B)}\frac{1}{T}\sum_t f_t(x^\star)\right] \succeq \frac{GB}{\sqrt{T}}.$$

And one such easy construction is choosing $f_t(x) = \langle v_t, x\rangle$ where $v_t \sim \frac{G}{\sqrt{d}}\cdot Unif(\{+1,-1\}^d)$. This ensures that the first term on the L.H.S. in the above inequality is zero after taking the expectation no matter which model is picked. The minimizer of the second quantity is $x^\star$ is $-\frac{GB}{\sqrt{T}}$ (c.f., Theorem 3.2 (Hazan, 2016)), which gives the lower bound in equation (8). Note that the construction is linear and lies in the class $\mathcal{F}^{G,B}$, which is an order-optimal lower bound since it doesn't have anything to do with the form of the algorithm. Furthermore, to get equality in equation (8), we recall the non-collaborative OGD upper bound (c.f., Theorem 3.1 (Hazan, 2016)).

Next, we draw our attention to the problem class of smooth functions $\mathcal{P}_{M,K,R}(\mathcal{F}^{H,B},\zeta,F_\star)$ where we show that

$$\mathcal{R}^1(\mathcal{P}_{M,K,R}(\mathcal{F}^{H,B},\zeta,F_\star)) \cong \frac{HB^2}{T} + \frac{\sqrt{HF_\star}B}{\sqrt{T}}. \tag{9}$$

We use the same lower-bounding strategy mentioned above but instead lower bound (P10), and note that $(P2) \succeq (P4) \succeq (P5) \succeq (P6) \succeq (P10)$. This already makes $\zeta = 0$ since our attack is coordinated. To lower bound (P10), we use the construction and distribution as used in the proof of Theorem 4 by Woodworth & Srebro (2021), which is a sample complexity lower bound that only depends on T, i.e., the number of samples observed from $\mathcal{D}$. For the upper bound, we use Theorem 3 by Srebro et al. (2010) proved in the same setting. Even though the upper bounds are first-order algorithms, the lower bounds are order-independent, i.e., sample complexity lower bounds.

Finally, we'd like to prove a lower bound for the problem class $\mathcal{P}_{MKR}(\mathcal{F}^{G,B})$ with two-point bandit feedback. In particular, we want to show that

$$\mathcal{R}^{0,2}(\mathcal{P}_{MKR}(\mathcal{F}^{G,B}), \zeta) \succeq \frac{GB}{\sqrt{KR}} + \frac{GB\sqrt{d}}{\sqrt{MKR}}. \tag{10}$$

To prove this, we'd use the reduction $(P2) \succeq (P4) \succeq (P5) \succeq (P6)$. Then we note for the problem (P6), $\zeta = 0$, and using 2-point feedback, we get in total $2MKR$ function value accesses to $\mathcal{D}$. We can directly use the lower bound in Proposition 2 by Duchi et al. (2015) for the problem (P6) for $2M$ points of feedback and KR iterations. Combined with the order-independent lower bound proved previously, this proves the required result.

C. Proof of Theorem 4.1

In this section, we provide the proofs of Theorem 4.1. We first introduce several notations, which will be used in our analysis. Let $d(x, y) = \|x\|_2^2/2 - \|\hat{y}\|_2^2/2 - \langle y, x - \hat{y}\rangle$, where $\|x\|_2 \le B$ and $\hat{y}$ is the projected point of y in to the ℓ_2-norm ball with radius B. We have the following holds

$$d(x, y) \ge \frac{1}{2}\|x - \hat{y}\|_2^2. \tag{11}$$

This is due to the following: if $\|y\| \le B$, (11) clearly holds. If $\|y\| > B$, we have

$$d(x, y) - \frac{1}{2}\|x - \hat{y}\|_2^2 = \langle x - \hat{y}, \hat{y} - y\rangle \ge \langle \hat{y} - \hat{y}, \hat{y} - y\rangle = 0,$$

where the inequality is due to the fact that $\hat{y} - y = (1 - \|y\|_2/B)\hat{y}$ lies in the opposite direction of $\hat{y}$, and $x = \hat{y}$ will minimize the inner product. Now, we are ready to prove the regret of Algorithm 2.

