Unifying (Federated) (Private) High-Dimensional Bandits via ADMM

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study all possible variants of the high dimensional stochastic linear contextual bandit problem in federated and private settings. We propose a unifying algorithm design and analysis framework built on ADMM. Our method achieves existing state-of-the art guarantees in either setting for the central model. For the federated model, our results are entirely new and near-optimal in either setting. We also establish a novel lower bound on privacy-utility tradeoff for the federated model in the private setting and demonstrate on suitable numerical experiments for all problem variants.

1 Introduction

We study a high-dimensional stochastic contextual linear bandit problem in central and federated learning settings under privacy constraints (Shukla, 2024; Chakraborty et al., 2024). At every epoch, the decision maker is given a set of stochastically generated exogeneous contexts and chooses from a finite set of actions $[K]$ to obtain the highest reward. Since the decision maker does not know the true underlying reward, it creates an estimator to select a decision. This induces an exploration-exploitation trade-off wherein the decision-maker faces the dilemma of exploring arms not played before and exploiting the information accumulated so far. The key underlying assumption in this framework is that the unknown parameter vector is s^*-sparse. This paper is motivated by the need for a unified algorithm and analysis framework for establishing utility and privacy guarantees for this problem under different learning models. From the perspective of this paper, existing work in high-dimensional bandits can be seen in central or federated settings with or without privacy constraints. We motivate the need for such a framework in the in Example 1.1 this is a very common learning setting.

Example 1.1. *A crucial step in assuring the drug safety of oligonucleotide drugs requires learning the relevant thermodynamics from large-scale data distributed across different organizations (Tavara et al., 2021). Preserving data privacy in this distributed setup requires limited and private communication between local nodes. Further, learning the safety curve online requires learning in an online fashion. Techniques proposed in this paper can be used to solve this problem of online, private, and federated learning with high-dimensional but sparse data.*

In the central setting with privacy constraints, the regret analysis follows the by-now standard framework: (i) constructing an estimator for the unknown parameter using LASSO or thresholding-based techniques, (ii) analysing per-step regret using underlying problem structure (such as margin condition, compatibilty or other similar conditions). In this paper, we eschew this line of arguments and propose a generic algorithmic design and analysis framework for this class of problems. In either settings, we view the statistical problem of constructing the LASSO-based estimator through an optimizer's lens. By exploiting the architecture of ADMM-based optimizers (Boyd et al., 2011), our proposed method can achieve state-of-the-art guarantees under all possible combinations of the

Submitted to the 18th European Workshop on Reinforcement Learning (EWRL 2025). Do not distribute.

Table 1: Contributions to existing work. New results are in blue. $\rho = \epsilon^{-2}\log(1/\delta)$ for (ϵ, δ)-DP.

		Central (Theorems 2, 3)	Federated (Theorems 2, 4)
Non-Private	Lower bound	$\Omega(\sqrt{s^*T\log(d/s^*)})$	$\Omega(\sqrt{sMT\log(d/s^*)})$
	Upper bound (existing)	$\mathcal{O}(\sqrt{s^*T\log(d/s^*)})$	–
	Upper bound (ours)	$\mathcal{O}(s^*\sqrt{T\log(d/s^*)})$	$\mathcal{O}(s^*\sqrt{MT\log(d/s^*)})$
Private ((ϵ, δ)-DP)	Lower bound	$\Omega(\log(d/s^*)\sqrt{s^*\rho})$	$\Omega(\log(d/s^*)\sqrt{s^*\rho})$
	Upper bound (existing)	$\mathcal{O}(\sqrt{\rho}(s^*)^{1.5}\log^{1.5}(d/s^*))$	–
	Upper bound (ours)	$\mathcal{O}(\sqrt{\rho}(s^*)^{1.5}\log^2(d/s^*))$	$\mathcal{O}(\sqrt{\rho}(s^*)1.5\log^2(d/s^*))$

learning models and privacy settings (see Table 1). Detailed literature is reviewed in Appendix A. In high-dimensional bandits in the central model, closest to our work are (Chakraborty et al., 2024; Shukla, 2024) that consider the high-dimensional bandit problem in the central model with privacy and propose different thresholding-based methods to solve it. For the federated setup the non-private setting was considered by (Wang et al., 2023). To the best of our knowledge, there is no known work on private high-dimensional federated contextual bandits. This motivates the question:

Is it possible to propose a generic algorithm design and analysis template for high-dimensional bandits that can operate both in central and federated settings, and with and without differential privacy constraints?

Our Contributions: We address these questions affirmatively and contribute to all the aforementioned strands of literature. Our contributions are summarised in Table 1 and detailed as follows:

1. We propose a general framework for designing algorithms for peeled LASSO (Section 3) and high-dimensional bandit problems (Section 4) applicable to several variants of the problem that have been considered independently until now. These include centralized/federated and private/non-private versions thereof. Our framework works for all combinations of these variants by tuning a few hyper-parameters.
2. In Algorithms 1, we propose an admm-based algorithm for the PeeledLASSO applicable for either communication model and establish its estimation error in Theorem 1. These are the first known recovery guarantees for the online private LASSO in private and federated setting. Algorithm 3 extends this to an algorithm for high-dimensional bandits in either setting using forgetting and the doubling trick. Privacy guarantees for these algorithms are established in Theorem 2 and utility guarantees in Theorem 3 and 4. Our utility proofs non-trivially combine iteration-based analysis of ADMM and peeling-based privacy arguments to accommodate bandit-feedback under both communication models.
3. In Theorem 5, we derive a problem independent regret lower bound for the federated model in both private and non-private setting. Our lower bound demonstrates a phase transition between the hardness of the problem depending on the privacy level, intrinsic dimension and sparsity. In the low-dimensional case, a similar phenomenon was shown by Azize and Basu (2022).

2 Model and Preliminaries

Central Model: We consider a linear contextual bandit problem with K arms, where the (unknown) underlying vector $\theta \in \mathbb{R}^d$ that parameterizes rewards is high-dimensional and sparse – often called the LASSO bandit setting (Bastani and Bayati, 2020). Central and federated models correspond to different modes of communication between central server and agents. We consider a time horizon T, where for each time $t \in [T]$, the algorithm is given an (exogenous) context vector $\mathcal{X}_t = \{X_{t,k}\}_{k\in[K]}$, where each $X_{t,k} \in \mathbb{R}^d$. The elements of the set $\mathcal{X}_t$ are drawn i.i.d. from an unknown distribution $\mathcal{D}$. After observing $\mathcal{X}_t$, the algorithm selects an arm $k_t \in [K]$ and observes a random reward r_t given by $r_t = \langle X_{t,k_t}, \theta\rangle + \eta_t$, where η_t is zero-mean bounded noise with variance σ^2. Let $\mathcal{H}_t$ denote the tupple of random variables generated by the past contexts, arm pulls, and observed rewards and $\mathcal{F}_t$ denote the corresponding filtration. We assume that the parameter vector $\theta \in \mathbb{R}^d$ is s-sparse, i.e., its support $S = \{i \in [d] : \theta_i \neq 0\}$ has cardinality s, which is known to the decision maker. Given the history of observations $\mathcal{H}_t = \{X_{k_s,s}, r_s\}_{s=1}^t$, let $\mathcal{F}_t$ denote the natural filtration associated with $\mathcal{H}_t$ and let Π denote the set of all $\mathcal{F}_t$-measurable policies. In the central setting,

our goal is to design a policy $\pi \in \Pi$ that minimizes the cumulative expected regret, defined as: $R^{\pi}(T) = \sum_{t=1}^{T} \mathbb{E}[\max_{k\in[K]} \langle X_{t,k} - X_{t,k_t}, \theta\rangle]$.

Federated Model: In the federated setup, $[M]$ is the set of clients, $[K]$ is the set of arms, $\mathcal{X}$ is the set of contexts. Each client is a K-armed bandit problem with a common parameter being shared across clients. At each time slot t, each client i observes a exogeneous contexts $X_{i,t} \in \mathcal{X}$, pulls arm $k_{i,t} \in [K]$ and receives a reward $r_{i,t} = X_{i,t}^{\top}\theta + \eta_{i,t}$. We assume there exists a central server in the system, and similar to FL, the clients can communicate with the server periodically with zero latency. Specifically, the clients can send "local model updates" to the central server, which then aggregates and broadcasts the updated "global model" to the clients. We also assume that clients and server are fully synchronized. The class of policies considered here are restricted to collinearly correlated policies Huang et al. (2021) (see Definition 6 in Appendix H). Intuitively, for two clients that are not collinear, their local observations on any arm cannot be utilized to improve each other's knowledge of their own local models. As a result, they should not affect each other's decision-making process. For a collinearly correlated policy, the regret in this setting is defined as: $R^{\pi}(T) = \sum_{s=1}^{T} \sum_{i\in[M]} \left(X_{i,k_t^*}^{\top}\theta - X_{i,k_t}^{\top}\theta\right)$.

Privacy: Differential privacy Dwork et al. (2010)is the most prominent notion of privacy. The contextual bandit problem involves two sets of variables that any agent must private to the other participating agents – the available decision sets $\{X_{k,s}\}_{k\in[K],s\in[T]}$ and observed rewards $\{r_s\}_{s\in[T]}$. In our problem, we are concerned with preserving privacy of continual observations Dwork et al. (2010) under different communication protocols, center and federated and bandit feedback. We instead use *joint differential privacy* (JDP), first introduced by Kearns et al. (2014) in the context of algorithmic mechanism design, and later extended to the online bandit setting by Shariff and Sheffet (2018). This slight relaxation of differential privacy allows the t-th component of the output (i.e., k_t) to depend arbitrarily on the t-th component of the input (i.e., $\mathcal{X}_t$), while ensuring differential privacy with respect to the joint distribution of all other components of the output.

