(Nearly) Optimal Algorithms for Private Online Learning in Full-information and Bandit Settings

Adam Smith*
Pennsylvania State University
asmith@cse.psu.edu

Abhradeep Thakurta†
Stanford University and
Microsoft Research Silicon Valley Campus
b-abhrag@microsoft.com

Abstract

We give differentially private algorithms for a large class of online learning algorithms, in both the full information and bandit settings. Our algorithms aim to minimize a *convex* loss function which is a sum of smaller convex loss terms, one for each data point. To design our algorithms, we modify the popular *mirror descent* approach, or rather a variant called *follow the approximate leader*.

The technique leads to the first nonprivate algorithms for private online learning in the bandit setting. In the full information setting, our algorithms improve over the regret bounds of previous work (due to Dwork, Naor, Pitassi and Rothblum (2010) and Jain, Kothari and Thakurta (2012)). In many cases, our algorithms (in both settings) match the dependence on the input length, T, of the optimal nonprivate regret bounds up to logarithmic factors in T. Our algorithms require logarithmic space and update time.

1 Introduction

This paper looks at the information leaked by online learning algorithms, and seeks to design accurate learning algorithms with rigorous privacy guarantees – that is, algorithms that provably leak very little about individual inputs.

Even the output of offline (batch) learning algorithms can leak private information. The dual form of a support vector machine's solution, for example, is described in terms of a small number of exact data points, revealing these individuals' data in the clear. Considerable effort has been devoted to designing batch learning algorithms satisfying *differential privacy* (a rigorous notion of privacy that emerged from the cryptography literature [DMNS06, Dwo06]), for example [BDMN05, KLN$^+$08, CM08, CMS11, Smi11, KST12, JT13, DJW13].

In this work we provide a general technique for making a large class of online learning algorithms differentially private, in both the full information and bandit settings. Our technique applies to algorithms that aim to minimize a *convex* loss function which is a sum of smaller convex loss terms, one for each data point. We modify the popular *mirror descent* approach (or rather a variant called *follow the approximate leader*) [Sha11, HAK07].

In most cases, the modified algorithms provide similar accuracy guarantees to their nonprivate counterparts, with a small (logarithmic in the stream length) blowup in space and time complexity.

Online (Convex) Learning: We begin with the *full information* setting. Consider an algorithm that receives a stream of inputs $F = \langle f_1, ..., f_T \rangle$, each corresponding to one individual's data. We interpret each input as a loss function on a parameter space $\mathcal{C}$ (for example, it might be one term

*Supported by NSF awards #0941553 and #0747294.
†Supported by Sloan Foundation fellowship and Microsoft Research.

in a convex program such as the one for logistic regression). The algorithm's goal is to output a sequence of parameter estimates $w_1, w_2, ...$, with each w_t in $\mathcal{C}$, that roughly minimizes the errors $\sum_t f_t(w_t)$. The difficulty for the algorithm is that it computes w_t based only on $f_1, ..., f_{t-1}$. We seek to minimize the *a posteriori regret*,

$$Regret(T) = \sum_{t=1}^{T} f_t(w_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} f_t(w) \tag{1}$$

In the *bandit* setting, the input to the algorithms consists only of $f_1(w_1), f_2(w_2), ...$ That is, at each time step t, the algorithm learns only the cost $f_{t-1}(w_{t-1})$ of the choice w_{t-1} it made at the previous time step, rather than the full cost function f_{t-1}.

We consider three types of adversarial input selection: An *oblivious* adversary selects the input stream $f_1, ..., f_T$ ahead of time, based on knowledge of the algorithm but not of the algorithm's random coins. A *(strongly) adaptive* adversary selects f_t based on the output so far $w_1, w_2, ..., w_t$ (but not on the algorithm's internal random coins).

Both the full-information and bandit settings are extensively studied in the literature (see, e.g., [Sha11, BCB12] for recent surveys). Most of this effort has been spent on online learning problems are *convex*, meaning that the loss functions f_t are convex (in w) and the parameter set $\mathcal{C} \subseteq \mathbb{R}^p$ is a convex set (note that one can typically "convexify" the parameter space by randomization). The problem dimension p is the dimension of the ambient space containing $\mathcal{C}$.

We consider various restrictions on the cost functions, such as Lipschitz continuity and strong convexity. A function $f : \mathcal{C} \to \mathbb{R}$ is L-Lipschitz with respect to the ℓ_2 metric if $|f(x) - f(y)| \leq L\|x - y\|_2$ for all $x, y \in \mathcal{C}$. Equivalently, for every $x \in \mathcal{C}^0$ (the interior of $\mathcal{C}$) and every subgradient $z \in \partial f(x)$, we have $\|z\|_2 \leq L$. (Recall that z is a subgradient of f at x if the function $\tilde{f}(y) = f(x) + \langle z, y - x \rangle$ is a lower bound for f on all of $\mathcal{C}$. If f is convex, then a subgradient exists at every point, and the subgradient is unique if and only if f is differentiable at that point.) The function f is H-strongly convex w.r.t. ℓ_2 if for every $y \in \mathcal{C}$, we can bound f below on $\mathcal{C}$ by a quadratic function of the form $\tilde{f}(y) = f(x) + \langle z, y - x \rangle + \frac{H}{2}\|y - x\|_2^2$. If f is twice differentiable, H-strong convexity is equivalent to the requirement that all eigenvalues of $\nabla^2 f(w)$ be at least H for all $w \in \mathcal{C}$.

We denote by $\mathcal{D}$ the set of allowable cost functions; the input sequence thus lies in $\mathcal{D}^T$.

Differential Privacy, and Challenges for Privacy in the Online Setting: We seek to design online learning algorithms that satisfy *differential privacy* [DMNS06, Dwo06], which ensures that the amount of information an adversary learns about a particular cost function f_t in the function sequence F is almost independent of its presence or absence in F. Each f_t can be thought as private information belonging to an individual. The appropriate notion of privacy here is when the entire sequence of outputs of the algorithms $(\hat{w}_1, ..., \hat{w}_T)$ is revealed to an attacker (the *continual observation* setting [DNPR10]). Formally, we say two input sequences $F, F' \in \mathcal{D}^T$ are *neighbors* if they differ only in one entry (say, replacing f_t by f_t').

Definition 2 (Differential privacy [DMNS06, Dwo06, DNPR10]). *A randomized algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private if for every two neighboring sequences $F, F' \in \mathcal{D}^T$, and for every event $\mathcal{O}$ in the output space $\mathcal{C}^T$,*

$$\Pr[\mathcal{A}(F) \in \mathcal{O}] \leq e^{\epsilon} \Pr[\mathcal{A}(F') \in \mathcal{O}] + \delta. \tag{2}$$

If δ is zero, then we simply say $\mathcal{A}$ is ϵ-differentially private.

Here $\mathcal{A}(F)$ refers to the entire sequence of outputs produced by the algorithm during its execution.[1] Our protocols all satisfy ϵ-differential privacy (that is, with $\delta = 0$). We include δ in the definition for comparison with previous work.

[1] As defined, differential privacy requires indistinguishable outputs only for nonadaptively chosen sequences (that is, sequences where the inputs at time t are fixed ahead of time and do not depend on the outputs at times $1, ..., t-1$). The algorithms in our paper (and in previous work) in fact satisfy a stronger *adaptive* variant, in which an adversary selects the input online as the computation proceeds. When $\delta = 0$, the nonadaptive and adaptive variants are equivalent [DNPR10]. Moreover, protocols based on "randomized response" or the "tree-based sum" protocol of [DNPR10, CSS10] are adaptively secure, even when $\delta > 0$. We do not define the adaptive variant here explicitly, but we use it implicitly when proving privacy.

Differential privacy provides meaningful guarantees in against an attacker who has access to considerable side information: the attacker learns the same things about someone whether or not their data were actually used (see [KS08, DN10, KM12] for further discussion).

Differential privacy is particularly challenging to analyze for online learning algorithms, since a change in a single input at the beginning of the sequence may affect outputs at all future times in ways that are hard to predict. For example, a popular algorithm for online learning is *online gradient descent*: at each time step, the parameter is updated as $w_{t+1} = \Pi_{\mathcal{C}}(w_{t-1} - \eta_t \nabla f_{t-1}(w_{t-1}))$, where $\Pi_{\mathcal{C}}(x)$ the nearest point to x in $\mathcal{C}$, and $\eta_t > 0$ is a parameter called the learning rate. A change in an input f_i (replacing it with f_i') leads to changes in all subsequent outputs $w_{i+1}, w_{i+2}, \ldots$, roughly pushing them in the direction of $\nabla f_i(w_i) - \nabla f_i'(w_i)$. The effect is amplified by the fact that the gradient of subsequent functions $f_{i+1}, f_{i+2}, \ldots$ will be evaluated at different points in the two streams.

Previous Approaches: Despite the challenges, there are several results on differentially private online learning. A special case, "learning from experts" in the full information setting, was discussed in the seminal paper of Dwork, Naor, Pitassi and Rothblum [DNPR10] on privacy under continual observation. In this case, the set of available actions is the simplex $\Delta(\{1, \ldots, p\})$ and the functions f_i are linear with coefficients in $\{0, 1\}$ (that is, $f_t(w) = \langle w, c_t \rangle$ where $c_t \in \{0, 1\}^p$). Their algorithm guarantees a weaker notion of privacy than the one we consider[2] but, when adapted to our stronger setting, it yields a regret bound of $O(p\sqrt{T}/\epsilon)$.