Proof. Define the following notations

$$\bar{x}_t = \frac{1}{M}\sum_{m=1}^{M} x_t^m, \;\; \bar{w}_t = \text{Proj}(\bar{x}_t), \;\; w_t^m = \text{Proj}(x_t^m).$$

We have

$$\begin{aligned}
d(x^\star, \bar{x}_{t+1}) &= \frac{1}{2}\|x^\star\|_2^2 - \frac{1}{2}\|\bar{w}_{t+1}\|_2^2 - \langle \bar{x}_{t+1}, x^\star - \bar{w}_{t+1}\rangle \\
&= \frac{1}{2}\|x^\star\|_2^2 - \frac{1}{2}\|\bar{w}_{t+1}\|_2^2 - \langle \bar{x}_t - \eta\frac{1}{M}\sum_{m=1}^{M} g_t^m, x^\star - \bar{w}_{t+1}\rangle \\
&= \underbrace{\frac{1}{2}\|x^\star\|_2^2 - \frac{1}{2}\|\bar{w}_{t+1}\|_2^2 - \langle \bar{x}_t, x^\star - \bar{w}_{t+1}\rangle}_{I_1} \\
&\quad \underbrace{-\eta\frac{1}{M}\sum_{m=1}^{M}\langle g_t^m, \bar{w}_{t+1} - x^\star\rangle}_{I_2},
\end{aligned} \tag{12}$$

where the second equality comes from the updating rule of Algorithm 2. For the term I_1, we have

$$I_1 = \frac{1}{2}\|x^\star\|_2^2 - \frac{1}{2}\|\bar{w}_{t+1}\|_2^2 - \langle \bar{x}_t, x^\star - \bar{w}_{t+1}\rangle$$

$$
\begin{aligned}
&= \frac{1}{2}\|x^\star\|_2^2 - \frac{1}{2}\|\bar{w}_t\|_2^2 - \langle \bar{x}_t, x^\star - \bar{w}_t\rangle - \langle \bar{x}_t, \bar{w}_t - \bar{w}_{t+1}\rangle \\
&\quad - \frac{1}{2}\|\bar{w}_{t+1}\|_2^2 + \frac{1}{2}\|\bar{w}_t\|_2^2 \\
&= d(x^\star, \bar{x}_t) - d(\bar{w}_{t+1}, \bar{x}_t) \\
&\leq d(x^\star, \bar{x}_t) - \frac{1}{2}\|\bar{w}_{t+1} - \bar{w}_t\|_2^2,
\end{aligned}
$$

where the last inequality is due to (11). For the term I_2, we have

$$
\begin{aligned}
I_2 &= -\eta \frac{1}{M}\sum_{m=1}^M \langle g_t^m, \bar{w}_{t+1} - x^\star\rangle \\
&= \underbrace{-\eta \frac{1}{M}\sum_{m=1}^M \langle g_t^m - \nabla f_t^m(w_t^m), \bar{w}_{t+1} - x^\star\rangle}_{I_{21}} \underbrace{-\eta \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m), \bar{w}_{t+1} - \bar{x}^\star\rangle}_{I_{22}}.
\end{aligned}
$$

For the term I_{21}, we have

$$
\begin{aligned}
\mathbb{E}_t[I_{21}] &= \eta \mathbb{E}_t \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m) - g_t^m, \bar{w}_{t+1} - x^\star\rangle \\
&= \eta \mathbb{E}_t \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m) - g_t^m, \bar{w}_{t+1} - \bar{w}_t\rangle \\
&\leq \eta \mathbb{E}_t \left\| \frac{1}{M}\sum_{m=1}^M (\nabla f_t^m(w_t^m) - g_t^m)\right\|_2 \cdot \|\bar{w}_{t+1} - \bar{w}_t\|_2 \\
&\leq \eta \frac{\sigma}{\sqrt{M}} \mathbb{E}_t \|\bar{w}_{t+1} - \bar{w}_t\|_2.
\end{aligned}
$$

Thus we have

$$
\mathbb{E}[I_{21}] \leq \eta \frac{\sigma}{\sqrt{M}} \mathbb{E}\|\bar{w}_{t+1} - \bar{w}_t\|_2.
$$

For the term I_{22}, we have

$$
\begin{aligned}
I_{22} &= -\eta \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m), w_t^m - \bar{x}^\star\rangle - \eta \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m), \bar{w}_t - w_t^m\rangle \\
&\quad - \eta \frac{1}{M}\sum_{m=1}^M \langle \nabla f_t^m(w_t^m), \bar{w}_{t+1} - \bar{w}_t\rangle \\
&\leq -\eta \frac{1}{M}\sum_{m=1}^M (f_t^m(w_t^m) - f_t^m(x^\star) + \eta \frac{1}{M}\sum_{m=1}^M \|\nabla f_t^m(w_t^m)\|_2 \cdot \|\bar{w}_t - w_t^m\|_2 \\
&\quad + \eta^2 \left\| \frac{1}{M}\sum_{m=1}^M \nabla f_t^m(w_t^m)\right\|_2^2 + \frac{1}{4}\|\bar{w}_{t+1} - \bar{w}_t\|_2^2
\end{aligned}
$$