Definition 1 (Joint Differential Privacy (Kearns et al., 2014; Shariff and Sheffet, 2018))**.** *A streaming algorithm $\mathcal{A} : \mathcal{H}^T \to [K]^T$ is said to be (ε, δ)-jointly differential private, for any t-neighboring streams $\mathcal{S}, \mathcal{S}'$ and any $\mathcal{T} \subseteq [K]^{T-1}$,*

$$\Pr[\mathcal{A}(\mathcal{S})_{\neq t} \in \mathcal{T}] \leq \exp(\varepsilon) \Pr[\mathcal{A}(\mathcal{S}')_{\neq t} \in \mathcal{T}] + \delta,$$

where $\mathcal{A}(\mathcal{S})_{\neq t}$ denotes all portions of the algorithm's output except at time t.

The adversary model assumed here is to prevent any two colluding agents j and k to obtain non-private information about any specific element in agent i's history. Therefore, the context set $\mathcal{X}_{i,t}$ and outcome $r_{i,s}$ are sensitive variables that the user trusts only with the agent i. Hence, we wish to keep $\mathcal{X}_{i,t} t \in [T]$ private. However, the agent only stores the chosen actions $X_{i_t,t}$ (and not all of $\mathcal{X}_{i,t}$), and hence making our technique differentially private with respect to $((x_{i,t}, r_{i,t})), t \in [T]$ will suffice.

Definition 2 ((ϵ, δ)-Fed-JDP (Dubey and Pentland, 2020))**.** *In a federated learning setting with $M \geq 2$ agents, a randomized multi-agent contextual bandit algorithm $A = (A_i), M_i = 1$ is (ϵ, δ, M)-federated differentially private under continual multi-agent observation if for any i, j s.t. $i \neq j$, any set of sequences $\mathbf{S}_i = (S_k)_{k=1}^{M}$ and $\mathbf{S'}_i = (S_k)_{k=1,k\neq i}^{M} \cap S'_i$ such that S_i and S'_i are neighboring and any subset of actions $S_j \subset \mathcal{D}_{j,1} \times \mathcal{D}_{j,2} \ldots \times \mathcal{D}_{j,T}$ it holds that*

$$\mathbb{P}\left(A_j(\mathbf{S}_i) \in \mathcal{S}_j\right) \leq \exp(\epsilon)\mathbb{P}\left(A_j(\mathbf{S'}_j) \in \mathcal{S}_j\right) + \delta$$

Remark 2.1 (Item-level vs. User-level DP in Federated Learning)**.** *Fed-JDP definition aims to protect each item from any of the users against the others. Thus, it is often called item-level DP and is widely studied in Federated bandits (Dubey and Pentland, 2020; Huang et al., 2021). Huang et al. (2023) proposes a stronger notion where privacy of the entire history is considered. This notion of user-level DP is out of the scope of present work.*

Remark 2.2 (Local DP vs. JDP in High-dimensions)**.** *Wang et al. (2020); Zhou and Chowdhury (2024) propose to use local DP for low-dimensional linear contextual bandits due to its strong privacy-preserving properties. But in high-dimensional settings, it causes significant degradation of utility because non-private high-dimensional regression depends on $\Omega(\sqrt{s \log d})$ but high-dimensional regression with local DP depends on $\Omega(\sqrt{d})$. Here, s is the sparsity parameter and $s \ll d$. In contrast, high-dimensional regression with JDP shows $\Omega(\sqrt{s} \log d)$ dependence (Chakraborty et al., 2024). Thus, we focus on JDP.*

High-dimensional Contextual Bandits. High-dimensional bandits with sparse linear structure are extensively studied (Bastani and Bayati, 2020; Oh et al., 2021; Hao et al., 2020). An algorithm for this problem follows a generic template. At every step t, (i) Observe context x_t, (ii) Construct an estimator, $\hat{\theta}_t$ from collected data, i.e. rewards and contexts $\{\{x_s, r_s\}_{s=1}^{t-1}$, (iii) Play arm greedily using $\hat{\theta}_t$ and observed context x_t, and (iv) Observe rewards r_t and store data. In this setting, the estimators of θ are often constructed using by solving the LASSO problem, i.e., $\min_{\theta,z} \frac{1}{t-1}\sum_{s=1}^{t-1}\left(r_s - \theta^\top x_s\right)^2 + \lambda\|z\|_1$ s.t.$\theta - zI_{d\times d} = 0$.For the private case, private thresholding-based mechanisms and Peeling algorithms Dwork et al. (2014b) are typically used. The analysis of these algorithms in centralized and federated, private and non-private settings differ significantly from each other motivating the following question: *Can we design an ADMM-based private LASSO algorithm, which is amenable to central, federated models with and without privacy for the high-dimensional bandit problem?*

Assumptions. We describe the assumptions that we use to establish regret guarantees of HiBPA. All of these assumptions are standard in high-dimensional bandit literature Zhang and Huang (2008); Zhang (2010); Bastani and Bayati (2020); Li et al. (2021); Chakraborty et al. (2023, 2024).

Assumption 2.1. *(a) Bounded context:* $\mathbb{P}_{x\sim\mathbb{P}_i}\left(\|x\|_\infty \le x_{\max}\right) = 1,\ \forall i \in [K]$ *and* $x_{\max} \in \mathbb{R}^+$. *(b) Bounded parameters:* $\|\theta\|_0 \le s_0$ *and* $\|\theta\|_1 \le s_{\max}$.

Assumption 2.2 (Observational Noise)**.** *We assume that the random variables* ϵ_t *are independent and each one is* σ *sub-Gaussian, i.e.,* $\mathbb{E}\left[\exp(\alpha\epsilon_t)\right] \le \exp\left(\frac{\sigma^2\alpha^2}{2}\right)$ *for all* $t \in [T]$ *and* $\alpha \in \mathbb{R}$.

Definition 3 (s-sparse eigenvalues)**.** *For a symmetric matrix* A*, its minimum and maximum* s*-sparse eigenvalues are defined as* $\phi_{\min}(s, A) = \inf_{u:u\neq 0, \|u\|_0\le s} \frac{u^\top A u}{\|u\|_2^2}$, $\phi_{\max}(s, A) = \sup_{u:u\neq 0,\|u\|_0\le s} \frac{u^\top A u}{\|u\|_2^2}$.

Assumption 2.3 (Context distribution)**.** *The contexts corresponding to each arm* i *are generated stochastically from* P_i *that satisfies the following conditions.*
1. Bounded Orlicz norm: For any arm $i \in [K]$*, the Orlicz norm of context distribution* P_i *is bounded, i.e.* $\|X_i\|_{\psi_2} \le \nu$ *for* $X \sim P_i$.
2. Sparsity: For all $i \in [K]$*, the design matrix* $\Sigma_i = \mathbb{E}_{x\sim\mathbb{P}_i}\left[xx^\top\right]$ *has bounded maximum sparse eigenvalue, i.e.* $\phi_{\max}\left(Cs^*, \Sigma_i\right) \le \phi_u \le \infty$.

Assumption 2.4 (Margin condition)**.** *There exists a positive constant* $\Delta_*, A, \alpha \in [0, \infty)$*, such that for an* $h \in [A\sqrt{\frac{\log d}{T}}, \Delta_*]$ *and* $\forall t \in [T]$*, the following holds* $\mathbb{P}\left(x_{k_t^*}^\top\theta \ge \max_{i\neq k} x_i^\top\theta + h\right) \le \left(\frac{h}{\Delta^*}\right)^\alpha$.

The zero-mean sub-Gaussian assumption on the noise (Assumption 2.2) is satisfied by various families of distributions, including normal distribution and bounded distributions, which are commonly chosen noise distributions. The margin condition (Assumption 2.4) controls the hardness of the bandit instance. For $\alpha \to \infty$, there is a deterministic gap between arms. $\alpha = 0$ implies that there is no apriori information about arm separation. The sparsity assumption (Assumption 3) on context distributions is needed to identify the sparse support of parameters (Chakraborty et al., 2023).

3 PeeledLASSO: Private ADMM with Peeling

In this section, we propose a template of designing PeeledLASSO algorithm for central and federated settings using ADMM-based optimizers (Boyd et al., 2011) and Peeling algorithm (Dwork et al., 2014b) in Algorithm 1. To obtain private LASSO estimates, we apply Peeling operator on each ADMM update. Introducing the dual variable $u = (u_1, \ldots, u_N) \in R^{N\times d}$ initialized to u_0 and exploiting the separable structure of the consensus problem, we obtain the updates for the i^{th} iteration of ADMM:

$$\hat{\theta}_i = \mathcal{P}_s\left(\frac{1}{N}\sum_{j=1}^{N} u_{i-1,j}, \sigma^2\right), z_{i,j} = \text{prox}_{\gamma,\frac{1}{N}(r_j - x_j^\top\cdot)^2}\left(2\hat{\theta}_i - u_{i-1,j}\right), \tag{1}$$

$$u_{i,j} = u_{i-1,j} + 2\lambda(z_{i,j} - \hat{\theta}_{i+1}) \tag{2}$$

where, $\text{prox}_{\gamma,f}(v_0) = \arg\min_v\left\{f(v) + \frac{1}{2\gamma}\|v - v_0\|_2^2\right\}$. The updated variable θ is given by the peeling operator denoted by $\mathcal{P}_s$, which can be seen as a noisy projection on an ℓ_0-ball of radii s,

Algorithm 1 `PeeledLASSO`$(\{(\mathbf{r}_{v,N}, \mathbf{X}_{v,N})\}_{v\in[M]}, \sigma^2, B, \lambda, \gamma, M, s)$