Jain, Kothari and Thakurta [JKT12] defined the general problem of private online learning, and gave algorithms for learning convex functions over convex domains in the full information setting. They gave algorithms that satisfy (ϵ, δ)-differential privacy with $\delta > 0$ (our algorithms satisfy the stronger variant with $\delta = 0$). Specifically, their algorithms have regret $\tilde{O}(\sqrt{T}\log(1/\delta)/\epsilon)$ for Lipshitz-bounded, strongly convex cost functions and $\tilde{O}(T^{2/3}\log(1/\delta)/\epsilon)$ for general Lipshitz convex costs. The idea of [JKT12] for learning strongly convex functions is to bound the sensitivity of the entire vector of outputs $w_1, w_2, \ldots$ to a change in one input (roughly, they show that when f_i is changed, a subsequent output w_j changes by $O(1/|j - i|)$).

Unfortunately, the regret bounds obtained by previous work remain far from the best nonprivate bounds. [Zin03] gave an algorithm with regret $O(\sqrt{T})$ for general Lipshitz functions, assuming L and the diameter $\|\mathcal{C}\|_2$ of $\mathcal{C}$ are constants. $\Omega(\sqrt{T})$ regret is necessary (see, e.g., [HAK07]), so the dependence on T of [Zin03] is tight. When cost functions in F are H-strongly convex for constant H, then the regret can be improved to $O(\log T)$ [HAK07], which is also tight. In this work, we give new algorithms that match these nonprivate bounds' dependence on T, up to $(\mathrm{poly}\log T)/\epsilon$ factors.

We note that [JKT12] give one algorithm for a specific strongly convex problem, online linear regression, with regret $\mathrm{poly}(\log T)$. One can view that algorithm as a special case of our results.

We are not aware of any previous work on privacy in the bandit setting. One might expect that bandit learning algorithms are *easier* to make private, since they access data in a much more limited way. However, even nonprivate algorithms for bandit learning are very delicate, and private versions had until now proved elusive.

Our Results: In this work we provide a technique for making a large class of online learning algorithms differentially private, in both the full information and bandit settings. In both cases, the idea is to search for algorithms whose decisions at time t depend only on previous time steps through a *sum* of observations made at times $1, 2, \ldots, t$. Specifically, our algorithms work by measuring the gradient $\nabla f_t(w_t)$ when f_t is learned, and maintaining a differentially private running sum of the gradients observed so far. We maintain this sum using the tree-based sum protocol of [DNPR10, CSS10]. We then show that a class of learning algorithms known collectively as *follow the approximate leader* (the version we use is due to [HAK07]) can be run given only these noisy sums, and that their regret can be bounded even when these sums are inaccurate.

Our algorithms can be run with space $O(\log T)$, and require $O(\log T)$ running time at each step.

[2] Specifically, Dwork et al. [DNPR10] provide single-entry-level privacy, in the sense that a neighboring data set may only differ in one entry of the cost vector for one round. In contrast, we allow the entire cost vector to change at one round. Hiding that larger set of possible changes is more difficult, so our algorithms also satisfy the weaker notion of Dwork et al.

Our contributions for the full information setting and their relation to previous work is summarized in Table 1. Our main algorithm, for strongly convex functions, achieves regret $O(\frac{\log^{2.5} T}{\epsilon})$, ignoring factors of the dimension p, Lipschitz continuity L and strong convexity H. When strong convexity is not guaranteed, we use regularization to ensure it (similar to what is done in nonprivate settings, e.g. [Sha11]). Setting parameters carefully, we get regret of $O(\sqrt{\frac{T \log^{2.5} T}{\epsilon}})$. These bounds essentially match the nonprivate lower bounds of $\Omega(\log T)$ and $\Omega(\sqrt{T})$, respectively.

The results in the full information setting apply even when the input stream is chosen adaptively as a function of the algorithm's choices at previous time steps. In the bandit setting, we distinguish between oblivious and adaptive adversaries.

Furthermore, in the bandit setting, we assume that $\mathcal{C}$ is sandwiched between two concentric L_2-balls of radii r and R (where $r < R$). We also assume that for all $w \in \mathcal{C}$, $|f_t(w)| \leq B$ for all $t \in [T]$. Similar assumption were made in [FKM05, ADX10].

Our results are summarized in Table 2. For most of the settings we consider, we match the dependence on T of the best nonprivate algorithm, though generally not the dependence on the dimension p.

Function class	Previous private upper bound.	Our algorithm	Nonprivate lower bound
Learning with experts (linear functions over $\mathcal{C} = \Delta(\{1, ..., p\})$	$\tilde{O}(p\sqrt{T}/\epsilon)$ [DNPR10]	$O(\sqrt{pT \log^{2.5} T}/\epsilon)$	$\Omega(\sqrt{T \log p})$
Lipshitz	$\tilde{O}(\sqrt{p}T^{2/3} \log(1/\delta)/\epsilon)$ [JKT12]	$O(\sqrt{pT \log^{2.5} T}/\epsilon)$	$\Omega(\sqrt{T})$
Lipshitz and strongly convex	$\tilde{O}(\sqrt{pT} \log^2(1/\delta)/\epsilon)$ [JKT12]	$O(p \log^{2.5} T/\epsilon)$	$\Omega(\log T)$

Table 1: Regret bounds for online learning in the full information setting. Bounds in lines 2 and 3 hide the (polynomial) dependencies on parameters L, H. Notation $\tilde{O}(\cdot)$ hides $poly(\log(T))$ factors.

Function class	Our result	Best nonprivate bound
Learning with experts (linear functions over $\mathcal{C} = \Delta(\{1, ..., p\})$	$\tilde{O}(pT^{3/4}/\epsilon)$	$O(\sqrt{T})$ [AHR08]
Lipschitz	$\tilde{O}(pT^{3/4}/\epsilon)$	$O(pT^{3/4})$ [FKM05]
Lipschitz and strongly convex (Adaptive)	$\tilde{O}(pT^{3/4}/\epsilon)$	$O(p^{2/3}T^{3/4})$[ADX10]
Lipschitz and strongly convex (Oblivious)	$\tilde{O}(pT^{2/3}/\epsilon)$	$O(p^{2/3}T^{2/3})$[ADX10]

Table 2: Regret bounds for online learning in the bandit setting. In all these settings, the best known nonprivate *lower* bound is $\sqrt{T}$. The $\tilde{O}(\cdot)$ notation hides $poly \log$ factors in T. Bounds hide polynomial dependencies on L, H, r and R.

In the remainder of the text, we refer to appendices for many of the details of algorithms and proofs. The appendices can be found in the "Supplementary Materials" associated to this paper.

2 Private Online Learning: Full-information Setting

In this section we adapt the *Follow The Approximate Leader* (FTAL) algorithm of [HAK07] to design a differentially private variant. Our modified algorithm, which we call *Private Follow The*

Approximate Leader (PFTAL), needs a new regret analysis as we have to deal with randomness due to differential privacy.

2.1 Private Follow The Approximate Leader (PFTAL) with Strongly Convex Costs

Algorithm 1 Differentially Private Follow the Approximate Leader (PFTAL)

Input: Cost functions: $\langle f_1, \cdots, f_T \rangle$ (in an online sequence), strong convexity parameter: H, Lipschitz constant: L, convex set: $\mathcal{C} \subseteq \mathbb{R}^p$ and privacy parameter: ϵ.

1: $\hat{w}_1 \leftarrow$ Any vector from $\mathcal{C}$. **Output** $\hat{w}_1$.

2: Pass $\bigtriangledown f_1(\hat{w}_1)$, L_2-bound L and privacy parameter ϵ to the *tree based aggregation protocol* and receive the current partial sum in $\hat{v}_1$.

3: **for** time steps $t \in \{1, \cdots, T-1\}$ **do**

4: $\quad \hat{w}_{t+1} \leftarrow \arg\min_{w \in \mathcal{C}} \langle \hat{v}_t, w \rangle + \frac{H}{2} \sum_{\tau=1}^{t} \|w - \hat{w}_\tau\|_2^2$. **Output** $\hat{w}_t$.

5: $\quad$ Pass $\bigtriangledown f_{t+1}(\hat{w}_{t+1})$, L_2-bound L and privacy parameter ϵ to the *tree-based protocol* (Algorithm 2) and receive the current partial sum in $\hat{v}_{t+1}$.

6: **end for**

The main idea in PFTAL algorithm is to execute the well-known Follow The Leader algorithm (FTL) algorithm [Han57] using quadratic approximations $\tilde{f}_1, \cdots, \tilde{f}_T$ of the cost functions $f_1, \cdots, f_T$. Roughly, at every time step $(t+1)$, PFTAL outputs a vector w that approximately minimizes the sum of the approximations $\tilde{f}_1, \cdots, \tilde{f}_t$ over the convex set $\mathcal{C}$.

Let $\hat{w}_1, \cdots, \hat{w}_t$ be the sequence of outputs produced in the first t time steps, and let f_t be the cost-function at step t. Consider the following quadratic approximation to f_t (as in [HAK07]). Define

$$\tilde{f}_t(w) = f_t(\hat{w}_t) + \langle \bigtriangledown f_t(\hat{w}_t), w - \hat{w}_t \rangle + \frac{H}{2}\|w - \hat{w}_t\|_2^2 \tag{3}$$

where H is the strong convexity parameter. Notice that f_t and $\tilde{f}_t$ have the same value and gradient at $\hat{w}_t$ (that is, $f_t(\hat{w}_t) = \tilde{f}_t(\hat{w}_t)$ and $\bigtriangledown f_t(\hat{w}_t) = \bigtriangledown \tilde{f}_t(\hat{w}_t)$). Moreover, $\tilde{f}_t$ is a lower bound for f_t everywhere on $\mathcal{C}$.