Therefore, combining (12) and the upper bound of I_1 and I_2, we have

$$
\begin{aligned}
\mathbb{E} d(x^\star, \bar{x}_{t+1}) &\leq \mathbb{E} d(x^\star, \bar{x}_t) - \frac{1}{4}\mathbb{E}\|\bar{w}_{t+1} - \bar{w}_t\|_2^2 + \eta \frac{\sigma}{\sqrt{M}} \mathbb{E}\|\bar{w}_{t+1} - \bar{w}_t\|_2 \\
&\quad - \eta \frac{1}{M}\sum_{m=1}^M \mathbb{E}(f_t^m(w_t^m) - f_t^m(x^\star) + \eta \frac{1}{M}\sum_{m=1}^M \mathbb{E}\|\nabla f_t^m(w_t^m)\|_2 \cdot \|\bar{w}_t - w_t^m\|_2
\end{aligned}
$$

$$+\eta^2\mathbb{E}\left\|\frac{1}{M}\sum_{m=1}^{M}\nabla f_t^m(w_t^m)\right\|_2^2.$$

Therefore, we have

$$\eta\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}(f_t^m(w_t^m)-f_t^m(x^\star)\le\mathbb{E}d(x^\star,\bar{x}_t)-\mathbb{E}d(x^\star,\bar{x}_{t+1})+\eta^2\frac{\sigma^2}{M}$$
$$+\eta\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\|\nabla f_t^m(w_t^m)\|_2\cdot\|\bar{w}_t-w_t^m\|_2$$
$$+\eta^2\mathbb{E}\left\|\frac{1}{M}\sum_{m=1}^{M}\nabla f_t^m(w_t^m)\right\|_2^2.$$

In addition, we have

$$\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\|\bar{w}_t-w_t^m\|_2\le\frac{1}{M}\sum_{m=1}^{M}\mathbb{E}\|\bar{x}_t-x_t^m\|_2\le 2\eta(\sigma\sqrt{K}+\zeta K).,$$

where the last inequality is due to the linear function and follows the similar proofs of Lemma 8 in Woodworth et al. (2020b). Thus, we can obtain (the indicator function comes from the fact that if $K=1$, there would be no consensus error)

$$\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[f_t^m(w_t^m)-f_t^m(x^\star)\right]\le\frac{1}{\eta}\left(\mathbb{E}d(x^\star,\bar{x}_t)-\mathbb{E}d(x^\star,\bar{x}_{t+1})\right)$$
$$+\eta\left(G^2+\frac{\sigma^2}{M}\right)+\mathbb{1}_{K>1}\cdot 2G(\sigma\sqrt{K}+\zeta K)\eta.$$

Since $\mathbb{E}d(x^\star,\bar{x}_T)\ge\mathbb{E}\|x^\star-\bar{w}_T\|_2^2/2\ge 0$ and $\mathbb{E}d(x^\star,\bar{x}_0)=\|x^\star\|_2^2/2$, summing the above inequality over t, we can get

$$\frac{1}{M}\sum_{t\in[KR],m\in[M]}\mathbb{E}\left[f_t^m(w_t^m)-f_t^m(x^\star)\right]\preceq\frac{B^2}{\eta}+\eta\left(G^2+\frac{\sigma^2}{M}+\mathbb{1}_{K>1}\cdot G(\sigma\sqrt{K}+\zeta K)\right)T.$$

If we choose η such that

$$\eta=\frac{B}{G\sqrt{T}}\cdot\min\left\{1,\frac{G\sqrt{M}}{\sigma},\frac{\sqrt{G}}{\mathbb{1}_{K>1}\sqrt{\sigma}K^{1/4}},\frac{\sqrt{G}}{\mathbb{1}_{K>1}\sqrt{\zeta K}}\right\},$$

We can get

$$\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\mathbb{E}\left[f_t^m(w_t^m)-f_t^m(x^\star)\right]\preceq\frac{GB}{\sqrt{KR}}+\frac{\sigma B}{\sqrt{MKR}}+\mathbb{1}_{K>1}\cdot\left(\frac{\sqrt{G\sigma}B}{K^{1/4}\sqrt{R}}+\frac{\sqrt{G\zeta}B}{\sqrt{R}}\right).$$

To get the regret, we just need to notice that we have the linear function, and thus we have: the smoothed function $\hat{f}=f$ and $\mathbb{E}f_t^m(w_t^m)=\mathbb{E}f_t^m(w_t^m+\delta u_t^m)$, where the expectation is over u_t^m. Furthermore, $\|g_t^m\|_2^2=d^2\left(f_t^m(w_t^m+\delta u_t^m)\right)^2\le d^2G^2(B+\delta)^2/\delta^2\le 4d^2G^2$, where the last inequality is due the choice of $\delta=B$. Since $\mathbb{E}g_t^m=\nabla\hat{f}_t^m(w_t^m)$ and $\mathbb{E}\|g_t^m-\nabla\hat{f}_t^m(w_t^m)\|_2^2\le\mathbb{E}\|g_t^m\|_2^2$ Therefore, we can plug in $\sigma^2=4d^2G^2$ to get our regret. □