1: **Input:** Noise variance σ^2, #iterations B, $\gamma \in (0, 1]$, λ, s
2: **Initialize** Collect $\{u_{0,v}\}_{v=1}^{M}$ and set $\hat{\theta}_1 = \frac{1}{M}\sum_{v=1}^{M} u_{0,v}$
3: **for** $i = 1 : B$ **do**
4: **for** $v \in [M]$ **do**
5: $u_{i,v} \leftarrow$ `LocalUpdate`$(\mathbf{X}_v, \mathbf{r}_v, \theta_i, M, \lambda, \gamma)$ //User level computations//
6: **end for**
7: $\hat{\theta}_{i+1} \leftarrow \mathcal{P}_s(\frac{1}{M}\sum_{v\in[M]} u_{i+1,v}, \sigma^2)$ //Server level computations//
8: Communicate $\hat{\theta}_{i+1}$ to each user
9: **end for**
10: **Return:** θ_B

Algorithm 2 `LocalUpdate`$(\mathbf{X}_v = \{X_j\}_{j=1}^{N}, \mathbf{r}_v = \{r_j\}_{j=1}^{N}, \theta_i, M, \lambda, \gamma)$

1: **Parameters:** Clipping parameter $C = 4r_{\max}X_{\max}$
2: **Initialize:** $u_{0,v}$
3: $\Sigma_v \leftarrow \left(\frac{2\gamma}{MN}\mathbf{X}_v\mathbf{X}_v^\top + \mathbf{I}_d\right)$
4: $z_{i,v} \leftarrow \Sigma_j^{-1}\left(2\gamma\mathbf{r}_v\mathbf{X}_v + u_{i,v} - 2\hat{\theta}_i\right)$
5: $u_{i,v} \leftarrow u_{i-1,v} + 2\lambda\left(\text{Clip}(z_{i,v} - \hat{\theta}_i, C)\right)$
6: **Return:** $u_{i,v}$

where noise variance σ^2 is calibrated to ensure the desired privacy level. We refer to Algorithm 5 in Appendix B for details. From these updates and together with the possibility to randomly sample the blocks in our general scheme, we can naturally obtain different variants of ADMM for the centralized and federated models in their private and non-private counterparts. We observe this universality of ADMM scheme as lines 5-7 of Algorithm 1 extends to Algorithm 4. We further use this flexibility to construct the PeeledLASSO estimators for both federated and central settings.

Proposition 1 (Privacy Guarantees)**.** *Under Assumptions 2.1 and 2.2, Algorithms 1 and 2 satisfy* (ϵ, δ)*-JDP (for* $M = 1$*) and* (ϵ, δ)*-FedJDP for* $\sigma^2 = \frac{24s\sigma_\eta^2 \log MN \log(\frac{1}{\delta})}{M^2N^2\epsilon^2}$*, where* σ_η^2 *is the variance of observational noise in the bandit model.*

Theorem 1 (Estimation Error)**.** *We are given* N *samples and* σ^2 *is tuned as per Proposition 1 to ensure* (ϵ, δ)*-JDP and* (ϵ, δ)*-Fed-JDP in centralised and federated settings, respectively.*

(a) ***In central setting****, Algorithm 1 yields estimation error* $\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2\right]$ *of* $\tilde{\mathcal{O}}\left(\frac{s^*\log d}{\sqrt{N_\ell}} + \frac{s^*\log d}{N_\ell\epsilon}\right)$.

(b) ***In federated setting****, Algorithm 1 yields estimation error* $\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2\right]$ *of* $\tilde{\mathcal{O}}\left(\frac{s^*\log d}{\sqrt{MN_\ell}} + \frac{s^*\log d}{\epsilon MN_\ell}\right)$.

We emphasize the estimation error in Theorem 1 is established differently from that in antecedent work such as Bastani and Bayati (2020); Oh et al. (2021). In particular, we analyze ADMM as a fixed-point iteration viewing peeling and regression as proximal operators (see Appendix D for the complete proof).

Discussions. We see that in the central and non-private case, the estimation error decreases as $\frac{1}{\sqrt{N}}$ and $\frac{1}{\epsilon N}$ for the private case. Due to homogeneity of contexts, the estimation error scales as $\frac{1}{M}$ with the number of clients in the private case and scales $\frac{1}{\sqrt{M}}$ in the non-private case. It remains an interesting open problem to establish the tightness of these error bounds by establishing suitable lower bounds and is left as future work. Further, a critical observation here is that we are not adding any noise during the dual update (step 5 in Algorithm 4). While this works in the low-dimensional setting, it would not lead to good utility guarantees for the high-dimensional problem since the noise variance would be proportional to d leading to the utility decay as $O(T)$. Further, we would like to remark that while the bounds presented in Theorem 1 are under the ℓ_2 norm, since the sparsity of θ is

Algorithm 3 HiBPA: High-dimensional Bandits with Peeled ADMM

Input: Privacy level ϵ, δ, users $[M]$, sparsity upper bound s, γ
Require: bounds $b_{\max}, X_{\max}$, sub-gaussian parameter σ_η^2
Initialize: $t_0 = 0$, $\ell = 0$, λ_0
if $\ell = 0$ **then**
  for user $v \in [M]$ **do**
    Play random arm $k_{v,1} \in [K]$, and observe $r_{v,\ell}$
    Set $\mathbf{X}_{v,0} = \{X_{v,k_{v,1}}\}, \mathbf{r}_{v,0} = \{r_{v,1}\}, t_1 = 1$
  end for
end if
for $\ell = 1, 2, \ldots$ **do**
  Set $\sigma_\ell^2 = \frac{24\overline{s}\sigma_\eta^2 \log(M(t_\ell - t_{\ell-1}))\log(\frac{1}{\delta})}{M^2(t_\ell - t_{\ell-1})^2\epsilon^2}$, clipping levels C_ℓ, #iterations B_ℓ, reg. γ_ℓ, dual step λ_ℓ
  $\hat{\theta}_\ell \leftarrow \texttt{PeeledLASSO}(\{\mathbf{X}_{v,\ell-1}, \mathbf{r}_{v,\ell-1}\}_v, \sigma_\ell^2, B_\ell, C_\ell, \lambda_\ell, \gamma, M)$
  Set $t_{\ell+1} = 2^{\ell+1}$
  for user $v \in [M]$ **do**
    $(\mathbf{X}_{v,\ell}, \mathbf{r}_{v,\ell}) \leftarrow \texttt{CollectBanditData}(t_{\ell+1} - t_\ell, \hat{\theta}_\ell, v)$
  end for
end for

bounded by s^*, these bounds can be translated in terms of s (i.e., independent of d). This fact will be used in the proof of Theorem 3, 4.

Remark 3.1 (Choices of Adding Noise)**.** *The noise to ensure privacy could be added at different points of communication, e.g. Step 5, 8, or 9 in Algorithm 1. We chose to add it in Step 5 after updating θ_i at each iteration i. This is a conscious choice as the ADMM iteration can be stopped at any time by a practitioner and in federated setting, θ_i is communicated to each user after each update. If we add noise in the other two steps, we have to calibrate it across each user and each data point. This leads to sampling noise more number of times. Adding it to θ_i minimizes the process while ensuring both privacy and sparsity.*

Remark 3.2 (Peeling vs. Soft Thresholding)**.** *In some variants of private LASSO with ADMM, soft thresholding is used in Step 5 of Algorithm 1 and Step 11 of Algorithm 1. Though this works perfectly in offline setting (Cyffers et al., 2023), in our bandit setting, we need to control the sparsity of PeeledLASSO estimate at any point of time (Chakraborty et al., 2024; Shukla, 2024). This is essential to ensure eventual recovery of support of the true parameter, and thus, correctness of the final bandit algorithm. While soft-thresholding does not exhibit this property, we use Peeling to ensure these requirements.*

4 Algorithm Deisgm: High-dimensional Bandits with PeeledLASSO

In this section, we present our main algorithm for privately learning high-dimensional bandits, (Algorithm 3). HiBPA is an episodic algorithm that computes θ_t^{priv} privately using ADMM updates. We divide the decision horizon into episodes of geometrically progressive lengths. Specifically, the ℓ^{th} episode begins at time 2^ℓ and the underlying algorithm restarts. In the private case, the doubling trick helps us reduce the amount of noise needed to preserve the privacy of the data. Since there are only logarithmic number of episodes. At each time t, the algorithm observes the set of context vectors $\mathcal{X}_t$ using $\hat{\theta}_t^{\text{priv}}$. The estimator $\hat{\theta}_t^{\text{priv}}$ is computed depending on the communication protocol used. Using this estimator, it selects an arm greedily at the current time-step. For both these subroutines, we use the forgetfulness technique, wherein the data collected *only* from the previous episodes is used to compute the estimates of the current episode. Given the estimator, we play an arm greedily using $\texttt{CollectBanditData}(N, \hat{\theta}, v)$ and collect resulting data.

4.1 Central model

In the centralized model, a trusted curator holds the dataset $\{(X_s, r_s)\}_{s=1}^t$ and creates a private estimator from this dataset. Our private ADMM algorithm for this centralized model follows the updates in (1) the `CentralOPLASSO` subroutine. The `CentralOPLASSO` sub-routine takes as input a random vector and uses initial data to compute admm-updates. These u-variable updates are then

Algorithm 4 `CollectBanditData`$(N, \hat{\theta}, v)$

Initialize $\mathbf{X}_{v,\ell} = \emptyset$, $\mathbf{r}_{v,\ell} = \emptyset$
for steps $t = 1, \ldots, N$ **do**
  Play $k_{v,t} = \arg\max_{k \in [K]} x_{v,k}^\top \hat{\theta}$
  Observe reward: $r_{v,t} = X_{v,k_{v,t}}^\top \theta + \eta_t$
  $\mathbf{X}_{v,\ell} = X_{v,\ell} \cup \{X_{v,k_{v,t}}\}$, $\mathbf{r}_{v,\ell} = \mathbf{r}_{v,\ell} \cup \{r_{v,t}\}$
end for
Return

peeled *every iteration* to compute the final θ update. One could add it after completing a pre-decided number of iterations but in practice one can stop the optimization prematurely. In this case, if we publish and use the estimated parameter, it will lead to privacy leakage. Thus, we claim to add it after each iteration. Additionally, as we will see in the next section, this is a mandatory requirement in federated model. Thus, peeling the estimated parameter after each iteration allows us to formulated a unified algorithmic template and analysis without losing anything in utility.