Let $\tilde{w}_{t+1} = \arg\min_{w \in \mathcal{C}} \sum_{\tau=1}^{t} \tilde{f}_\tau(w)$ be the "leader" corresponding to the cost functions $\tilde{f}_1, \cdots, \tilde{f}_t$. Minimizing the sum of $\tilde{f}_t(w)$ is the same as minimizing the sum of $\tilde{f}_t(w) - f_t(\hat{w}_t)$, since subtracting a constant term won't change the minimizer. We can thus write $\tilde{w}_{t+1}$ as

$$\tilde{w}_{t+1} = \arg\min_{w \in \mathcal{C}} \langle \sum_{\tau=1}^{t} \bigtriangledown f_t(\hat{w}_\tau), w \rangle + \frac{H}{2} \sum_{\tau=1}^{t} \|w - \hat{w}_\tau\|_2^2 \tag{4}$$

Suppose, $\hat{w}_1, \cdots, \hat{w}_t$ have been released so far. To release a private approximation to $\tilde{w}_{t+1}$, it suffices to approximate $v_{t+1} = \sum_{\tau=1}^{t} \bigtriangledown f_t(\hat{w}_\tau)$ while ensuring differential privacy. If we fix the previously released information $\hat{w}_\tau$, then changing any one cost function will only change one of the summands in v_{t+1}.

With the above observation, we abstract out the following problem: Given a set of vectors $z_1, \cdots, z_T \in \mathbb{R}^p$, compute all the partial sums $v_t = \sum_{\tau=1}^{t} z_\tau$, while preserving privacy. This problem is well studied in the privacy literature. Assuming each z_t has L_2-norm of at most L', the following *tree-based aggregation* scheme will ensure that in expectation, the noise (in terms of L_2-error) in each of v_t is $O\left(pL' \log^{1.5} T/\epsilon\right)$ and the whole sequence $v_1, \cdots, v_T$ is ϵ-differentially private. We now describe the tree-based scheme.

Tree-based Aggregation [DNPR10, CSS10]: Consider a complete binary tree. The leaf nodes are the vectors $z_1, \cdots, z_T$. (For the ease of exposition, assume T to be a power of two. In general, we can work with the smallest power of two greater than T). Each internal node in the tree stores the sum of all the leaves in its sub-tree. In a differentially private version of this tree, we ensure that each node's sub-tree sum is $(\epsilon/\log_2 T)$-differentially private, by adding a noise vector $b \in \mathbb{R}^p$

whose L_2-norm is Gamma distributed and has standard deviation $O(\frac{\sqrt{p}L' \log T}{\epsilon})$. Since each z_t only affects $\log_2 T$ nodes in the tree, by the *composition property* [DMNS06], the complete tree will be ϵ-differentially private. Moreover, the algorithm's error in estimating any partial sum $v_t = \sum_{\tau=1}^{t} z_\tau$ grows as $O(\frac{\sqrt{p}L' \log^2 T}{\epsilon})$, since one can compute v_t from at most $\log T$ nodes in the tree. A formal description of the tree based aggregation scheme in given in Appendix A.

Now we complete the PFTAL algorithm by computing the private version $\hat{w}_{t+1}$ of $\tilde{w}_{t+1}$ in (4) as the minimizer of the perturbed loss function:

$$\hat{w}_{t+1} = \arg\min_{w \in \mathcal{C}} \langle \hat{v}_t, w \rangle + \frac{H}{2} \sum_{\tau=1}^{t} \|w - \hat{w}_\tau\|_2^2 \tag{5}$$

Here $\hat{v}_t$ is the noisy version of v_t, computed using the tree-based aggregation scheme. A formal description of the algorithm is given in Algorithm 1.

Note on space complexity: For simplicity, in the description of tree based aggregation scheme (Algorithm 2 in Appendix A) we maintain the complete binary tree. However, it is not hard to show at any time step t, it suffices to keep track of the vectors (of partial sums) in the path from z_t to the root of the tree. So, the amount of space required by the algorithm is $O(\log T)$.

2.1.1 Privacy and Utility Guarantees for PFTAL (Algorithm 1)

In this section we provide the privacy and regret guarantees for the PFTAL algorithm (Algorithm 1). For detailed proofs of the theorem statements, see Appendix B.

Theorem 3 (Privacy guarantee)**.** *Algorithm 1 is ϵ-differentially private.*

Proof Sketch. Given the binary tree, the sequence $\hat{w}_2, \cdots, \hat{w}_T$ is completely determined. Hence, it suffices to argue privacy for the collection of noisy sums associated to nodes in the binary tree. At first glance, it seems that each loss function affects only one leaf in the tree, and hence at most $\log T$ of the nodes' partial sums. If it were true, that statement would make the analysis simple. The analysis is delicate, however, since the value (gradient z_τ) at a leaf τ in the tree depends on the partial sums that are released before time τ. Hence, changing one loss function f_t actually affects *all* subsequent partial sums. One can get around this by using the fact that differential privacy composes adaptively [DMNS06]: we can write the computations done on a particular loss function f_t as a sequence of $\log T$ smaller differentially private computations, where the each computation in the sequence depends on the outcome of previous ones. See Appendix B for details. ☐

In terms of regret guarantee, we show that our algorithm enjoys regret of $O(p \log^{2.5} T)$ (assuming other parameters to be constants). Compared to the non-private regret bound of $O(\log T)$, our regret bound has an extra $\log^{1.5} T$ factor and an *explicit* dependence on the dimensionality (p). A formal regret bound for PFTAL algorithm is given in Theorem 4.

Theorem 4 (Regret guarantee)**.** *Let $f_1, \cdots, f_T$ be L-Lipschitz, H-strongly convex functions and let $\mathcal{C} \subseteq \mathbb{R}^p$ be a fixed convex set. For adaptive adversaries, the expected regret satisfies:*

$$\mathbb{E}\left[Regret(T)\right] = O\left(\frac{p(L + H\|\mathcal{C}\|_2)^2 \log^{2.5} T}{\epsilon H}\right).$$

Here expectation is taken over the random coins of the algorithm and adversary.

Results for Lipschitz Convex Costs: Our algorithm for strongly convex costs can be adapted to arbitrary Lipschitz convex costs by executing Algorithm 1 on functions $h_t(w) = f_t(w) + \frac{H}{2}\|w\|_2^2$ instead of the f_t's. Setting $H = O(p \log^{2.5} T/(\epsilon\sqrt{T}))$ will give us a regret bound of $\tilde{O}(\sqrt{pT}/\epsilon)$. See Appendix C for details.

3 Private Online Learning: Bandit Setting

In this section we adapt the Private Follow the Approximate Leader (PFTAL) from Section 2 to the bandit setting. Existing (nonprivate) bandit algorithms for online convex optimization follow a generic reduction to the full-information setting [FKM05, ADX10], called the "one-point" (or "one-shot") gradient trick. Our adaptation of PFTAL to the bandit setting also uses this technique. Specifically, to define the quadratic lower bounds to the input cost functions (as in (3)), we replace the exact gradient of f_t at $\hat{w}_t$ with a one-point approximation.

In this section we describe our results for strongly convex costs. Specifically, to define the quadratic lower bounds to the input cost functions (as in (3)), we replace the exact gradient of f_t at $\hat{w}_t$ with a one-point approximation. As in the full information setting, one may obtain regret bounds for general convex functions in the bandit setting by adding a strongly convex regularizer to the cost functions.

One-point Gradient Estimates [FKM05]: Suppose one has to estimate the gradient of a function $f : \mathbb{R}^p \to \mathbb{R}$ at a point $w \in \mathbb{R}^p$ via a single query access to f. [FKM05] showed that one can approximate $\bigtriangledown f(w)$ by $\frac{p}{\beta} f(w+\beta u)u$, where $\beta > 0$ is a small real parameter and u is a uniformly random vector from the p-dimensional unit sphere $\mathbb{S}^{p-1} = \{a \in \mathbb{R}^p : \|a\|_2 = 1\}$. More precisely, $\bigtriangledown f(w) = \lim_{\beta \to 0} \mathbb{E}_u \left[\frac{p}{\beta} f(w+\beta u)u\right]$.

For finite, nonzero values of β, one can view this technique as estimating the gradient of a smoothed version of f. Given $\beta > 0$, define $\hat{f}(w) = \mathbb{E}_{v \sim \mathbb{B}^p}\left[f(w+\beta v)\right]$ where $\mathbb{B}^p$ is the unit ball in $\mathbb{R}^p$. That is, $\hat{f} = f * U_{\beta \mathbb{B}^p}$ is the convolution of f with the uniform distribution on the ball $\beta\mathbb{B}^p$ of radius β. By Stokes' theorem, we have $\mathbb{E}_{u \sim \mathbb{S}^{p-1}}\left[\frac{p}{\beta} f(w+\beta u)u\right] = \bigtriangledown \hat{f}(w)$.

3.1 Follow the Approximate Leader (Bandit version): Non-private Algorithm

Let $\tilde{W} = \langle \tilde{w}_1, \cdots, \tilde{w}_T \rangle$ be a sequence of vectors in $\mathcal{C}$ (the outputs of the algorithm). Corresponding to the smoothed function $\hat{f}_t = f * U_{\beta\mathbb{B}^p}$, we define a quadratic lower bound $\hat{g}_t$:

$$\hat{g}_t(w) = \hat{f}_t(\tilde{w}_t) + \langle \bigtriangledown \hat{f}_t(\tilde{w}_t), w - \tilde{w}_t \rangle + \tfrac{H}{2}\|w - \tilde{w}_t\|_2^2 \tag{6}$$

Notice that $\hat{g}_t$ is a uniform lower bound on $\hat{f}_t$ satisfying $\hat{g}_t(\tilde{w}_t) = \hat{f}_t(\tilde{w}_t)$ and $\bigtriangledown \hat{g}_t(\tilde{w}_t) = \bigtriangledown \hat{f}_t(\tilde{w}_t)$.