D. Proof of Theorem 5.1

In this section and the next one, we consider access to a first-order stochastic oracle as an intermediate step before considering the zeroth-order oracle. Specifically, each machine has access to a **stochastic gradient** g_t^m of f_t^m at point x_t^m, such that it is unbiased and has bounded variance, i.e., for all $x\in\mathcal{X}$,

$$\mathbb{E}[g_t^m(x_t^m)|x_t^m]=\nabla f_t^m(x_t^m)\text{ and }\mathbb{E}\left[\|g_t^m(x_t^m)-\nabla f_t^m(x_t^m)\|_2^2\,|x_t^m\right]\le\sigma^2.$$

In Algorithm 3, we constructed a particular stochastic gradient estimator at x_t^m with $\sigma^2=G^2d$. We can define the corresponding problem class $\mathcal{P}_{M,K,R}^{1,\sigma}(\mathcal{F}^{G,B},\zeta)$ when the agents can access a stochastic first-order oracle. We have the following lemma about this problem class:

Lemma D.1. *Consider the problem class $\mathcal{P}^{1,\sigma}_{M,K,R}(\mathcal{F}^{G,B},\zeta)$. If we choose $\eta = \frac{B}{G\sqrt{T}}\cdot\min\left\{1, \frac{G\sqrt{M}}{\sigma}, \frac{\sqrt{G}}{\mathbb{1}_{K>1}\sqrt{\sigma K}}, \frac{1}{\mathbb{1}_{K>1}\sqrt{K}}\right\}$, then the models $\{x_t^m\}_{t,m=1}^{T,M}$ of Algorithm 3 satisfy the following guarantee:*

$$\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\mathbb{E}\left[f_t^m(x_t^m)-f_t^m(x^\star)\right] \preceq \frac{GB}{\sqrt{KR}}+\frac{\sigma B}{\sqrt{MKR}}+\mathbb{1}_{K>1}\cdot\left(\frac{\sqrt{\sigma G}B}{\sqrt{R}}+\frac{GB}{\sqrt{R}}\right),$$

where $x^\star \in \arg\min_{x\in\mathbb{R}^d}\sum_{t\in[KR]} f_t(x)$, and the expectation is w.r.t. the stochastic gradients.

Remark D.2. Note that when $K=1$, the upper bound in Lemma D.1 reduces to the first two terms, both of which are known to be optimal due to lower bounds in the stochastic setting, i.e., against a stochastic online adversary (Nemirovski, 1994; Hazan, 2016). We now use this lemma to guarantee bandit two-point feedback oracles for the same function class. We recall that one can obtain a stochastic gradient for a *"smoothed-version"* $\hat{f}$ of a Lipschitz function f at any point $x\in\mathcal{X}$, using two function value calls to f around the point x (Shamir, 2017; Duchi et al., 2015).

With this lemma, we can prove Theorem 5.1.

Proof of Theorem 5.1. First, we consider smoothed functions

$$\hat{f}_t^m(x) := \mathbb{E}_{u\sim Unif(S_{d-1})}[f_t^m(x+\delta u)],$$

for some $\delta>0$ and S_{d-1} denoting the euclidean unit sphere. Based on the gradient estimator proposed by Shamir (2017) (which can be implemented with two-point bandit feedback) and Lemma D.1, we can get the following regret guarantee (noting that $\sigma \le c_1\sqrt{d}G$ for a numerical constant c_1, c.f., (Shamir, 2017)):

$$\mathbb{E}\left[\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\hat{f}_t^m(\hat{x}_t^m)\right]-\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\hat{f}_t^m(x^\star)\preceq\frac{GB}{\sqrt{KR}}+\frac{GB\sqrt{d}}{\sqrt{MKR}}+\mathbb{1}_{K>1}\cdot\frac{GBd^{1/4}}{\sqrt{R}},$$

where the expectation is w.r.t. the stochasticity in the stochastic gradient estimator. To transform this into a regret guarantee for f we need to account for two things:

1. The difference between the smoothed function $\hat{f}$ and the original function f. This is easy to handle because both these functions are pointwise close, i.e., $\sup_{x\in\mathcal{X}}|f(x)-\hat{f}(x)|\le G\delta$.