4.2 Federated model

For the Federated Learning (FL) model Kairouz et al. (2021), we assume that there exists a central server that coordinates the behaviors of all the different clients. The server has access to the same partition of the parameter space used by all the clients, and can communicate with the clients. Every client is a K-armed high-dimensional contextual bandit problem, where the true parameter is shared across clients. Due to privacy concerns, the client-side algorithm should keep the reward of each evaluation confidential. The only things that can be transmitted to the server are the local statistical summary of the rewards. The clients are not allowed to communicate with each other. Accordingly, the `FedOPLASSO` subroutine samples a set of users and uses local data to perform admm-updtes. Then, the u-variables from every user are aggregated and peeled in order to obtain the global theta updates. These updates are executed in a federated fashion since (i) the blocks x_k and u_k associated to each arm k can be updated and perturbed locally and in parallel, and (ii) if each arm k shares $u_{t+1} - u_t$ with the server, then the latter can execute the rest of the updates to compute z_{t+1}. On top of this vanilla version, we can natively accommodate user sampling (often called "client sampling" in the literature), which is a key property for cross-device FL as it allows to improve efficiency and to model partial user availability Kairouz et al. (2021). The communication cost of such updates is $Md \log T$ over T episodes if all users participate at every step. Finally, in the federated setup it is important to note that parameters λ, γ are tuned centrally not locally since the final updates happen centrally.

Numerical Experiments: We compare HiBA with existing baselines and show that it performs competitively in settings considered in the paper. Due to space constraints all our experimental studies are deferred to Appendix **??**.

5 Privacy and Utility Guarantees of HiBPA

Now, we establish privacy and utility guarantees for HiBPA in both central and federated models.

Theorem 2 (Privacy Guarantees)**.** *Under Assumption 2.1 and 2.2, Algorithm 3 preserves* (ϵ, δ)*-JDP and* (ϵ, δ)*-Fed-JDP in the central and federated models, respectively.*

HiBPA satisfies due to two components. First, the parameters estimated in each episode ℓ using `CentralOPLASSO` and `FedOPLASSO` are already (ϵ, δ)-JDP and (ϵ, δ)-Fed-JDP due to Proposition 1. Second component is forgetting with doubling episodes as proposed by (Azize and Basu, 2022). Since estimated parameters are the only thing communicated to the server in each episode and are used for further data collection, ensuring their privacy protects the data collected in the previous episode. Additionally, since we do not reuse data of any past episode except the last one, their privacy is not leaked any further. Additionally, HiBPA also enjoys reward-DP ensuring privacy protection for collected downstream rewards. In the reward-DP model, the adversary is only allowed to perturb the reward stream (Hanna et al., 2024; Azize and Basu, 2024). In Theorem 3 and 4, we establish utility guarantees for HiBPA under both central and federated models, and also with and without (ϵ, δ)-JDP.

Theorem 3 (Central model)**.** *Let us set* $C_\ell = x_{\max} b_{\max} + 2\sigma\sqrt{\log N_\ell}$, $\lambda_\ell = \frac{1}{1-\sqrt{\gamma_\ell}}$.

1. *Non-private: Under Assumption 2.1, 2.2, 2.3, and 2.4, Algorithm 3 with parameters* $\gamma_\ell = \sqrt{\frac{\log dN_\ell}{N_\ell}}, B_\ell = \mathcal{O}\left(\log_{1/(1+\gamma_\ell)}(N_\ell)\right)$ *incurs a regret*

$$\mathbb{E}[R(T)] \leq x_{\max} b_{\max} s \log d + \psi(\alpha),$$

such that $\psi(\alpha) = \frac{T^{(1-\alpha)/2}}{(1-\alpha)\Delta_*^\alpha}\left(s^3 \log d\right)^{\frac{1+\alpha}{2}},\ \alpha \in [0,\infty),\ s^{3/2}\log d,\ \alpha = \infty.$

2. *Private: Under assumptions 2.1, 2.2, 2.3, and 2.4, Algorithm 3 with parameters* $\gamma_\ell = \sqrt{\frac{\log dN_\ell}{N_\ell}}, B_\ell = \mathcal{O}\left(\log_{1/(1+\gamma_\ell)}\left(N_\ell\epsilon^2\right)\right), \sigma_\ell^2 = \frac{24s\sigma_\eta^2 \log(\frac{N_\ell}{\delta})}{(N_\ell\epsilon)^2}$ *incurs a regret*

$$\mathbb{E}[R(T)] \leq x_{\max} b_{\max} s \log d + \psi(\alpha) \wedge \phi(\epsilon),$$

such that $\psi(\alpha)$ *is as above and* $\phi(\epsilon) = \Psi_\alpha(T)\left(\frac{s^3 \log^2 d \log(\frac{1}{\delta})}{\epsilon^2}\right)^{\frac{(1+\alpha)}{2}},\ \alpha \in [0,\infty)$, *and* $\frac{s^3 \log^2 d \log(\frac{1}{\delta})}{\epsilon^2},\ \alpha = \infty$. *For* $\alpha > 0$, $\Psi_\alpha(T) \lesssim \frac{\Gamma(3+2\alpha)}{\Delta_*^\alpha(\alpha \log 2)^{3+2\alpha}}$, *and* $\Psi_0(T) \lesssim \log^3 T$.

Discussions: Implications of Theorem 3. *1. Dependence on dimension and sparsity:* We observe that with respect to the dimension the privacy term dominates the non-privacy term by a factor of $(\log d)^{1+\alpha}/2$. In contrast, they depend similarly on the sparsity parameter $s^\star + s$, i.e. $(s + s^\star)^{\frac{3(1+\alpha)}{2}}$. Finally, note that the regularization parameter is independent of the privacy level since in PeeledLASSO privacy is achieved through the peeling step.

2. Dependence on T*:* The non-private depends on $T^{-\frac{1+\alpha}{2}}$ whereas the private part has approximately $\log^3 T$ dependence. Thus, as $T \to \infty$, the effect of T fades away. Additionally, for $\alpha > 0$, the private term of regret is T independent. The last two observations resonate with the similar impact of privacy in other bandit problem such as finite-armed (Azize and Basu, 2022), linear (Azize and Basu, 2024), and stochastic low-dimensional contextual bandits (Azize and Basu, 2024; Hanna et al., 2022).

3. Minimax optimality: For $\alpha = 0$, Chakraborty et al. (2024) provide the lower bound for (ϵ, δ)-JDP high-dimensional linear bandits in the central model. We observe that our regret is *order-optimal in terms of* T *and dimension both with and without privacy*. For sparsity, the non-private and private regret bounds has an additional multiplicative $s + s^*$ factor. Now, we state the regret upper bounds of HiBPA for non-private and private federated model.

4. *Regularization for Exploration.* Here, we adaptively tune the regularization parameter of LASSO as $\sqrt{\frac{\log dN_\ell}{N_\ell}}$. This is commonly used in hihg-dimensional bandits for enforcing exploration (Oh et al., 2021; Ariu et al., 2020).

Theorem 4 (Federated model)**.**

1. *Non-private: Under assumptions 2.1, 2.2, 2.3, and 2.4, Algorithm 3 with parameters* $\gamma_\ell = \sqrt{\frac{\log dN_\ell}{N_\ell}}, B_\ell = \mathcal{O}\left(\log_{1/(1+\gamma_\ell)}(MN_\ell)\right)$ *incurs a regret*

$$\mathbb{E}[R(T)] \leq x_{\max} b_{\max} s \log d + \psi(\alpha),$$

such that $\psi(\alpha) = \frac{(MT)^{(1-\alpha)/2}}{(1-\alpha)\Delta^\alpha}\left(s^3 \log d\right)^{\frac{1+\alpha}{2}},\ \alpha \in [0,\infty)$, *and* $s^3 \log d,\ \alpha = \infty$.

2. *Private: Under assumptions 2.1, 2.2, 2.3, and 2.4, Algorithm 3 with parameters* $B_\ell = \mathcal{O}\left(\log_{1/(1+\gamma_\ell)}\left(MN_\ell\epsilon^2\right)\right), \sigma_\ell^2 = \frac{24s\sigma_\eta^2 \log(MN_\ell)\log(\frac{1}{\delta})}{(MN_\ell\epsilon)^2}$ *incurs a regret*

$$\mathbb{E}[R(T)] \leq x_{\max} b_{\max} s \log d + \psi(\alpha) \wedge \phi(\epsilon),$$

such that $\psi(\alpha)$ *is as above and* $\phi(\epsilon) = \Psi_\alpha(MT)\left(\frac{s^3 \log^2 d \log(\frac{1}{\delta})}{\epsilon^2}\right)^{\frac{(1+\alpha)}{2}},\ \alpha \in [0,\infty)$, *and* $\frac{s^3 \log^2 d \log(\frac{1}{\delta})}{\epsilon^2},\ \alpha = \infty$.