To define $\hat{g}_t$, one needs access to $\bigtriangledown \hat{f}_t(\tilde{w}_t)$. As suggested above, we replace the true gradient with the one-point estimate. Consider the following proxy $\tilde{g}_t$ for $\hat{g}_t$:

$$\tilde{g}_t(w) = \underbrace{\hat{f}_t(\tilde{w}_t) - \langle \bigtriangledown \hat{f}_t(\tilde{w}_t), \tilde{w}_t \rangle}_{A} + \langle \frac{p}{\beta} f_t(\tilde{w}_t + \beta u_t)u_t, w \rangle + \frac{H}{2}\|w - \tilde{w}_t\|_2^2 \tag{7}$$

where u_T is drawn uniformly from the unit sphere $\mathbb{S}^{p-1}$. Note that in (7) we replaced the gradient of $\hat{f}_t$ with its one-point approximation only in one of its two occurrences (the inner product with w).

We would like to define $\tilde{w}_{t+1}$ as the minimizer of the sum of proxies $\sum_{\tau=1}^{t} \tilde{g}_\tau(w)$. One difficulty remains: because f_t is only assumed to be defined on $\mathcal{C}$, the approximation $\frac{p}{\beta} f_t(\tilde{w}_t + \beta u_t)u_t$ is only defined when $\tilde{w}_t$ is sufficiently far inside $\mathcal{C}$. Recall from the introduction that we assume $\mathcal{C}$ contains $r\mathbb{B}^p$ (the ball of radius r). To ensure that we only evaluate f on $\mathcal{C}$, we actually minimize over a smaller set $(1-\xi)\mathcal{C}$, where $\xi = \frac{\beta}{r}$. We obtain:

$$\tilde{w}_{t+1} = \arg\min_{w \in (1-\xi)\mathcal{C}} \sum_{\tau=1}^{t} \tilde{g}_\tau(w) = \arg\min_{w \in (1-\xi)\mathcal{C}} \langle \sum_{\tau=1}^{t} \left(\frac{p}{\beta} f_t(\tilde{w}_t + \beta u_t)u_t\right), w \rangle + \tfrac{H}{2} \sum_{\tau=1}^{t} \|w - \tilde{w}_\tau\|_2^2 \tag{8}$$

(We have use the fact that to minimize $\tilde{g}_t$, one can ignore the constant term A in (7).)

We can now state the bandit version of FTAL. At each step $t = 1, ..., T$:

1. Compute $\tilde{w}_{t+1}$ using (8).
2. Output $\hat{w}_t = \tilde{w}_t + \beta u_t$.

Theorem 12 (in Appendix D) gives the precise regret guarantees for this algorithm. For adaptive adversaries the regret is bounded by $\tilde{O}(p^{2/3}T^{3/4})$ and for oblivious adversaries the regret is bounded by $\tilde{O}(p^{2/3}T^{2/3})$.

3.2 Follow the Approximate Leader (Bandit version): Private Algorithm

To make the bandit version of FTAL ϵ-differentially private, we replace the value $v_t = \sum_{\tau=1}^{t}\left(\frac{p}{\beta} f_t(w_t^\dagger + \beta u_t) u_t\right)$ with a private approximation $v_t^\dagger$ computed using the tree-based sum protocol. Specifically, at each time step t we output

$$w_{t+1}^\dagger = \arg\min_{w\in(1-\xi)\mathcal{C}} \langle v_t^\dagger, w\rangle + \frac{H}{2}\sum_{\tau=1}^{t}\|w - w_\tau^\dagger\|_2^2 \,. \tag{9}$$

See Algorithm 3 (Appendix E.1) for details.

Theorem 5 (Privacy guarantee)**.** *The bandit version of Private Follow The Approximate Leader (Algorithm 3) is ϵ-differentially private.*

The proof of Theorem 5 is exactly the same as of Theorem 3, and hence we omit the details.

In the following theorem we provide the regret guarantee of the Private FTAL (bandit version). For a complete proof, see Appendix E.2.

Theorem 6 (Regret guarantee)**.** *Let $\mathbb{B}^p$ be the p-dimensional unit ball centered at the origin and $\mathcal{C} \subseteq \mathbb{R}^p$ be a convex set such that $r\mathbb{B}^p \subseteq \mathcal{C} \subseteq R\mathbb{B}^p$ (where $0 < r < R$). Let $f_1, \cdots, f_T$ be L-Lipschitz, H-strongly convex functions such that for all $w \in \mathcal{C}$, $|f_i(w)| \leq B$. Setting $\xi = \beta/r$ in the bandit version of Private Follow The Approximate Leader (Algorithm 3 in Appendix E.1), we obtain the following regret guarantees.*

1. **(Oblivious adversary)** *With $\beta = \frac{p}{T^{1/3}}$, $\mathbb{E}\left[Regret(T)\right] \leq \tilde{O}\left(pT^{2/3}\chi\right)$*
2. **(Adaptive adversary)** *With $\beta = \frac{p}{T^{1/4}}$, $\mathbb{E}\left[Regret(T)\right] \leq \tilde{O}\left(pT^{3/4}\chi\right)$*

Here $\chi = \left(BR + (1 + R/r)L + \frac{(H\|\mathcal{C}\|_2 + B)^2}{H}\left(1 + \frac{B}{\epsilon}\right)\right)$. The expectations are taken over the randomness of the algorithm and the adversary.

One can remove the dependence on r in Thm. 6 by rescaling $\mathcal{C}$ to isotropic position. This increases the expected regret bound by a factor of $(LR + \|\mathcal{C}\|_2)$. See [FKM05] for details.

Bound for general convex functions: Our results in this section can be extended to the setting of arbitrary Lipshitz convex costs via regularization, as in Section C (by adding $\frac{H}{2}\|w\|_2^2$ to each cost function f_t) . With the appropriate choice of H the regret scales as $\tilde{O}(T^{3/4}/\epsilon)$ for both oblivious and adaptive adversaries. See Appendix E.3 for details.

4 Open Questions

Our work raises several interesting open questions: First, our regret bounds with general convex functions have the form $\tilde{O}(\sqrt{T}/\epsilon)$. We would like to have a regret bound where the parameter $1/\epsilon$ is factored out with lower order terms in the regret, *i.e.,* we would like to have regret bound of the form $O(\sqrt{T}) + o(\sqrt{T}/\epsilon)$.

Second, our regret bounds for convex bandits are worse than the non-private bounds for linear and multi-arm bandits. For multi-arm bandits [ACBF02] and for linear bandits [AHR08], the non-private regret bound is known to be $O(\sqrt{T})$. If we use our private algorithm in this setting, we will incur a regret of $\tilde{O}(T^{2/3})$. Can we get $O(\sqrt{T})$ regret for multi-arm or linear bandits?

Finally, bandit algorithms require internal randomness to get reasonable regret guarantees. Can we harness the randomness of non-private bandit algorithms in the design private bandit algorithms? Our current privacy analysis ignores this additional source of randomness.

References

[ACBF02] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 2002.

[ADX10] Alekh Agarwal, Ofer Dekel, and Lin Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *COLT*, 2010.

[AHR08] Jacob Abernethy, Elad Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *COLT*, 2008.

[BCB12] Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv:1204.5721*, 2012.

[BDMN05] Avrim Blum, Cynthia Dwork, Frank McSherry, and Kobbi Nissim. Practical privacy: The SuLQ framework. In *PODS*, 2005.

[CM08] Kamalika Chaudhuri and Claire Monteleoni. Privacy-preserving logistic regression. In *NIPS*, 2008.

[CMS11] Kamalika Chaudhuri, Claire Monteleoni, and Anand D. Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12:1069–1109, 2011.

[CSS10] TH Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. In *ICALP*, 2010.

[DJW13] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. Local privacy and statistical minimax rates. In *IEEE Symp. on Foundations of Computer Science (FOCS)*, 2013. http://arxiv.org/abs/1302.3203.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *TCC*, 2006.

[DN10] Cynthia Dwork and Moni Naor. On the difficulties of disclosure prevention in statistical databases or the case for differential privacy. *J. Privacy and Confidentiality*, 2(1), 2010.

[DNPR10] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N Rothblum. Differential privacy under continual observation. In *Proceedings of the 42nd ACM symposium on Theory of computing*, 2010.

[Dwo06] Cynthia Dwork. Differential privacy. In *ICALP*, 2006.

[FKM05] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. In *SODA*, 2005.

[HAK07] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Journal of Machine Learning Research*, 2007.

[Han57] James Hannan. Approximation to bayes risk in repeated play. 1957.

[JKT12] Prateek Jain, Pravesh Kothari, and Abhradeep Thakurta. Differentially private online learning. In *COLT*, 2012.

[JT13] Prateek Jain and Abhradeep Thakurta. Differentially private learning with kernels. In *ICML*, 2013.

[KLN+08] Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? In *FOCS*, 2008.

[KM12] Daniel Kifer and Ashwin Machanavajjhala. A rigorous and customizable framework for privacy. In *PODS*, 2012.

[KS08] Shiva Prasad Kasiviswanathan and Adam Smith. A note on differential privacy: Defining resistance to arbitrary side information. *CoRR*, arXiv:0803.39461 [cs.CR], 2008.

[KST12] Daniel Kifer, Adam Smith, and Abhradeep Thakurta. Private convex empirical risk minimization and high-dimensional regression. In *COLT*, 2012.

[Sha11] Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 2011.

[Smi11] Adam Smith. Privacy-preserving statistical estimators with optimal convergence rates. In *STOC*, 2011.

[Zin03] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, 2003.

A Algorithm for Tree based Aggregation Protocol

Algorithm 2 Private Tree based aggregation protocol

Input: Vectors: $\langle z_1, \cdots, z_T \in \mathbb{R}^p \rangle$ (in an online sequence), μ : L_2-norm bound on z_i's, privacy parameter: ϵ.