2. The difference between the points $\hat{x}_t^m$ at which the stochastic gradient is computed for $\hat{f}_t^m$ and the actual points $x_t^{m,1}$ and $x_t^{m,2}$ on which we incur regret while making zeroth-order queries to f_t^m. This is also easy to handle because due to the definition of the estimator, $x_t^{m,1}, x_t^{m,1}\in B_\delta(\hat{x}_t^m)$, where $B_\delta(x)$ is the L_2 ball of radius δ around x.

In light of the last two observations, the average regret between the smoothed and original functions only differs by a factor of $2G\delta$, i.e.,

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2MKR}\sum_{t\in[KR],m\in[M],j\in[2]}f_t^m(x_t^{m,j})\right]-\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}f_t^m(x^\star)\\
&\preceq G\delta+\frac{GB}{\sqrt{KR}}+\frac{GB\sqrt{d}}{\sqrt{MKR}}+\mathbb{1}_{K>1}\cdot\frac{GBd^{1/4}}{\sqrt{R}}\\
&\preceq\frac{GB}{\sqrt{KR}}+\frac{GB\sqrt{d}}{\sqrt{MKR}}+\mathbb{1}_{K>1}\cdot\frac{GBd^{1/4}}{\sqrt{R}},
\end{aligned}$$

where the last inequality is due to the choice of δ such that $\delta\preceq\frac{Bd^{1/4}}{\sqrt{R}}\left(1+\frac{d^{1/4}}{\sqrt{MK}}\right)$. $\square$

E. Proof of Theorem 5.2

Similar to before, we start by looking at $\mathcal{P}^{1,\sigma}_{M,K,R}(\mathcal{F}^{G,H,B},\zeta,F_\star)$. We have the following lemma.

Lemma E.1. *Consider the problem class $\mathcal{P}^{1,\sigma}_{M,K,R}(\mathcal{F}^{G,H,B},\zeta,F_\star)$. The models $\{x_t^m\}_{t,m=1}^{T,M}$ of Algorithm 3 with appropriate η (specified in the proof) satisfy the following regret guarantee:*

$$\frac{1}{MKR}\sum_{t\in[KR],m\in[M]}\mathbb{E}\left[f_t^m(x_t^m)-f_t^m(x^\star)\right]\preceq\frac{HB^2}{KR}+\frac{\sigma B}{\sqrt{MKR}}+\min\left\{\frac{GB}{\sqrt{KR}},\frac{\sqrt{HF_\star}B}{\sqrt{KR}}\right\},$$
$$+\mathbb{1}_{K>1}\cdot\min\left\{\frac{H^{1/3}B^{4/3}\sigma^{2/3}}{K^{1/3}R^{2/3}}+\frac{H^{1/3}B^{4/3}\zeta^{2/3}}{R^{2/3}}+\frac{\sqrt{\zeta\sigma}B}{K^{1/4}\sqrt{R}}+\frac{\zeta B}{\sqrt{R}},\right.$$
$$\left.\frac{\sqrt{G\sigma}B}{K^{1/4}\sqrt{R}}+\frac{\sqrt{G\zeta}B}{\sqrt{R}}\right\},$$

where $x^\star\in\arg\min_{x\in\mathbb{R}^d}\sum_{t\in[KR]}f_t(x)$, and the expectation is w.r.t. the stochastic gradients. The models also satisfy the guarantee of Lemma D.1 with the same step-size.

Proof of Theorem 5.2. Given Lemma E.1, it is now straightforward to prove Theorem 5.2 similar to the proof for Theorem 5.1 by replacing σ^2 with G^2d ad choosing small enough δ such that $G\delta\cong$ the r.h.s. of the theorem statement. □

F. Proof of Lemma D.1

In this section, we prove Lemma D.1.

Proof of Lemma D.1. Consider any time step $t\in[KR]$ and define ghost iterate $\bar{x}_t=\frac{1}{M}\sum_{m\in[M]}x_t^m$ (which not might actually get computed). If $K=1$, the machines calculate the stochastic gradient at the same point, $\bar{x}_t$. Then using the update rule of Algorithm 3, we can get the following:

$$\begin{aligned}
\mathbb{E}_t\left[\|\bar{x}_{t+1}-x^\star\|_2^2\right]&=\mathbb{E}_t\left[\left\|\bar{x}_t-\frac{\eta_t}{M}\sum_{m\in[M]}\nabla f_t^m(x_t^m)-x^\star+\frac{\eta_t}{M}\sum_{m=1}^{M}\left(\nabla f_t^m(x_t^m)-g_t^m(x_t^m)\right)\right\|_2^2\right]\\
&=\|\bar{x}_t-x^\star\|_2^2+\frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2-\frac{2\eta_t}{M}\sum_{m\in[M]}\langle\bar{x}_t-x^\star,\nabla f_t^m(x_t^m)\rangle+\frac{\eta_t^2\sigma^2}{M}\\
&=\|\bar{x}_t-x^\star\|_2^2+\frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2-\frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m-x^\star,\nabla f_t^m(x_t^m)\rangle\\
&\quad+\mathbb{1}_{K>1}\cdot\frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m-\bar{x}_t,\nabla f_t^m(x_t^m)\rangle+\frac{\eta_t^2\sigma^2}{M}\\
&\leq\|\bar{x}_t-x^\star\|_2^2+\frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2-\frac{2\eta_t}{M}\sum_{m\in[M]}\left(f_t^m(x_t^m)-f_t^m(x^\star)\right)\\
&\quad+\mathbb{1}_{K>1}\cdot\frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m-\bar{x}_t,\nabla f_t^m(x_t^m)\rangle+\frac{\eta_t^2\sigma^2}{M},
\end{aligned}$$

where $\mathbb{E}_t$ is the expectation conditioned on the filtration at time t under which x_t^m's are measurable, and the last inequality is due to the convexity of each function. Re-arranging this leads to

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M]}\left(f_t^m(x_t^m)-f_t^m(x^\star)\right) &\le \frac{1}{2\eta_t}\left(\|\bar{x}_t-x^\star\|_2^2-\mathbb{E}_t\left[\|\bar{x}_{t+1}-x^\star\|_2^2\right]\right)+\frac{\eta_t}{2M^2}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2 \\
&\quad +\mathbb{1}_{K>1}\cdot\frac{1}{M}\sum_{m\in[M]}\mathbb{E}_t\left\langle x_t^m-\bar{x}_t,\nabla f_t^m(x_t^m)\right\rangle+\frac{\eta_t\sigma^2}{2M} \\
&\le \frac{1}{2\eta_t}\left(\|\bar{x}_t-x^\star\|_2^2-\mathbb{E}_t\left[\|\bar{x}_{t+1}-x^\star\|_2^2\right]\right)+\frac{\eta_t}{2}\left(G^2+\frac{\sigma^2}{M}\right) \\
&\quad +\mathbb{1}_{K>1}\cdot\frac{G}{M}\sum_{m\in[M]}\mathbb{E}\left[\|x_t^m-\bar{x}_t\|_2\right].
\end{aligned}
\tag{13}
$$

The last inequality comes from each function's G-Lipschitzness. For the last term in (13), we can upper bound it similar to Lemma 8 in Woodworth et al. (2020b) to get that

$$
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[\|x_t^m-\bar{x}_t\|_2\right]\le 2(\sigma+G)K\eta. \tag{14}
$$

Plugging (14) into (13) and choosing a constant step-size η, and taking full expectation we get

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[f_t^m(x_t^m)-f_t^m(x^\star)\right] &\le \frac{1}{2\eta}\left(\left\|\mathbb{E}\left[\bar{x}_t-x^\star\right]^2\right\|_2-\mathbb{E}\left[\|\bar{x}_{t+1}-x^\star\|_2^2\right]\right) \\
&\quad +\frac{\eta}{2}\left(G^2+\frac{\sigma^2}{M}\right)+\mathbb{1}_{K>1}\cdot 2G(\sigma+G)K\eta.
\end{aligned}
$$

Summing this over time $t\in[KR]$ we get,

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M],t\in[T]}\mathbb{E}\left[f_t^m(x_t^m)-f_t^m(x^\star)\right] &\preceq \frac{\|\bar{x}_0-x^\star\|_2^2}{\eta}+\eta\left(G^2+\frac{\sigma^2}{M}+\mathbb{1}_{K>1}\cdot\sigma GK+\mathbb{1}_{K>1}\cdot\zeta GK\right)T \\
&\preceq \frac{B^2}{\eta}+\eta\left(G^2+\frac{\sigma^2}{M}+\mathbb{1}_{K>1}\cdot\sigma GK+\mathbb{1}_{K>1}\cdot G^2K\right)T.
\end{aligned}
$$

Finally choosing,

$$
\eta=\frac{B}{G\sqrt{T}}\cdot\min\left\{1,\frac{G\sqrt{M}}{\sigma},\frac{\sqrt{G}}{\mathbb{1}_{K>1}\sqrt{\sigma K}},\frac{1}{\mathbb{1}_{K>1}\sqrt{K}}\right\},
$$

we can obtain,

$$
\frac{1}{M}\sum_{m\in[M],t\in[T]}\mathbb{E}\left[f_t^m(x_t^m)-f_t^m(x^\star)\right]\preceq GB\sqrt{T}+\mathbb{1}_{K>1}\cdot\sqrt{\sigma G}B\sqrt{KT}+\mathbb{1}_{K>1}\cdot GB\sqrt{KT}+\frac{\sigma B\sqrt{T}}{\sqrt{M}}. \tag{15}
$$

Dividing by KR finishes the proof. □

G. Proof of Lemma E.1

In this section, we provide the proofs for Lemma E.1.