Discussion: Dependence of HiBPA on M**.** We observe that without privacy regret of HiBPA in federated model varies with $M^{\frac{1-\alpha}{2}}$. For $\alpha = 0$, it is $\sqrt{M}$ and fades away as $M \to \infty$. In contrast,

due to the nature of Ψ_α being a bounded function, the privacy term of regret is independent of M for $\alpha > 0$. For $\alpha = 0$, it depends on $\log^3 M$. In the next section, we prove a lower bound on regret of federated bandit algorithms for stochastic high-dimensional linear contextual bandits satisfying (ϵ, δ)-Fed-JDP. The lower bounds indicate that non-private regret of HiBPA achieves order-optimal dependence on M for $\alpha = 0$, whereas it is loose by a $\log^3 M$ factor for regret term due to privacy.

6 Lower Bound

In this section, we report the cost of privacy in the high-dimensional bandit in the feederated model with privacy constraints. We propose a lower bound on the regret of any (ϵ, δ)-JDP algorithm for the high-dimensional bandit problem. This will shed further light on the cost of privacy in these settings when compared with other existing non-private lower bounds.

Definition 4 (Minimax regret)**.** *We define the minimax regret in the (ϵ, δ)-JDP setting as:*

$$R^{\text{minimax}}_{(\epsilon,\delta)} := \inf_{\pi \in \Pi_{(\epsilon,\delta)}} \sup_{P_{x,r} \in \mathcal{B}} \mathbb{E}\left[R(T)\right] ,$$

where $\mathcal{B}$ denotes the set of all sparse high-dimensional bandit instances satisfying Assumptions 2.1 and 2.2, and $\Pi_{(\epsilon,\delta)}$ denotes the set of all collinearly dependent (ϵ, δ)-Fed-JDP policies.

Following Huang et al. (2021), we consider the collinear policies to rigorously proof the lower bound. Intuitively, if two clients that are not collinear under a policy, their local observations on any arm a cannot be utilized to improve the others' knowledge about the arm. Thus, it is equivalent to running M independent bandits. The policy used by HiPBA to interact is also a collinear policy.

Theorem 5 (Lower bound: Federated, Private)**.** *Give homogeneous context distributions across users, privacy parameters $\epsilon, \delta > 0$ such that $\epsilon^2 < \log(\frac{1}{\delta})$, and $d > 9$ and sufficiently large $s\log(d/s)$, we have that*

$$R^{\text{minimax}}_{(\epsilon,\delta)}(T) = \Omega\left(\max\left\{ \underbrace{\sqrt{\frac{s^* \log^2(d/s^*)\log(1/\delta)}{\epsilon^2}}}_{\textit{private and high-d}}, \underbrace{\sqrt{s^* MT \log(d/s^*)}}_{\textit{non-private and high-d}} \right\} \right).$$

Discussions: *1. Impact of M.* This shows that for federated bandit with M users with homogeneous context distributions the non-private regret scales with $\tilde{\Omega}(\sqrt{M})$. HiPBA achieves this scaling w.r.t. M. In low-dimensional federated bandits settings, $\sqrt{M}$ dependence is exhibited by (Zhou and Chowdhury, 2024) but there exists no lower bound to support this phenomenon. Our lower bound shows that $\sqrt{M}$ is the optimal dependence.

2. Private vs. Non-private Bound. In contrast to the non-private lower bound, the private lower bound does not depend on M and T. This indicates a regime change depending on the privacy budget ϵ and the non-private bound dominates for $\epsilon \geq \sqrt{\frac{\log d}{MT}}$. For $M = 1$, we retrieve the lower bound of (Chakraborty et al., 2024) for sparse high-dimensional bandits in centralised setting.

3.*Context Homogeneity.* Homogeneity of contexts is fundamentally important in this lower bound. Since we collect sample from M users at each step and their context distributions are homogeneous, conceptually we can treat them together as MT samples from the bandit environment. This intuition is reflected in the lower bound. Though homogeneity is prevalent in federated bandit literature (Dubey and Pentland, 2020; Wang et al., 2023, 2020), context heterogeneity is gaining interest due to its practicality (Blaser et al., 2024). It is still an open problem to develop lower bounds for federated bandits with heterogeneous users and we leave it as a future work.

7 Conclusion

In this paper, we considered the high-dimensional bandit problem under different communication models and privacy constraints. We proposed a unifying algorithmic design and analysis template that can be used for all these models. This is accompanied by a novel privacy and utility analysis of this framework. We also show that our performance matches the lower bounds up to a factor of s. As a future work, it would be interesting to extend this problem to the case went the clients are heterogeneous.

References

Acharya, J., Sun, Z., and Zhang, H. (2021). Differentially private assouad, fano, and le cam. In *Algorithmic Learning Theory*, pages 48–78. PMLR.

Ariu, K., Abe, K., and Proutière, A. (2020). Thresholded lasso bandit. *arXiv preprint arXiv:2010.11994*.

Asi, H., Feldman, V., Koren, T., and Talwar, K. (2021). Private stochastic convex optimization: Optimal rates in l1 geometry. In *International Conference on Machine Learning*, pages 393–403. PMLR.

Azize, A. and Basu, D. (2022). When privacy meets partial information: A refined analysis of differentially private bandits. *Advances in Neural Information Processing Systems*, 35:32199–32210.

Azize, A. and Basu, D. (2024). Concentrated differential privacy for bandits. In *2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, pages 78–109. IEEE.

Bastani, H. and Bayati, M. (2020). Online decision making with high-dimensional covariates. *Operations Research*, 68(1):276–294.

Bauschke, H. and Combettes, P. (2019). Convex analysis and monotone operator theory in hilbert spaces, corrected printing.

Blaser, E., Li, C., and Wang, H. (2024). Federated linear contextual bandits with heterogeneous clients. In *International Conference on Artificial Intelligence and Statistics*, pages 631–639. PMLR.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122.

Bun, M., Dwork, C., Rothblum, G. N., and Steinke, T. (2018). Composable and versatile privacy via truncated cdp. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 74–86.

Byrne, C. (2003). A unified treatment of some iterative algorithms in signal processing and image reconstruction. *Inverse problems*, 20(1):103.

Cai, T. T., Wang, Y., and Zhang, L. (2020). The cost of privacy in generalized linear models: Algorithms and minimax lower bounds. *arXiv preprint arXiv:2011.03900*.

Carpentier, A. and Munos, R. (2012). Bandit theory meets compressed sensing for high dimensional stochastic linear bandit. In *Artificial Intelligence and Statistics*, pages 190–198. PMLR.

Chakraborty, S., Roy, S., and Basu, D. (2024). Fliphat: Joint differential privacy for high dimensional sparse linear bandits. *arXiv preprint arXiv:2405.14038*.

Chakraborty, S., Roy, S., and Tewari, A. (2023). Thompson sampling for high-dimensional sparse linear contextual bandits. In *International Conference on Machine Learning*, pages 3979–4008. PMLR.

Chaudhuri, K., Monteleoni, C., and Sarwate, A. D. (2011). Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3).

Combettes, P. L. and Pesquet, J.-C. (2021). Fixed point strategies in data science. *IEEE Transactions on Signal Processing*, 69:3878–3905.

Cyffers, E., Bellet, A., and Basu, D. (2023). From noisy fixed-point iterations to private admm for centralized and federated learning. In *International Conference on Machine Learning*, pages 6683–6711. PMLR.

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

Duchi, J. C. and Wainwright, M. J. (2013). Distance-based and continuum fano inequalities with applications to statistical estimation. *arXiv preprint arXiv:1311.2669*.

Dwork, C., Naor, M., Pitassi, T., and Rothblum, G. N. (2010). Differential privacy under continual observation. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 715–724.

Dwork, C., Roth, A., et al. (2014a). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.

Dwork, C., Su, W. J., and Zhang, L. (2018). Differentially private false discovery rate control. *arXiv preprint arXiv:1807.04209*.

Dwork, C., Talwar, K., Thakurta, A., and Zhang, L. (2014b). Analyze gauss: optimal bounds for privacy-preserving principal component analysis. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 11–20.

Eckstein, J. and Yao, W. (2015). Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives. *Pac. J. Optim.*, 11(4):619–644.

Hanna, O., Girgis, A. M., Fragouli, C., and Diggavi, S. (2024). Differentially private stochastic linear bandits:(almost) for free. *IEEE Journal on Selected Areas in Information Theory*.

Hanna, O. A., Girgis, A. M., Fragouli, C., and Diggavi, S. (2022). Differentially private stochastic linear bandits:(almost) for free. *arXiv preprint arXiv:2207.03445*.

Hao, B., Lattimore, T., and Wang, M. (2020). High-dimensional sparse linear bandits. *Advances in Neural Information Processing Systems*, 33:10753–10763.

Huang, R., Wu, W., Yang, J., and Shen, C. (2021). Federated linear contextual bandits. *Advances in neural information processing systems*, 34:27057–27068.

Huang, R., Zhang, H., Melis, L., Shen, M., Hejazinia, M., and Yang, J. (2023). Federated linear contextual bandits with user-level differential privacy. In *International Conference on Machine Learning*, pages 14060–14095. PMLR.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. *Foundations and trends® in machine learning*, 14(1–2):1–210.

Kamath, G., Liu, X., and Zhang, H. (2022). Improved rates for differentially private stochastic convex optimization with heavy-tailed data. In *International Conference on Machine Learning*, pages 10633–10660. PMLR.

Kearns, M., Pai, M., Roth, A., and Ullman, J. (2014). Mechanism design in large games: Incentives and privacy. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 403–410. ACM.

Kifer, D., Smith, A., and Thakurta, A. (2012). Private convex empirical risk minimization and high-dimensional regression. In *Conference on Learning Theory*, pages 25–1. JMLR Workshop and Conference Proceedings.

Lai, T. L., Robbins, H., et al. (1985). Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22.

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

Li, W., Barik, A., and Honorio, J. (2021). A simple unified framework for high dimensional bandit problems. *arXiv preprint arXiv:2102.09626*.

Lions, P.-L. and Mercier, B. (1979). Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979.