1: **Initialization:** Define a binary tree of size $2^{\lceil \log_2 T \rceil + 1} - 1$ with leaves $z_1, \cdots, z_T$.
2: **Online Phase:** At each iteration $t \in [T]$, execute Steps 3 to 18.
3: Accept z_t from the data stream.
4: Let $L = \{z_t \to \cdots \to root\}$ be the path from z_t to the root.
5: **Tree update:** Steps 6 till10.
6: $\Lambda \leftarrow$ First node in L that is a left-child in A. Let $L_\Lambda = \{a_t \to \cdots \to \Lambda\}$.
7: **for all** α in L **do**
8: $\quad \alpha \leftarrow \alpha + z_t$.
9: $\quad$ **If** $\alpha \in L_\Lambda$, **then** $\alpha \leftarrow \alpha + n$, where $n \sim \lambda e^{-\frac{\|n\|_2 \epsilon}{\mu(\lceil \log_2 T \rceil + 1)}}$ and λ is the proportionality constant.
10: **end for**
11: **Output private partial sum:** Steps 12 till 18.
12: Initialize vector $v \in \mathbb{R}^p$ to zero. Let $b \leftarrow \lceil \log_2 T \rceil + 1$-bit binary representation of t.
13: **for all** i in $[\lceil \log_2 T \rceil + 1]$ **do**
14: $\quad$ **if** bit $b_i = 1$ **then**
15: $\quad\quad$ **If** i-th node in L (denoted by $L(i)$) is the left child in A, **then** $v \leftarrow v + L(i)$, **else** $v \leftarrow v +$ left sibling$(L(i))$.
16: $\quad$ **end if**
17: **end for**
18: **return** The noisy partial sum v.

B Privacy and Utility Guarantees of PFTAL Algorithm (Algorithm 1)

B.1 Privacy guarantee for Algorithm 1

Proof of Theorem 3. Notice that given $\hat{v}_2, \cdots, \hat{v}_{t+1}$ (where $\hat{v}_{t+1}$ is the noisy version of $v_{t+1} = \sum_{\tau=1}^{t} \triangledown f_t(\hat{w}_\tau)$), the outputs $\hat{w}_2, \cdots \hat{w}_{t+1}$ are completely determined. Hence, it suffices to argue for the privacy of $\hat{v}_2, \cdots, \hat{v}_T$. Let F and F' be any two sequences of L-Lipschitz, H-strongly convex cost functions differing in exactly one cost function. Let $\hat{V} = \langle v_2, \cdots, v_T \rangle$. For ϵ-differential privacy, we need to argue that for any set $S = \langle s_2, \cdots, s_T \rangle$ of T vectors, the following is true.

$$\frac{\Pr[\hat{V}(F) = S]}{\Pr[\hat{V}(F') = S]} = \prod_{t=2}^{T} \frac{\Pr[\hat{v}_t(F) = s_t | \hat{v}_2 = s_2, \cdots, \hat{v}_{t-1} = s_{t-1}]}{\Pr[\hat{v}_t(F') = s_t | \hat{v}_2 = s_2, \cdots, \hat{v}_{t-1} = s_{t-1}]} \leq e^\epsilon \tag{10}$$

Now in (10), each $\hat{v}_t$ is computed using the tree A (see Algorithm 2) and hence fixing the values of the nodes in the tree A completely determines $V(F)$. Let $A(F) = \langle \alpha_1(F), \cdots \alpha_{(2^{\lceil lgT \rceil + 1} - 1)}(F) \rangle$ be the in-order tree traversal of $A(F)$. To prove (10), it suffices to prove that for all possible assignments $A = \langle \alpha_1, \cdots \rangle$ to the tree, the following holds.

$$\frac{\Pr[A(F) = A]}{\Pr[A(F') = A]} = \prod_{t=1}^{(2^{\lceil \log_2 T \rceil + 1} - 1)} \frac{\Pr[\alpha_t(F) = \alpha_t | \alpha_1(F) = \alpha_1, \cdots, \alpha_{t-1}(F) = \alpha_{t-1}]}{\Pr[\alpha_t(F') = \alpha_t | \alpha_1(F') = \alpha_1, \cdots, \alpha_{t-1}(F') = \alpha_{t-1}]} \leq e^\epsilon \tag{11}$$

In the above ratio, changing one entry in the data set F affects only $(\lceil \log_2 T \rceil + 1)$ terms in the product in (11). By the amount of noise added to each node of the tree, each of the ratio in the product of (11) is bounded by $e^{\left(\frac{\epsilon}{\lceil \log_2 T \rceil + 1}\right)}$. (See Line 9 in Algorithm 2). Here we have used the fact that for any vector $w \in \mathcal{C}$, $\| \triangledown f_t(w) \|_2$ is at most L (by the Lipschitz property) and $\| \triangledown \tilde{f}_t(w) \|_2$ (in (3)) is at most $L + H\|\mathcal{C}\|_2$ (by the bound on the convex set $\mathcal{C}$).

Hence, we can conclude that in overall, Algorithm 1 is ϵ-differentially private. $\square$

B.2 Regret guarantee for Algorithm 1

Proof of Theorem 4. Recall that regret is given by the following expression

$$Regret(T) = \sum_{t=1}^{T} f_t(\hat{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} f_t(w). \tag{12}$$

We will prove the required regret bound via the following three stage argument. We will first show in Lemma 7 that the regret in (12) is upper bounded by the regret for the cost functions $\tilde{f}_t$ (see (3) for notation). Next in Lemma 8, we show that the regret for $\tilde{f}_t$'s with respect to $\hat{w}_t$'s is not "too much" higher compared to the regret with $\tilde{w}_t$'s (see (4) for notation). Finally we bound the regret with respect to $\tilde{w}_t$'s.

Lemma 7. $\sum\limits_{t=1}^{T} f_t(\hat{w}_t) - \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} f_t(w) \leq \sum\limits_{t=1}^{T} \tilde{f}_t(\hat{w}_t) - \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} \tilde{f}_t(w).$

Proof. First notice that by definition, $f_t(\hat{w}_t) = \tilde{f}_t(\hat{w}_t)$. Also, notice that $\tilde{f}_t(w) \leq f_t(w)$ for all $w \in \mathbb{R}^p$. There fore, i) $\sum\limits_{t=1}^{T} f_t(\hat{w}_t) = \sum\limits_{t=1}^{T} \tilde{f}_t(\hat{w}_t)$ and ii) $\min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} \tilde{f}_t(w) \leq \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} f_t(w)$.

This completes the proof. □

In the next lemma we show that the regret with the outputs $\hat{w}_1, \cdots, \hat{w}_T$ is not much different from with respect to $\tilde{w}_1, \cdots, \tilde{w}_T$.

Lemma 8. *Under the randomness of Algorithm 1, the following is true.*

$$\mathbb{E}\left[\sum_{t=1}^{T} \tilde{f}_t(\hat{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w)\right] \leq \mathbb{E}\left[\sum_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w)\right] + \frac{4p(L + H\|\mathcal{C}\|_2)^2 \log^{2.5} T}{\epsilon H}.$$

Proof. Recall that $\tilde{w}_{t+1} = \arg\min\limits_{w \in \mathcal{C}} J(w)$, where $J(w) = \langle \sum\limits_{\tau=1}^{t} \triangledown f_t(\hat{w}_\tau), w \rangle + \frac{H}{2} \sum\limits_{\tau=1}^{t} \|w - \hat{w}_\tau\|_2^2$. We can equivalently write $\hat{w}_{t+1} = \arg\min\limits_{w \in \mathcal{C}} J(w) + \langle n, w \rangle$, where n is the noise added in the noisy computation of $v_{t+1} = \sum\limits_{\tau=1}^{t} \triangledown f_t(\hat{w}_\tau)$ in Line 9 (via the tree-aggregation scheme). By the Ht-strong convexity property of $J(w)$, we have

$$\|\tilde{w}_{t+1} - \hat{w}_{t+1}\|_2 \leq \frac{2\|n\|_2}{Ht}. \tag{13}$$

Now, since f_t is assumed to be L-Lipschitz and the L_2 norm of any vector in $\mathcal{C}$ is bounded by $\|\mathcal{C}\|_2$, it directly follows that $\tilde{f}_t$ is $(L + H\|\mathcal{C}\|_2)$-Lipschitz. Therefore, from (13) and using the Lipschitz property of $\tilde{f}_t$, we have

$$|\tilde{f}_t(\hat{w}_t) - \tilde{f}_t(\tilde{w}_t)| \leq \frac{2\|n\|_2(L + H\|\mathcal{C}\|_2)}{Ht}.$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \tilde{f}_t(\hat{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w)\right] &\leq \sum_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w) + \frac{2\mathbb{E}[\|n\|_2](L + H\|\mathcal{C}\|_2)}{H} \sum_{t=1}^{T} \frac{1}{t} \\
&\leq \sum_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w) + \frac{2\mathbb{E}[\|n\|_2](L + H\|\mathcal{C}\|_2) \log T}{H}.
\end{aligned}$$

(14)

(15)

To bound $\mathbb{E}[\|n\|_2]$, notice that n is formed by adding at most $\lceil \log T \rceil + 1$ vectors whose norms are drawn from the Gamma distribution with scale p and shape $\frac{(\lceil \log T \rceil + 1)(L+H\|\mathcal{C}\|_2)}{\epsilon}$. Therefore,$\mathbb{E}[\|n\|_2] \leq \frac{4p \log^{1.5} T(L+H\|\mathcal{C}\|_2)}{\epsilon}$.

Plugging in the above bound in (15), we complete the proof. □

Next, we prove the following fact which will be useful in proving the regret bound. In the online learning literature this fact is also called the bound on regret via the bound on forward regret [HAK07].

Fact 9. $\sum\limits_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} \tilde{f}_t(w) \leq \sum\limits_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \tilde{f}_t(\tilde{w}_{t+1})$.