Proof of Lemma E.1. Consider any time step $t\in[KR]$ and define ghost iterate $\bar{x}_t=\frac{1}{M}\sum_{m\in[M]}x_t^m$ (which not might actually get computed). Then using the update rule of Algorithm 3, we can get:

$$
\mathbb{E}_t\left[\|\bar{x}_{t+1}-x^\star\|_2^2\right]=\mathbb{E}_t\left[\left\|\bar{x}_t-\frac{\eta_t}{M}\sum_{m\in[M]}\nabla f_t^m(x_t^m)-x^\star+\frac{\eta_t}{M}\sum_{m=1}^{M}\left(\nabla f_t^m(x_t^m)-g_t^m(x_t^m)\right)\right\|_2^2\right],
$$

$$
\begin{aligned}
&= \|\bar{x}_t - x^\star\|_2^2 + \frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]} \nabla f_t^m(x_t^m)\right\|_2^2 - \frac{2\eta_t}{M}\sum_{m\in[M]}\langle \bar{x}_t - x^\star, \nabla f_t^m(x_t^m)\rangle + \frac{\eta_t^2\sigma^2}{M} \\
&= \|\bar{x}_t - x^\star\|_2^2 + \frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]} \nabla f_t^m(x_t^m)\right\|_2^2 - \frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m - x^\star, \nabla f_t^m(x_t^m)\rangle \\
&\quad + \mathbb{1}_{K>1}\cdot\frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m)\rangle + \frac{\eta_t^2\sigma^2}{M} \\
&\le \|\bar{x}_t - x^\star\|_2^2 + \frac{\eta_t^2}{M^2}\left\|\sum_{m\in[M]} \nabla f_t^m(x_t^m)\right\|_2^2 - \frac{2\eta_t}{M}\sum_{m\in[M]}\left(f_t^m(x_t^m) - f_t^m(x^\star)\right) \\
&\quad + \mathbb{1}_{K>1}\cdot\frac{2\eta_t}{M}\sum_{m\in[M]}\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m)\rangle + \frac{\eta_t^2\sigma^2}{M},
\end{aligned}
$$

where $\mathbb{E}_t$ is the expectation taken with respect to the filtration at time t, and the last line comes from the convexity of each function. Re-arranging this and taking expectation gives leads to

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left(f_t^m(x_t^m) - f_t^m(x^\star)\right) &\le \frac{1}{2\eta_t}\left(\mathbb{E}\|\bar{x}_t - x^\star\|_2^2 - \mathbb{E}\left[\|\bar{x}_{t+1} - x^\star\|_2^2\right]\right) + \frac{\eta_t}{2M^2}\mathbb{E}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2 \\
&\quad + \mathbb{1}_{K>1}\cdot\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m)\rangle + \frac{\eta_t\sigma^2}{2M}
\end{aligned} \tag{16}
$$

Bounding the blue term. We consider two different ways to bound the term. First note that similar to Lemma D.1 we can just use the following bound,

$$
\frac{\eta_t}{2M^2}\mathbb{E}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2 \le \frac{\eta_t G^2}{2} \tag{17}
$$

However, since we also have smoothness, we can use the self-bounding property (c.f., Lemma 4.1 (Srebro et al., 2010)) to get,

$$
\frac{\eta_t}{2M^2}\mathbb{E}\left\|\sum_{m\in[M]}\nabla f_t^m(x_t^m)\right\|_2^2 \le \frac{\eta_t H}{2M}\sum_{m\in[M]}\left(f_t^m(x_t^m) - f_t^m(x^\star)\right) + \frac{\eta_t H}{2M}\sum_{m\in[M]} f_t^m(x^\star) \tag{18}
$$

Bounding the red term. We will bound the term in three different ways. Similar to lemma D.1, we can bound the term after taking expectation and then bounding the consensus term similar to Lemma 8 in Woodworth et al. (2020b) as follows,