McMahan, B., Moore, E., Ramage, D., Hampson, S., and y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR.

Oh, M.-h., Iyengar, G., and Zeevi, A. (2021). Sparsity-agnostic lasso bandit. In *International Conference on Machine Learning*, pages 8271–8280. PMLR.

Raff, E., Khanna, A., and Lu, F. (2024). Scaling up differentially private lasso regularized logistic regression via faster frank-wolfe iterations. *Advances in Neural Information Processing Systems*, 36.

Shariff, R. and Sheffet, O. (2018). Differentially private contextual linear bandits. *Advances in Neural Information Processing Systems*, 31.

Shukla, A. (2024). Differentially private high dimensional bandits. *arXiv preprint arXiv:2402.03737*.

Talwar, K., Guha Thakurta, A., and Zhang, L. (2015). Nearly optimal private lasso. *Advances in Neural Information Processing Systems*, 28.

Tavara, S., Schliep, A., and Basu, D. (2021). Federated learning of oligonucleotide drug molecule thermodynamics with differentially private admm-based svm. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 459–467. Springer.

Wang, C.-H., Li, W., and Lin, G. (2023). Federated high-dimensional online decision making. *Transactions on Machine Learning Research*.

Wang, H., Zhao, Q., Wu, Q., Chopra, S., Khaitan, A., and Wang, H. (2020). Global and local differential privacy for collaborative bandits. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 150–159.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, pages 894–942.

Zhang, C.-H. and Huang, J. (2008). The sparsity and bias of the lasso selection in high-dimensional linear regression. *The Annals of Statistics*, pages 1567–1594.

Zhou, X. and Chowdhury, S. R. (2024). On differentially private federated linear contextual bandits. In *The Twelfth International Conference on Learning Representations*.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: They describe the paper, its motivation and main contribution.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Discussed throughout the presentation of results.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: All proofs are relegated to the appendix and some details are described throughout the text.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Details in appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: Details in the appendix.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Details in the appendix.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Details in the appendix.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Details in the appendix.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper has impact on resource allocation problems of interest to society.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: [NA]

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA]

    Justification: [NA]

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: [NA]

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.

Part I

Appendix

A Related Work

A.1 Existing Literature

Multi-armed bandits have been studied since the foundational work of (Lai et al., 1985) and a comprehensive overview can be found in (Lattimore and Szepesvári, 2020). In this work, we consider the high-dimensional bandit problem (Li et al., 2021; Hao et al., 2020; Carpentier and Munos, 2012), under the necessary assumption that the unknown parameter θ is s^* sparse. (Bastani and Bayati, 2020; Oh et al., 2021; Ariu et al., 2020) propose LASSO or thresholding-based algorithms for this problem with competitive regret guarantees. Extending this line work to ensure privacy requires constructing estimators of the mean reward by injecting well-calibrated noise. This however implies that the variance of noise – and therefore the error in the constructed estimate – scales linearly with the dimension of the parameter. A straightforward adaptation of these tools to the high-dimensional problem setting where $d \gg T$ would lead to a regret bound that scales $\mathcal{O}(d)$, which would implicitly be super-linear in T. Another approach to designing private algorithms is objective or gradient perturbation (Chaudhuri et al., 2011; Kifer et al., 2012). (Talwar et al., 2015; Asi et al., 2021; Cyffers et al., 2023; Raff et al., 2024) consider this problem in the offline setting.

Federated learning (FL) (McMahan et al., 2017) has become a popular distributed machine learning paradigm in which numerous clients collaboratively train a prediction model under the coordination of a central server while maintaining the local training data at each client. FL is motivated by various applications where real-world data are exogenously generated at edge devices, and it is desirable to protect the privacy of local data by only sharing model updates instead of the raw data. In the low-dimensional setting, one of the first works to consider Federated linear contextual bandits was (Huang et al., 2021) and (Dubey and Pentland, 2020) extended this to a Fed-DP guarantee for their algorithm.

B Peeling

Algorithm 5 Peeling

Input: Vector $v \in \mathbb{R}^d$, sparsity s, privacy level ϵ, δ, sensitivity D
Initialise: $S = \emptyset$ and set $\xi = \frac{2D\sqrt{3s\log(1/\delta)}}{\epsilon}$
for $i = 1, 2, \ldots, s$ **do**
  Generate $\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{id}) \overset{iid}{\sim} \text{Lap}(\xi)$
  $j_i^* = \arg\max_{j \in [d] \setminus S} |v_j| + w_{ij}$
  Update $S \leftarrow S \cup \{j_i^*\}$
end for
Set $\mathcal{P}_s = v_S$
Generate $\tilde{\mathbf{w}} = (w_{i1}, w_{i2}, \ldots, w_{id}) \overset{iid}{\sim} \text{Lap}(\xi)$
Return: $\mathcal{P}_s(v; \epsilon, \delta, D) = \mathcal{P}_s(v) + \tilde{\mathbf{w}}_S$

Lemma 1 ((ϵ, δ)-DP Dwork et al. (2018))**.** *If for every pair of adjacent datasets $\bar{D}, \bar{D}'$, we have $\|v(\bar{D}) - v(\bar{D}')\| \leq D$, then Algorithm 5 is (ϵ, δ)-DP.*

We further have the following bounds on that establish that peeling is indeed a contraction Cai et al. (2020).

Lemma 2 (Peeling)**.** *For any index set $I \subset [d]$ and $\theta \in \mathbb{R}^d$ with supp$(\theta) \subset I$ and any $\hat{\theta} \in \mathbb{R}^d$ such that $\|\theta\|_0 \leq s$, we have for every $c > 0$, we have:*

$$\|\hat{\theta} - \theta\|_2^2 \leq \left(1 + \frac{1}{c}\right) \frac{I - s}{I - \hat{s}} \|\hat{\theta} - \theta\|_2^2 + 4(1 + c) \sum_i \|w_i\|_\infty$$

Lemma 3 (Peeling is a contraction)**.** *Given $\theta^* \in \mathbb{R}^d$ with $\|\theta^*\|_0 \leq s$ and θ such that supp$(\theta^*) \subset$ supp(θ) with $\|\theta\|_0 \leq \tilde{s} + s$, for any $c_1, c_2 > 0$ and $c > 0$, we have:*

$$\|\mathcal{P}_s(\theta, W, \tilde{w}) - \theta^*\|_2^2 \leq a\|\theta - \theta^*\|_2^2 + b$$

where,

$$\begin{aligned} a &= \left(1 + \frac{1}{c_1}\right)\left[\left(1 + \frac{1}{c_2}\right)\left(1 + \frac{1}{c}\right) + (1 + c_2)\right] \\ b &= (1 + c_1)\|\tilde{w}_S\|_2^2 + 4\left(1 + \frac{1}{c_1}\right)\left(1 + \frac{1}{c_2}\right)(1 + c)\sum_{i \in [s]}\|w_i\|_\infty^2 \end{aligned}$$

C Preliminaries

Many optimization algorithms such as proximal methods, ADMM belong to this family Bauschke and Combettes (2019). Formulating optimization algorithms as λ-averaged operators allows us to use generic convergence results by the Krasnosel'skii Mann theorem Byrne (2003). In the setting of differential privacy, each application of T is perturbed by noise due to stochasticity in the data. A generic convergence analysis of fixed-point iterations under both inexact and block updates Combettes and Pesquet (2021).

C.1 ADMM

Consider the problem of minimizing a function $f : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X} \subseteq \mathbb{R}^p$, The problem reduces to finding a fixed-point, $x^* \in \mathcal{X}$, such that $0 \in \partial f(x^*)$ or $\nabla f(x^*) = 0$ under differentiability assumptions. Alternatively, we can view these updates as reaching the fixed point of an operator $T : \mathcal{X} \to \mathcal{X}$. Starting with initial point x_0, these updates are given by:

$$x_{t+1} = T(x_t)$$

We now define the type of operators considered in this paper.

Definition 5 (Fixed Point Operator)**.** *Let $T : \mathcal{X} \to \mathcal{X}$ and $\lambda \in [0, 1]$, then:*

1. *T is non-expansive, i.e., $\|T(x) - T(x')\| \leq \|u - u'\|$, $\forall\, x, x' \in \mathcal{X}$. It is τ contractive if it is τ-Lipschitz.*
2. *T is λ-averaged if there exists a contractive operator R such that $T = \lambda R + (1 - \lambda)I$.*

We now present how ADMM can be defined as a fixed-point iteration. ADMM minimizes the sum of two (possibly non- smooth) convex functions with linear constraints between the variables of these functions, which can be formulated as:

$$\begin{aligned} \min_{x,z} \quad & f(x) + g(z) \\ \text{s.t.} \quad & Ax + Bz = c \end{aligned}$$

ADMM is often presented as an approximate version of the augmented Lagrangian method, where the minimization of the sum in the primal is approximated by the alternating minimizations on x and z. However, for the purpose of analysis, it is useful to view ADMM as a splitting algorithm Eckstein and Yao (2015). i.e., an approach to find a fixed point of the composition of two (proximal) operators by performing operations that involve each operator separately.