Proof. We prove the above fact by proving that $\sum\limits_{t=1}^{T} \tilde{f}_t(\tilde{w}_{t+1}) \leq \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} \tilde{f}_t(w)$. We prove this by induction. Clearly the base case is true by definition of $\tilde{w}_2$ (see (4)). Now assume correctness for $T-1$, and

$$
\begin{aligned}
\sum_{t=1}^{T} \tilde{f}_t(\tilde{w}_{t+1}) &\leq \min_{w \in \mathcal{C}} \sum_{t=1}^{T-1} \tilde{f}_t(w) + \tilde{f}_T(w_{T+1}) \quad \text{(by induction hypothesis)} \\
&\leq \sum_{t=1}^{T-1} \tilde{f}_t(w_{T+1}) + \tilde{f}_T(w_{T+1}) \\
&= \min_{w \in \mathcal{C}} \sum_{t=1}^{T} \tilde{f}_t(w) \quad \text{(by definition).}
\end{aligned}
$$

□

Let $\zeta = \sum\limits_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \min\limits_{w \in \mathcal{C}} \sum\limits_{t=1}^{T} \tilde{f}_t(w)$. Using Fact 9 above and the Lipschitz property of $\tilde{f}_t$'s, we can conclude that $\zeta \leq (L + H\|\mathcal{C}\|_2) \sum\limits_{t=1}^{T} \|\tilde{w}_t - \tilde{w}_{t+1}\|_2$. All we now need to do is bound $\|\tilde{w}_t - \tilde{w}_{t+1}\|_2$ for all t.

Claim 10. *For all t,* $\|\tilde{w}_t - \tilde{w}_{t+1}\|_2 \leq \frac{2(L+H\|\mathcal{C}\|_2)}{Ht}$.

Proof. Notice that

$$
\tilde{w}_t = \arg\min_{w \in \mathcal{C}} \sum_{\tau=1}^{t-1} \tilde{f}_\tau(w)
$$

and

$$
\tilde{w}_{t+1} = \arg\min_{w \in \mathcal{C}} \sum_{\tau=1}^{t-1} \tilde{f}_\tau(w) + \tilde{f}_t(w).
$$

Let $J(w) = \sum\limits_{\tau=1}^{t-1} \tilde{f}_\tau(w) + f_t(w)$. Therefore,

$$
\begin{aligned}
J(\tilde{w}_t) &\geq J(\tilde{w}_{t+1}) + \frac{Ht}{2}\|\tilde{w}_t - \tilde{w}_{t+1}\|_2^2 \\
\Leftrightarrow \frac{Ht}{2}\|\tilde{w}_t - \tilde{w}_{t+1}\|_2^2 &\leq \left(\sum_{\tau=1}^{t-1} \tilde{f}_\tau(\tilde{w}_t) - \sum_{\tau=1}^{t-1} \tilde{f}_\tau(\tilde{w}_{t+1})\right) + \tilde{f}_t(\tilde{w}_t) - \tilde{f}_t(\tilde{w}_{t+1}) \\
\Leftrightarrow \frac{Ht}{2}\|\tilde{w}_t - \tilde{w}_{t+1}\|_2^2 &\leq f_t(\tilde{w}_t) - f_t(\tilde{w}_{t+1}) \leq (L + H\|\mathcal{C}\|_2)\|\tilde{w}_t - \tilde{w}_{t+1}\|_2 \\
\Leftrightarrow \|\tilde{w}_t - \tilde{w}_{t+1}\|_2 &\leq \frac{2(L + H\|\mathcal{C}\|_2)}{Ht}.
\end{aligned}
$$

□

Using the above claim and Fact 9, we can conclude that

$$\sum_{t=1}^{T} \tilde{f}_t(\tilde{w}_t) - \tilde{f}_t(\tilde{w}_{t+1}) \leq \frac{2(L + H\|\mathcal{C}\|_2)^2 \log T}{H}.$$

Combining the above expression with Lemma 8, we obtain the required regret bound. □

C Results for General Convex Costs

In this section we will adapt the Private Follow the Approximate Leader (Algorithm 1) for H-strongly convex costs from previous section to the to the case of general convex functions. The idea is to add a L_2-regularizer to the cost functions while running the PFTAL algorithm, and then tune H for the optimal regularization parameter. To be more precise, for every cost function f_t, we will have Algorithm 1 work with the cost function $h_t(w) = f_t(w) + \frac{H}{2}\|w\|_2^2$ (instead of f_t). Clearly, each h_t is now H-strongly convex. So, the privacy and regret guarantees in Section 2.1.1 will hold for the cost sequence $h_1, \cdots, h_T$. Notice that the following is always true for any sequence of vectors $w_1, \cdots, w_T \in \mathcal{C}$, since the diameter of the convex set $\mathcal{C}$ is bounded.

$$\sum_{t=1}^{T} f_t(w_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} f_t(w) \leq \left(\sum_{t=1}^{T} h_t(w_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} h_t(w)\right) + \frac{HT}{2}\|\mathcal{C}\|_2^2. \tag{16}$$

If $\hat{w}_1, \cdots, \hat{w}_T$ be the sequence of outputs of Algorithm 1 on the cost sequence $h_1, \cdots, h_T$, then (17) follows from Theorem 4 and (16).

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T} f_t(\hat{w}_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} f_t(w)\right] &\leq \mathbb{E}\left[\sum_{t=1}^{T} h_t(\hat{w}_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} h_t(w)\right] + \frac{HT}{2}\|\mathcal{C}\|_2^2 \\ &= O\left(\frac{p(L + H\|\mathcal{C}\|_2)^2 \log^{2.5} T}{\epsilon H}\right) + \frac{HT}{2}\|\mathcal{C}\|_2^2. \end{aligned} \tag{17}$$

Theorem 11 (Regret guarantee)**.** *Let $f_1, \cdots, f_T$ be L-Lipschitz convex functions and let $\mathcal{C} \subseteq \mathbb{R}^p$ be a fixed convex set. Setting the parameter H in the regularizer $\frac{H}{2}\|w\|_2^2$ optimally, we have the following regret bound.*

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\hat{w}_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} f_t(w)\right] = O\left(\frac{\sqrt{p\log^{2.5} T}\left(L + \sqrt{\frac{p\log^{2.5} T}{\epsilon T}}\|\mathcal{C}\|_2\right)^2}{\epsilon}\sqrt{T}\right). \tag{18}$$

The expectation is over the randomness of the algorithm and the adversary.

Proof. Setting $H = \sqrt{\frac{p\log^{2.5} T}{\epsilon T}}$ in the right hand side of (17), we get the regret guarantee in (18) for the sequence of outputs $\hat{w}_1, \cdots, \hat{w}_T$. □

Notice that the regret bound in (18) is a factor or $\sqrt{p\log^{2.5} T}/\epsilon$ worse than the non-private regret bound of $O(\sqrt{T})$, assuming other parameters to be constants and $T = \omega\left(\frac{p}{\epsilon}\right)$. The assumption on T is benign, since if $T = O\left(\frac{p}{\epsilon}\right)$, then the regret guarantee in (18) will no longer be sublinear.

We believe it is unlikely that one can remove the explicit dependence on the dimensionality in the regret bound for general convex costs, while preserving differential privacy.

D Regret Guarantees for Follow The Approximate Leader (Bandit version)

Theorem 12 (Regret guarantee)**.** *Let $\mathbb{B}^p$ be a d-dimensional unit ball centered at the origin and $\mathcal{C}$ be a convex set such that $r\mathbb{B}^p \subseteq \mathcal{C} \subseteq R\mathbb{B}^p$ (where $0 < r < R$).*

- ***Adaptive adversary:*** *Setting $\beta = \frac{p^{2/3}}{T^{1/4}}$ and $\xi = \beta/r$, the expected regret is at most*

$$\tilde{O}\left(p^{2/3}T^{3/4}\left(BR + \left(1 + \frac{R}{r}\right)L + \frac{(H\|\mathcal{C}\|_2 + B)^2}{H}\right)\right).$$

- ***Oblivious adversary:*** *Setting $\beta = \frac{p^{2/3}}{T^{1/3}}$ and $\xi = \beta/r$, the expected regret is at most*

$$\tilde{O}\left(p^{2/3}T^{2/3}\left((1 + R/r)\,L + \frac{(H\|\mathcal{C}\|_2 + B)^2}{H}\right)\right).$$

The expectation is over the randomness of the algorithm and the adversary.

D.1 Proof: Regret guarantee for Adaptive Adversary

Proof. We prove the regret bound in the following three stages: i) In Lemma 13, we show that the regret for the output sequence $\hat{w}_1, \cdots, \hat{w}_t$ with respect to the original cost functions f_t's is not much higher compared to $\hat{f}_t$'s with parameter vectors $\tilde{w}_1, \cdots, \tilde{w}_T$ (defined in (8)), ii) We show in Lemma 14 that the regret of $\hat{f}_t$'s with the parameter vectors $\tilde{w}_t$'s is at most the regret of the cost functions $\hat{g}_t$'s with the same parameter vectors (defined in (6)). iii) In Lemma 15, we directly bound the regret on $\hat{g}_t$'s with parameter vectors $\tilde{w}_t$'s.