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m)\rangle\right] &\le \frac{G}{M}\sum_{m\in[M]}\mathbb{E}\left[\|x_t^m - \bar{x}_t\|_2\right] \\
&\le 2G(\sigma + G)\sum_{t'=\tau(t)}^{\tau(t)+K-1}\eta_{t'},
\end{aligned} \tag{19}
$$

where $\tau(t)$ maps t to the last time step when communication happens. Alternatively, we can use smoothness as follows after assuming $\eta_t \le 1/2H$,

$$
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m)\rangle\right] = \frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[\langle x_t^m - \bar{x}_t, \nabla f_t^m(x_t^m) - \nabla f_t(\bar{x}_t)\rangle\right],
$$

$$
\begin{aligned}
&\leq \sqrt{\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left\|x_t^m-\bar{x}_t\right\|_2^2}\sqrt{\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left\|\nabla f_t^m(x_t^m)-\nabla f_t(\bar{x}_t)\right\|_2^2},\\
&\leq \sqrt{\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left\|x_t^m-\bar{x}_t\right\|_2^2}\sqrt{\frac{2}{M}\sum_{m\in[M]}H^2\mathbb{E}\left\|x_t^m-\bar{x}_t\right\|_2^2+2\zeta^2},\\
&\leq \frac{2H}{M}\sum_{m\in[M]}\mathbb{E}\left\|x_t^m-\bar{x}_t\right\|_2^2+2\zeta\sqrt{\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left\|x_t^m-\bar{x}_t\right\|_2^2},\\
&\leq 2\eta_t^2H(\sigma^2K+\zeta^2K^2)+2\eta_t\zeta(\sigma\sqrt{K}+\zeta K),
\end{aligned}
\tag{20}
$$

where we used lemma 8 from Woodworth et al. (2020b) in the last inequality. We can also use the lipschitzness and smoothness assumption together with a constant step size $\eta < 1/2H$ to obtain,

$$
\begin{aligned}
\frac{1}{M}\sum_{m\in[M]}\mathbb{E}\left[\langle x_t^m-\bar{x}_t,\nabla f_t^m(x_t^m)\rangle\right] &\leq \frac{G}{M}\sum_{m\in[M]}\mathbb{E}\left[\left\|x_t^m-\bar{x}_t\right\|_2\right]\\
&\leq \eta G(\sigma\sqrt{K}+\zeta K).
\end{aligned}
\tag{21}
$$

Combining everything. After using a constant step-size η, summing (16) over time, we can use the upper bound of the red and blue terms in different ways. If we plug in (17) and (19) we recover the guarantee of lemma D.1. This is not surprising because $\mathcal{F}^{G,H,B}\subseteq\mathcal{F}^{G,B}$. Combining the upper bounds in all other combinations assuming $\eta<\frac{1}{2H}$, we can show the following upper bound

$$
\begin{aligned}
\frac{Reg(M,K,R)}{KR} \preceq\ & \frac{HB^2}{KR}+\frac{\sigma B}{\sqrt{MKR}}+\min\left\{\frac{GB}{\sqrt{KR}},\frac{\sqrt{HF_\star}B}{\sqrt{KR}}\right\},\\
&+\mathbb{1}_{K>1}\min\left\{\frac{H^{1/3}B^{4/3}\sigma^{2/3}}{K^{1/3}R^{2/3}}+\frac{H^{1/3}B^{4/3}\zeta^{2/3}}{R^{2/3}}+\frac{\sqrt{\zeta\sigma}B}{K^{1/4}\sqrt{R}}+\frac{\zeta B}{\sqrt{R}},\frac{\sqrt{G\sigma}B}{K^{1/4}\sqrt{R}}+\frac{\sqrt{G\zeta}B}{\sqrt{R}}\right\},
\end{aligned}
$$

where we used step size,

$$
\begin{aligned}
\eta = \min&\left\{\frac{1}{2H},\frac{B\sqrt{M}}{\sigma\sqrt{KR}},\max\left\{\frac{B}{G\sqrt{KR}},\frac{B}{\sqrt{HF_\star KR}}\right\},\right.\\
&\frac{1}{\mathbb{1}_{K>1}}\cdot\max\left\{\min\left\{\frac{B^{2/3}}{H^{1/3}\sigma^{2/3}K^{2/3}R^{1/3}},\frac{B^{2/3}}{H^{1/3}\zeta^{2/3}KR^{1/3}},\frac{B}{K^{3/4}\sqrt{\zeta\sigma R}},\frac{B}{\zeta K\sqrt{R}}\right\},\right.\\
&\left.\left.\min\left\{\frac{B}{K^{3/4}\sqrt{G\sigma R}},\frac{B}{K\sqrt{\zeta GR}}\right\}\right\}\right\}
\end{aligned}
$$

This finishes the proof. □

G.1. Modifying the Proof for Federated Adversarial Linear Bandits

To prove the guarantee for the adversarial linear bandits, we first note that the self-bounding property can't be used anymore as the functions are not non-negative. Thus we proceed with the lemma's proof with the following changes:

- We don't prove the additional upper bound in (18) for blue term.
- While upper bounding the red term in (20), we set $H=0$ and use this single bound for the red term.

After making these changes, combining all the terms, and tuning the learning rate, we recover the correct lemma for federated adversarial linear bandits.