ADMM is defined through the Lions- Mercier operator Lions and Mercier (1979). Given two proximable functions P_1 and P_2 and parameter $\eta > 0$, the Lions-Mercier operator is: $T_{\eta P_1, \eta P_2} = \lambda R_{P_1} \eta R_{P_2} + (1 - \lambda)I$, where, $R_{P_1} = 2\text{prox}(P_1) - I, R_{P_2} = 2\text{prox}(P_2) - I$. In this work, we study fixed-point iterations with Differential Privacy (DP) Dwork et al. (2014a). DP relies on a notion of neighboring datasets. We denote a private dataset of size n by $D := (d_1, \ldots, d_n)$. Two datasets D, D' are neighboring if they differ in at most one element. We refer to each d_i as a data item. Consider the problems of the form:

$$\min_{x \in \mathcal{X}} \frac{1}{n} \sum_{i=1}^{n} f(u; d_i) + r(u),\ i = 1, \ldots, n \tag{3}$$

where $f(; d_i)$ is a (typically smooth) loss function computed on data item d_i and r is a (typically non-smooth) regularizer. We denote $f(u; D) := \frac{1}{n} \sum_{i=1}^{n} f(u; d_i), i = 1, \ldots, n$. To solve this problem, we consider the general noisy fixed-point iteration described in Algorithm 1. The core of each update applies a λ_k-averaged operator constructed from a non-expansive operator R, and a Gaussian noise term added to ensure differential privacy via the Gaussian mechanism. Algorithm 1 can use (possibly randomized) block-wise updates ($B > 1$) and accommodate additional errors in operator evaluation.

D Proof of Theorem 1

Proof. For every iteration b given by (1) from Lemma 3, we have:

$$\begin{aligned}
\mathbb{E}\left[\|\hat{\theta}_b - \theta^*\|_2^2\right] &= \mathbb{E}\left[\|\mathcal{P}_S(\hat{\theta}_{b+0.5}, \mathbf{W}^{(b)}, \tilde{w}^b) - \theta^*\|_2^2\right] \\
&\leq \mathbb{E}\left[\|\mathcal{P}_S((\hat{\theta}_{b+0.5})_{S_b}, \mathbf{W}^{(b)}, \tilde{w}^b) - \theta^*\|_2^2\right] \\
&\leq c_3\mathbb{E}\left[\|(\hat{\theta}_{b+0.5})_{S_b} - \theta^*\|_2 62\right] + c_4 \\
&\leq c_3\mathbb{E}\left[\|u_b - u^*\|_2\right] + c_4\,,
\end{aligned}$$

where,

$$\begin{aligned}
c_3 &= \frac{6}{7}\left(1+\frac{1}{c_1}\right)\left[\left(1+\frac{1}{c_2}\right)\left(1+\frac{1}{c}\right)\frac{s^*}{s} + (1+c_2)\right] \\
c_4 &= (1+c_1)\mathbb{E}\left[\|\tilde{\mathbf{w}}_S\|_2^2\right] + 4\left(1+\frac{1}{c_1}\right)\left(1+\frac{1}{c_2}\right)(1+c)\mathbb{E}\left[\sum_{i\in[s]}\|\mathbf{w}_i\|_\infty^2\right]
\end{aligned}$$

Now we analyse the convergence step for which we use the analysis from ADMM. For the soft-thresholding operator we have $\tau = \frac{1}{\gamma}$ and $D =$ and B_ℓ is the number of iterations for which ADMM is run in a single episode. From Algorithm 4 and using Theorem 6, we have that:

$$\mathbb{E}\left[\|u_b - u^*\|_2^2\right] \leq \left(1 - \frac{(1-\tau)}{8}\right)^{B_\ell} D$$

Plugging this back, we get:

$$\mathbb{E}\left[\|\hat{\theta}_{B_\ell} - \theta^*\|_2^2\right] \leq c_3\left(1 - \frac{(1-\tau)}{8}\right)^{B_\ell} D + c_4 B_\ell \mathbb{E}\left[\|\tilde{w}_{S_m}^{(m)}\|_2^2 + \sum_{i\in[s]}\|\mathbf{w}_i^{(m)}\|_\infty^2\right]$$

The first step in Algorithm 1 is Peeling Algorithm 5. The noise added at every iteratino in order to preserve the privacy is given by $\mathbf{W}_b = \left(\sum_{i\in[s]}\|\mathbf{w}_i^{(b)}\|_\infty^2 + \|\tilde{\mathbf{w}}_S^{(b)}\|_2^2\right)$ where each $\mathbf{w}_i^{(b)} \in \mathbb{R}^d$ and $\tilde{\mathbf{w}}_S$ has $|S| = s$ non-zero co-ordindates. Noting that $\xi = \frac{\theta_{\max}}{\underline{\kappa}}\frac{\sqrt{3s\log\frac{B_\ell}{\delta}}}{N_\ell\epsilon}$, using Theorem 7, for suitably large constant, we have:

$$\mathbb{P}\left(\|\mathbf{w}_S\|_\infty^2 \geq \frac{s\log^2 d\left(\log\frac{B_\ell}{\delta}\right)}{N_\ell^2\epsilon^2}\right) \leq \frac{1}{d^8}$$

Using a union bound, we get:

$$\begin{aligned}
&\mathbb{P}\left(\max_{b\in[B_\ell]}\mathbf{W}_b > K\frac{s^2\log^2 d\left(\log\frac{B_\ell}{\delta}\right)}{N_\ell^2\epsilon^2}\right) \\
\leq &\sum_{b\in[B_\ell]}\mathbb{P}\left(\mathbf{W}_b > \frac{s^2\log^2 d\left(\log\frac{B_\ell}{\delta}\right)}{N_\ell^2\epsilon^2}\right) \\
\leq &\frac{2sB_\ell}{d^8} = O\left(\frac{1}{d^6}\right)
\end{aligned}$$

Plugging this back, we get:

$$\begin{aligned}
\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2^2\right] &\leq c_3\left(1 - \frac{(1-\gamma)}{8}\right)^{B_\ell} D + c_4 B\mathbb{E}\left(\|\tilde{w}_{S_m}^{(m)}\|_2^2 + \sum_{i\in[s]}\|\mathbf{w}_i^{(m)}\|_\infty^2\right) \\
&\leq c_3\left(1 - \frac{(1-\gamma)}{8}\right)^{B_\ell} D + c_4 B_\ell K' \cdot \frac{(s^*\log d)^2\log\frac{\log B_\ell}{\delta}}{N_\ell^2\epsilon^2}
\end{aligned}$$

1. For the central setting, let B such that

$$\left(1 - \frac{1-\gamma}{8}\right)^{B_\ell} D = K' B_\ell \cdot \frac{(s^*)^2 \log^2 d \log \frac{B_\ell}{\delta}}{N_\ell \epsilon^2}$$
$$B_\ell = \mathcal{O}\left(\log_{1/(1+\gamma)}\left(\frac{N_\ell \epsilon}{(s^* \log d)^2}\right)\right).$$

Ignoring $\log$ and $\log\log$ terms, this leads to:

$$\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2^2\right] \lesssim C^2 \cdot \frac{(s^* \log d)^2}{N_\ell}\left(1 + \frac{1}{\epsilon^2 N_\ell}\right)$$
$$\implies \mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2\right] \lesssim C\left(\frac{s^* \log d}{\sqrt{N_\ell}} + \frac{s^* \log d}{N_\ell \epsilon}\right).$$

2. In the Federated setting,

$$\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2^2\right] \leq c_3\left(1 - \frac{(1-\gamma)}{8}\right)^{B_\ell} D + c_4 B_\ell K' \cdot \frac{(s^* \log d)^2 \log \frac{\log B_\ell}{\delta}}{M^2 N_\ell^2 \epsilon^2}$$

$$B = \mathcal{O}\left(\log_{1/(1+\gamma)}\left(M N_\ell\right)\right).$$

Leading to the following bound error per-iteration:

$$\mathbb{E}\left[\|\hat{\theta}_B - \theta^*\|_2\right] \lesssim \frac{s^* \log d}{\sqrt{M N_\ell}} + K' \cdot \frac{s^* \log d}{\epsilon M N_\ell}.$$

□

E Proof of Theorem 2

Proof. Federated Model: Let $\mathbf{r}_t = (r_{1,t}, r_{2,t}, \ldots, r_{MK,t})$ and $\mathcal{X}_t = (X_{1,t}, \ldots, X_{MK,t})$. Consider two neighboring datasets $\mathcal{D}_1 = \{(\mathbf{r}_t, \mathcal{X}_t)\}_{t\in[T]}, \mathcal{D}_2 = \{\mathbf{r}'t, \mathcal{X}'_t\}_{t\in[T]}$. Let the randomized mapping introduced by Algorithm 1 be $\mathcal{M}_t$. At time t, $\mathcal{M}_t$ takes a set of observed reward-context pairs from the previous episodes and recommends action $k_t \in [K]$. Let $\alpha(t)$ denote the episode to which time t belongs:

$$\alpha(t) := \sum_{\ell=0}^{\lfloor \log_2 T \rfloor} \ell \mathbf{1}\{t_\ell \le t \le t_{\ell+1}\}$$

$$\mathbb{P}\left(\mathcal{M}_\tau = k_{-\tau}\right) = \Pi_{\ell\ge 0}\mathbb{P}\left(k_t() = k_t; t_\ell \le t \le t_{\ell+1}, t \ne \tau\right)$$

Now, we will analyze this product. First note that because of forgetfulness, changing reward and context at time τ only affects the private estimate $\hat{\theta}_{\ell+1}$ as it only depends on the data from the previous episodes. We now prove this case by case