Lemma 13. *For any sequence of parameter vectors $\tilde{w}_1, \cdots, \tilde{w}_T$ from the convex set $(1-\xi)\mathcal{C}$ and vectors $\hat{w}_1, \cdots, \hat{w}_T$ such that for all $t \in [T]$, $\hat{w}_t = \tilde{w}_t + \beta u_t$ (where u_t is a uniform vector drawn from the unit sphere $\mathbb{S}^{p-1}$), the following is true.*

$$\sum_{t=1}^{T} f_t(\hat{w}_t) - \min_{w \in \mathcal{C}} \sum_{t=1}^{T} f_t(w) \leq \sum_{t=1}^{T} \hat{f}_t(\tilde{w}_t) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{f}_t(w) + 3\beta LT + \xi RLT$$

Proof. First notice that for any $w \in \mathcal{C}$, the following is true for any $t \in [T]$ by the Lipschitz property of f_t's.

$$\begin{aligned}\left|f_t(w) - \hat{f}_t(w)\right| &= \left|f_t(w) - \mathbb{E}_{v\sim\mathbb{B}^p}[f_t(w+\beta v)]\right| \\ &= \left|\mathbb{E}_v\left[f_t(w) - f_t(w+\beta v)\right]\right| \\ &\leq L\beta \cdot \mathbb{E}_v\left[\|v\|_2\right] \leq \beta L \end{aligned} \tag{19}$$

Now for any $w \in \mathcal{C}$, by the Lipschitz property of f_t, we can obtain the following bound $|f_t(w) - f_t((1-\xi)w)| \leq \xi LR$. This means that $\min\limits_{w\in(1-\xi)\mathcal{C}} \sum\limits_{t=1}^{T} f_t(w) \leq \min\limits_{w\in\mathcal{C}} \sum\limits_{t=1}^{T} f_t(w) + \xi LRT$. Therefore, by (19) we directly have

$$\min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{f}_t(w) \leq \min_{w\in\mathcal{C}} \sum_{t=1}^{T} f_t(w) + \beta LT + \xi RLT \tag{20}$$

By Lipschitz property of f_t, we have $|f_t(\hat{w}_t) - f_t(\tilde{w}_t)| \leq \beta L$. Additionally, by (22) we have $|\hat{f}_t(\tilde{w}_t) - f_t(\tilde{w}_t)| \leq \beta L$. Combining these two observations, we get

$$\sum_{t=1}^{T} f_t(\hat{w}_t) \leq \sum_{t=1}^{T} \hat{f}_t(\tilde{w}_t) + 2\beta LT \tag{21}$$

Combining (20) and (21) we get the required error guarantee. □

Lemma 14. *For any sequence of parameter vectors $\tilde{w}_1, \cdots, \tilde{w}_T$ from the convex set $(1-\xi)\mathcal{C}$, the following is true.*

$$\sum_{t=1}^{T} \hat{f}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{f}_t(w) \leq \sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)$$

Proof. First notice that by definition, $\hat{f}_t(\tilde{w}_t) = \hat{g}_t(\tilde{w}_t)$. Also, notice that $\hat{g}_t(w) \leq \hat{f}_t(w)$ for all $w \in \mathbb{R}^p$. There fore, i) $\sum_{t=1}^{T} \hat{f}_t(\tilde{w}_t) = \sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t)$ and ii) $\min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{f}_t(w) \leq \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)$. This completes the proof. □

Using the above lemma we directly get (22) below. In order to obtain the final regret guarantee, we just need to bound $\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)$ and appropriately set β and ξ.

$$\sum_{t=1}^{T} f_t(\hat{w}_t) - \min_{w\in\mathcal{C}} \sum_{t=1}^{T} f_t(w) \leq \sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) + 3\beta LT + \xi RLT \tag{22}$$

Lemma 15. $\mathbb{E}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)\right] \leq \frac{2(H\|\mathcal{C}\|_2+pB/\beta)^2}{H} \log T + 2BR\sqrt{T}\frac{p}{\beta}$. *The expectation is over the random unit vectors $u_1, \cdots, u_T$.*

Proof. Since $\tilde{g}_t$'s are H-strongly convex functions and $\left(H\|\mathcal{C}\|_2 + \frac{p}{\beta}B\right)$-Lipschitz, from the regret analysis in Lemma 8 we directly have the following.

$$\sum_{t=1}^{T} \tilde{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \tilde{g}_t(w) \leq \frac{2(H\|\mathcal{C}\|_2 + \frac{p}{\beta}B)^2}{H} \log T \tag{23}$$

Let $w^* = \arg\min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)$. Therefore by (23), we have (24).

$$\sum_{t=1}^{T} \tilde{g}_t(\tilde{w}_t) - \sum_{t=1}^{T} \tilde{g}_t(w^*) \leq \frac{2(H\|\mathcal{C}\|_2 + \frac{p}{\beta}B)^2}{H} \log T \tag{24}$$

Notice that

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)\right] &= \mathbb{E}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \sum_{t=1}^{T} \hat{g}_t(w^*)\right] \\ &= \mathbb{E}[\sum_{t=1}^{T} \tilde{g}_t(\tilde{w}_t)] - \mathbb{E}[\sum_{t=1}^{T} \hat{g}_t(w^*)] \end{aligned} \tag{25}$$

The last inequality follows from the observation that $\mathbb{E}_{u_t}[\tilde{g}_t(w)] = \hat{g}_t(w)$ for all $w \in \mathcal{C}$. Let $\alpha_t = \bigtriangledown\mathbb{E}_{v\sim\mathbb{B}^p}\left[f_t(\tilde{w}_t + \beta v)\right] - \frac{p}{\beta} f_t(\tilde{w}_t + \beta u_t)u_t$. For any $w \in (1-\xi)\mathcal{C}$,

$$\left|\sum_{t=1}^{T} (\hat{g}_t(w) - \tilde{g}_t(w))\right| = \left|\langle w, \sum_{t=1}^{T} \alpha_t\rangle\right| \leq R\|\sum_{t=1}^{T} \alpha_t\|_2$$

Now,

$$\begin{aligned}\mathbb{E}\left[\|\sum_{t=1}^{T} \alpha_t\|_2\right]^2 &\leq \mathbb{E}\left[\|\sum_{t=1}^{T} \alpha_t\|_2^2\right] \\ &= \sum_{t=1}^{T} \mathbb{E}[\|\alpha_t\|_2^2] + 2\sum_{t<t'} \mathbb{E}\left[\alpha_t \alpha_{t'}\right] \leq 4T\frac{p^2}{\beta^2}B^2\end{aligned}$$

The last inequality is true because $\mathbb{E}\left[\alpha_t\alpha_{t'}\right] = 0$. Therefore,

$$E\left[\min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)\right] \geq \mathbb{E}\left[\min_{w\in(1-\xi)\mathcal{C}} \sum_{t=1}^{T} \tilde{g}_t(w)\right] - \frac{2p}{\beta}BR\sqrt{T}$$

Using this bound in (25), we have

$$\mathbb{E}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right] \le \mathbb{E}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\right] + \frac{2p}{\beta}BR\sqrt{T}$$

Plugging in the bound from (24) completes the proof. □

Combining Lemmas 13, 14 and 15, we obtain the following.

$$\mathbb{E}\left[\sum_{t=1}^{T}f_t(\hat{w}_t) - \min_{w\in\mathcal{C}}\sum_{t=1}^{T}f_t(w)\right] \le 3\beta LT + \xi RLT + \frac{2(H\|\mathcal{C}\|_2 + \frac{p}{\beta}B)^2}{H}\log T + \frac{2p}{\beta}BR\sqrt{T}$$

Setting, $\beta = \frac{p^{2/3}}{T^{1/4}}$ and $\xi = \frac{\beta}{r}$ gives the required regret bound. □

D.2 Proof: Regret guarantee for Oblivious Adversary

Proof. The proof of this theorem is similar to the proof with adaptive adversary, except we will be prove a tighter bound corresponding to Lemma 15.

Lemma 16. $\mathbb{E}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right] \le \frac{2(H\|\mathcal{C}\|_2+pB/\beta)^2}{H}\log T$. *The expectation is over the random unit vectors* $u_1,\cdots,u_T$.

Proof. Similar to the proof of Lemma 15, let $w^* = \arg\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)$. Notice that $\mathbb{E}_{u_t}[\tilde{g}_t(w)] = \hat{g}_t(w)$ for all $w\in\mathcal{C}$. Therefore,

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t) - \sum_{t=1}^{T}\hat{g}_t(w^*)\right] &= \mathbb{E}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)\right] - \mathbb{E}\left[\sum_{t=1}^{T}\tilde{g}_t(w^*)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t) - \sum_{t=1}^{T}\tilde{g}_t(w^*)\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t) - \min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\right] \qquad (26)
\end{aligned}$$

Now, using the bound from (23) in (26), we get the required regret bound. □

Combining Lemmas 13, 14 and 16, we obtain the following.

$$\mathbb{E}\left[\sum_{t=1}^{T}f_t(\hat{w}_t) - \min_{w\in\mathcal{C}}\sum_{t=1}^{T}f_t(w)\right] \le 3\beta LT + \xi RLT + \frac{2(H\|\mathcal{C}\|_2 + \frac{p}{\beta}B)^2}{H}\log T$$

Setting, $\beta = \frac{p^{2/3}}{T^{1/3}}$ and $\xi = \frac{\beta}{r}$ gives the required regret bound. □

E Algorithm and Regret Guarantees for Private Follow The Approximate Leader (Bandit version)

E.1 Private Follow The Approximate Leader (Bandit version) Algorithm

Algorithm 3 Differentially Private Follow the Approximate Leader (PFTAL): Bandit Version

Input: Cost functions: $\langle f_1, \cdots, f_T \rangle$ (in an online sequence), strong convexity parameter: H, bound on the costs: B, convex set: $\mathcal{C} \subseteq \mathbb{R}^p$, scaling parameter: ξ, sampling radius: β, and privacy parameter: ϵ.

1: $w_1^\dagger \leftarrow$ Any vector from $\mathcal{C}$. **Output** $w_1^\dagger$.
2: Sample u_1 uniformly from the sphere $\mathbb{S}^{p-1} = \{w \in \mathbb{R}^p : \|w\|_2 = 1\}$.
3: Pass $\frac{p}{\beta} f_1(w_1^\dagger + \beta u_1)u_1$, L_2-bound $\frac{pB}{\beta}$ and privacy parameter ϵ to the *tree based protocol* (Algorithm 2) and receive the current partial sum in ${v^\dagger}_1$.
4: **for** time steps $t \in \{1, \cdots, T-1\}$ **do**
5: $\quad w_{t+1}^\dagger = \arg\min\limits_{w \in (1-\xi)\mathcal{C}} \langle {v^\dagger}_t, w \rangle + \frac{H}{2} \sum\limits_{\tau=1}^{t} \|w - w_\tau^\dagger\|_2^2$. **Output** $\hat{w}_t$.
6: $\quad$ Sample u_{t+1} uniformly from the sphere $\mathbb{S}^{p-1}$.
7: $\quad$ Pass $\frac{p}{\beta} f_{t+1}(w_{t+1}^\dagger + \beta u_{t+1})u_{t+1}$, L_2-bound $\frac{pB}{\beta}$ and privacy parameter ϵ to *the tree based protocol* (Algorithm 2) and receive the current partial sum in ${v^\dagger}_{t+1}$.
8: **end for**

E.2 Regret Analysis

Proof of Theorem 6. Corresponding to definitions of $\hat{g}_t$ and $\tilde{g}_t$'s in (6), (7), and (8) (in Section 3.1), we redefine them while using the Taylor expansion around $w_{t+1}^\dagger$.

$$\hat{g}_t(w) = \hat{f}_t(w_t^\dagger) + \langle \triangledown \hat{f}_t(w_t^\dagger), w - w_t^\dagger \rangle + \frac{H}{2}\|w - w_t^\dagger\|_2^2 \tag{27}$$

$$\tilde{g}_t(w) = \hat{f}_t(w_t^\dagger) - \langle \triangledown \hat{f}_t(w_t^\dagger), w_t^\dagger \rangle + \langle \frac{p}{\beta} f_t(w_t^\dagger + \beta u_t)u_t, w \rangle + \frac{H}{2}\|w - w_t^\dagger\|_2^2 \tag{28}$$

$$\tilde{w}_{t+1} = \arg\min_{w \in (1-\xi)\mathcal{C}} \sum_{\tau=1}^{t} \tilde{g}_\tau(w) \tag{29}$$

With the above equations in hand, we can rewrite the definition of $w_{t+1}^\dagger$ in (9) as follows. Here $n_t = v_t^\dagger - v_t$, where ${v^\dagger}_t$ and v_t are as defined in Section 3.2.

$$w_{t+1}^\dagger = \arg\min_{w \in (1-\xi)\mathcal{C}} \sum_{\tau=1}^{t} \tilde{g}_\tau(w) + \langle n_t, w \rangle \tag{30}$$

Using a similar argument we used in Lemma 8, we get the following.

$$\sum_{t=1}^{T} \hat{g}_t(w_t^\dagger) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) \leq \sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) + \frac{2(pB/\beta + H\|\mathcal{C}\|_2)}{H} \sum_{t=1}^{T} \frac{\|n_t\|_2}{t} \tag{31}$$

From (31) and using an expectation bound on $\|n_t\|_2$ similar to Lemma 8, we obtain the following.

$$\begin{aligned} &\mathbb{E}_{n_1,\cdots,n_T}\left[\sum_{t=1}^{T} \hat{g}_t(w_t^\dagger) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) \middle| u_1, \cdots, u_T\right] \\ &\leq \mathbb{E}_{n_1,\cdots,n_T}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) \middle| u_1, \cdots, u_T\right] + \frac{2p(pB/\beta + H\|\mathcal{C}\|_2)^2 \log^{2.5} T}{\beta\epsilon H} \end{aligned} \tag{32}$$

Now,

$$\begin{aligned} &\mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w)\right] \\ &= \mathbb{E}_{n_1,\cdots,n_T}\left[\mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T} \hat{g}_t(\tilde{w}_t) - \min_{w \in (1-\xi)\mathcal{C}} \sum_{t=1}^{T} \hat{g}_t(w) \middle| n_1, \cdots, n_T\right]\right] \end{aligned} \tag{33}$$

If the adversary is adaptive, then by the same line of argument in Lemma 15, we have

$$\begin{aligned}
&\mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\middle| n_1,\cdots,n_T\right]\\
&\le \mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\middle| n_1,\cdots,n_T\right]+\frac{2p}{\beta}BR\sqrt{T}
\end{aligned}\tag{34}$$

If the adversary is oblivious, then by the same line of argument in Lemma 16, we have

$$\begin{aligned}
&\mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\middle| n_1,\cdots,n_T\right]\\
&\le \mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\middle| n_1,\cdots,n_T\right]
\end{aligned}\tag{35}$$

For the purpose of brevity, we combine (34) and (35) into one expression (36), where the term γ equals $\frac{2d}{\beta}\sqrt{TRB}$ for adaptive adversary and zero for oblivious adversary. For the rest of the proof, we will set γ according to the assumption about the adversary.

$$\begin{aligned}
&\mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\middle| n_1,\cdots,n_T\right]\\
&\le \mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\middle| n_1,\cdots,n_T\right]+\gamma
\end{aligned}\tag{36}$$

Plugging (36) back in (33), we get

$$\begin{aligned}
&\mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right]\\
&\le \mathbb{E}_{n_1,\cdots,n_T}\left[\mathbb{E}_{u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\middle| n_1,\cdots,n_T\right]\right]+\gamma\\
&= \mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\right]+\gamma
\end{aligned}\tag{37}$$

Combining (32) and (37), we have

$$\begin{aligned}
&\mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(w_t^{\dagger})-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right]\\
&= \mathbb{E}_{u_1,\cdots,u_T}\left[\mathbb{E}_{n_1,\cdots,n_T}\left[\sum_{t=1}^{T}\hat{g}_t(w_t^{\dagger})-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\middle| u_1,\cdots,u_T\right]\right]\\
&\le \mathbb{E}_{u_1,\cdots,u_T}\left[\mathbb{E}_{n_1,\cdots,n_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\middle| u_1,\cdots,u_T\right]\right]+\frac{2p(pB/\beta+H\|\mathcal{C}\|_2)^2\log^{2.5}T}{\beta\epsilon H}\\
&= \mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right]+\frac{2p(pB/\beta+H\|\mathcal{C}\|_2)^2\log^{2.5}T}{\beta\epsilon H}\\
&\le \mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\tilde{g}_t(\tilde{w}_t)-\min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\tilde{g}_t(w)\right]+\frac{2p(pB/\beta+H\|\mathcal{C}\|_2)^2\log^{2.5}T}{\beta\epsilon H}+\gamma
\end{aligned}\tag{38}$$

Plugging in the absolute bound on $\sum\limits_{t=1}^{T}\tilde{g}_t(\tilde{w}_t) - \min\limits_{w\in(1-\xi)\mathcal{C}}\sum\limits_{t=1}^{T}\tilde{g}_t(w)$ from (23), we obtain the following.

$$\mathbb{E}_{n_1,\cdots,n_T,u_1,\cdots,u_T}\left[\sum_{t=1}^{T}\hat{g}_t(w_t^{\dagger}) - \min_{w\in(1-\xi)\mathcal{C}}\sum_{t=1}^{T}\hat{g}_t(w)\right]$$
$$\leq \frac{2(H\|\mathcal{C}\|_2+\frac{p}{\beta}B)^2}{H}\log T + \frac{2p(pB/\beta+H\|\mathcal{C}\|_2)^2\log^{2.5}T}{\beta\epsilon H}+\gamma \tag{39}$$

Combining Lemmas 13, 14 and (39), we obtain the following. The expectation is over the complete randomness of the private FTAL (bandit version).

$$\mathbb{E}\left[\sum_{t=1}^{T}f_t(\hat{w}_t) - \min_{w\in\mathcal{C}}\sum_{t=1}^{T}f_t(w)\right]$$
$$\leq 3\beta LT + \xi RLT + \frac{2(H\|\mathcal{C}\|_2+\frac{p}{\beta}B)^2}{H}\log T + \frac{2p(pB/\beta+H\|\mathcal{C}\|_2)^2\log^{2.5}T}{\beta\epsilon H}+\gamma$$

Recall that if the adversary is adaptive, then $\gamma=\frac{2p}{\beta}BR\sqrt{T}$ and zero otherwise. Setting $\beta=\frac{p}{T^{1/4}}$ for adaptive adversary and $\beta=\frac{p}{T^{1/3}}$ for oblivious adversary, and setting $\xi=\frac{\beta}{r}$, we get the required regret bound. □

E.3 Private Bandit Learning for General Convex Functions

Our results in this section can be extended to the setting with general convex costs via the regularization "trick" from Appendix C (by adding $\frac{H}{2}\|w\|_2^2$ to each cost function f_t) . One can show that under optimal choice of H, both for oblivious and adaptive adversary, the regret scales as $\tilde{O}(T^{3/4}/\epsilon)$, which is also the best known nonprivate bound [FKM05]. We provide the formal regret guarantee below.

Theorem 17 (Regret guarantee)**.** *Let $\mathbb{B}^p$ be a p-dimensional unit ball centered at the origin and $\mathcal{C}\subseteq\mathbb{R}^p$ be a convex set such that $r\mathbb{B}^p\subseteq\mathcal{C}\subseteq R\mathbb{B}^p$ (where $0<r<R$). Let $f_1,\cdots,f_T$ be L-Lipschitz functions and for all $w\in\mathcal{C}$, $|f_i(w)|\leq B$. Additionally assume that the regularizing parameter H is set to $1/T^{1/4}$. Setting $\beta=\frac{p}{T^{1/4}}$ and $\xi=\beta/r$ in the Private Follow The Approximate Leader (bandit version) algorithm (Algorithm 3), we obtain the following regret guarantee.*

$$\mathbb{E}\left[\sum_{t=1}^{T}f_t(\hat{w}_t) - \min_{w\in\mathcal{C}}\sum_{t=1}^{T}f_t(w)\right]\leq\tilde{O}\left(pT^{3/4}\chi\right).$$

Here $\chi=\left(BR+(1+R/r)L+\frac{B^3}{\epsilon}\right)$. The expectation is over the randomness of the algorithm and the adversary.