1. Case-I: For a fixed action sequence $k_{-\tau} = \{k_{i,1}, \ldots k_{i,\tau-1}, k_{i,\tau+1} \ldots k_{i,T}\}_{i\in[M]} \cup \{k_{j,\tau}\}_{j\in[M], j\ne m}$ for some $m \in [M]$. We have $\mathcal{H}_{\ell-1} = \mathcal{H}'_{\ell-1}$. Fixing a time t in episode ℓ, by the nature of Algorithm 1, the randomness in $\mathcal{M}_t$ only comes through the estimator $\hat{\theta}_\ell$. Let the distribution function of $\hat{\theta}_\ell$ on dataset $\mathcal{D}$ be F and those on $\mathcal{D}'$ be F'. For a fixed action sequence, since $\mathcal{H}_{\ell-1} = \mathcal{H}'_{\ell-1}$, these distribution functions are identical. Define the set: $\mathcal{O} = \{\hat{\theta} : \mathcal{M}(\hat{\theta}_t, C_t) = k_t\}$, then we have the following for $\ell \le \ell_0$:

$$\begin{aligned}
\mathbb{P}\left(\hat{k}_t = k_t; t_\ell \le t \le t_{\ell+1}\right) &= \int_{\mathbb{R}^d} \mathbf{1}\left(; t_\ell \le t \le t_{\ell 1}; t \ne \tau\right) dF_{\hat{\theta}}(z) \\
&= \mathbb{P}\left(\hat{\theta}_t(\mathcal{H}_{t-1}) \in \cap_{t_{\ell\le t\le t_{\ell+1}}; t\ne\tau}\mathcal{O}_t\right) \\
&= \mathbb{P}\left(\hat{\theta}_t(\mathcal{H}'_{t-1}) \in \cap_{t_{\ell\le t\le t_{\ell+1}}; t\ne\tau}\mathcal{O}_t\right)
\end{aligned}$$

2. Case-II ($\ell > \ell + 1$): This argument follows from the previous case and using the Forgetfulness property, we get:

$$\mathbb{P}\left(\hat{k}_t(\mathcal{H}_t) = k_t; t_\ell \le t \le t_{\ell+1}\right) = \mathbb{P}\left(\hat{k}_t(\mathcal{H}'_t) = k_t; t_\ell \le t \le t_{\ell+1}\right)$$

3. Case-III ($\ell = \ell_0 + 1$): Noting that $\mathcal{H}_{t-1}$ and $\mathcal{H}'_{t-1}$ are neighbouring datasets and that $\hat{\theta}_{t-1}$ is an (ϵ, δ)-DP estimator. Hence,

$$\begin{aligned}
\mathbb{P}\left(\hat{k}_t(\mathcal{H}) = k_t; t_\ell \le t \le t_{\ell+1}\right) &= \int_{\mathbb{R}^d} \mathbf{1}\left(t_{\ell_0+1\le t\le t_{\ell_0}+2}; t \ne \tau\right) dF(\hat{\theta}_\ell)(z) \\
&= \mathbb{P}\left(\hat{\theta}_t(\mathcal{H}_{\ell_0}) \in \cap_{t_{\ell_0}\le t\le t_{\ell_0+12; t\ne\tau}}\mathcal{O}_t\right) \\
&= \exp(\epsilon)\cdot\mathbb{P}\left(\hat{\theta}_t(\mathcal{H}'_{\ell_0}) \in \cap_{t_{\ell_0}\le t\le t_{\ell_0+12; t\ne\tau}}\mathcal{O}_t\right) + \delta \\
&= \exp(\epsilon)\cdot\mathbb{P}\left(\hat{k}_t = k_t; t_\ell \le t \le t_{\ell+1}; t \ne \tau\right) + \delta
\end{aligned}$$

Combining these cases, we have:

$$\begin{aligned}
\mathbb{P}\left(\mathcal{M}_{-\tau}(\mathcal{D}) = a_{-\tau}\right) &= \Pi_{\ell\ge 0}\mathbb{P}\left(\hat{k}_t = k_t; t_\ell \le t \le t_{\ell+1}; t \ne \tau\right) \\
&= \exp(\epsilon)\cdot\Pi_{\ell\ge 0}\mathbb{P}\left(\hat{k}_t = k_t; t_\ell \le t \le t_\ell + 1; t \ne \tau\right) \\
&+ \delta\Pi_{\ell\ne\ell+1}\mathbb{P}\left(\hat{k}_t(\mathcal{H}_t) \ne k_t; t_\ell \le t \le t_\ell + 1; t \ne \tau\right) \\
&= \exp(\epsilon)\cdot\mathbb{P}\left(\mathcal{A}_{-\tau} = \mathcal{M}_{-\tau}\right) + \delta
\end{aligned}$$

The proof for privacy in the central setting can be obtained by setting $M = 1$. □

F Proof of Theorem 3

Proof. Here, $\mathcal{G} \triangleq \cap_\ell \left\{ \max_b \mathbf{W}_b \leq K \frac{s^2 \log^2 d \left(\log \frac{B_\ell}{\delta}\right)}{N_\ell^2 \epsilon^2} \right\}$, which holds with probability $1-\delta$.

The regret decomposition is given by:

$$\begin{aligned}
\mathbb{E}\left[R(T)\right] &= \sum_{s=1}^{t} \mathbb{E}\left[X_{t,k_t^*}^\top \theta^* - X_{t,k_t}\theta^*\right] \\
&= 2X_{\max}\|\theta^*\|_1 \sum_{1\leq \ell \leq L} t_\ell + \sum_{s=t_L+1}^{T} \mathbb{E}\Bigg[\left(X_{t,\theta^*}^\top \theta^* - X_{t,k_t^*}^\top \hat{\theta}_{\ell(t)}\right) + \left(X_{t,k_t^*}^\top \hat{\theta}_{\ell(t)} - X_{t,k_t}^\top \hat{\theta}_{\ell(t)}\right) \\
&\qquad\qquad + \left(X_{t,k_t}^\top \hat{\theta}_{\ell(t)} - X_{t,k_t}^\top \theta^*\right) \Bigg] \\
&= 2X_{\max} b_{\max}(2^L - 1) + \sum_{\ell \geq L} \sum_{t_\ell \leq t_{\ell+1}} \mathbb{E}\left[\left(X_{t,k_t^*}\theta^* - X_{t,k_t^*}\hat{\theta}_{\ell(t)}\right)\mathbf{1}_{\mathcal{G}} + \left(X_{t,k_t}\hat{\theta}_{\ell(t)} - X_{t,k_t}\theta^*\right)\mathbf{1}_{\mathcal{G}}\right] \\
&\qquad + X_{\max} b_{\max} \mathbb{P}(\mathbf{1}_{\overline{\mathcal{G}}}) \\
&\leq 2X_{\max} b_{\max}(2^L - 1) + 2X_{\max} \sum_{\ell \geq L} \sum_{t_\ell \leq t_{\ell+1}} \mathbb{E}\left[\|\theta^* - \hat{\theta}_{\ell(t)}\|_1 \mathbf{1}_{\mathcal{G}}\right] + X_{\max} b_{\max} \mathbb{P}(\mathbf{1}_{\overline{\mathcal{G}}}) \\
&\leq 2X_{\max} b_{\max} 2^L + 2X_{\max}\sigma^2\sqrt{s} \sum_{\ell \geq L} (t_{\ell+1} - t_\ell) \sqrt{K' \frac{(s^*)^2 \log^2 d}{N_\ell}} \\
&+ 2X_{\max}\sigma^2\sqrt{s} \sum_{\ell \geq L} (t_{\ell+1} - t_\ell) \sqrt{K' \frac{s^2 \log^2 d \log N_\ell^2}{\epsilon^2}} \log\left(\frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{D N_\ell}\right) \\
&\leq 2X_{\max} b_{\max} \log T + 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log\log T}{\delta} \sum_{\ell \geq L} \sqrt{K' \cdot N_\ell} \\
&+ 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log\log T}{\delta} \log\left(\frac{s^2 \log^2 d \log \frac{\log\log T}{\delta}}{D}\right) \sqrt{\frac{K'}{\epsilon^2}} \\
&\leq 2X_{\max} b_{\max} \log T + 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log\log T}{\delta} \sqrt{K'T} \\
&+ 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log\log T}{\delta} \log\left(\frac{s^2 \log^2 d \log \frac{\log\log T}{\delta}}{D}\right) \sqrt{\frac{K'}{\epsilon^2}}
\end{aligned}$$

The per-step regret is given by:

$$\begin{aligned}
\Delta_{k_t}(t) &= x_{k_t^*}^\top \theta^* - x_{k_t}^\top \theta^* \\
&\leq x_{k_t^*}^\top \theta^* - x_{k_t^*}^\top \hat{\theta}_{\ell(t)} + x_{k_t^*}^\top \hat{\theta}_{\ell(t)} - x_{k_t}^\top \hat{\theta}_{\ell(t)} + x_{k_t}^\top \hat{\theta}_{\ell(t)} - x_{k_t}^\top \theta^* \\
&\leq \|x_{k_t^*}\|_\infty \|\theta^* - \hat{\theta}_{\ell(t)}\|_1 + \|x_{k_t}\|_\infty \|\theta^* - \hat{\theta}_{\ell(t)}\|_1 \\
&\leq 4\sigma X_{\max} s\sqrt{T \log d} + 4\sigma X_{\max} \sqrt{\frac{s^3 (\log d)^2 \log\left(\frac{1}{\delta}\right) \log^6 T)}{\epsilon^2}}
\end{aligned}$$

where, the last-inequality follows from the previous lemma. Consider the event:

$$\mathcal{H}_t \triangleq \{x_{k_t^*}^\top \theta^* > \max_{k \neq k_t^*} x_k^\top \theta^* + p_t\}$$

Under $\mathcal{M}_\ell \cap \mathcal{A}_{\ell(t)}$, we have the following for any $k \neq k_t$:

$$\begin{aligned}
x_{k_t^*}^\top \hat{\theta}_{\ell(t)} - x_k^\top \hat{\theta}_{\ell(t)} &= \langle x_{k_t^*}, \hat{\theta} - \theta^* \rangle + \langle x_{k_t^*} - x_{k_t}, \theta^* \rangle + \langle x_{k_t}, \theta^* - \hat{\theta}_{\ell(t)} \rangle \\
&\geq -g_t + p_{t-1} - g_t
\end{aligned}$$

Choosing $p_t = 3g_{\ell(t)}$, with $g_{\ell(t)} = \sqrt{K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{N_\ell}} + \sqrt{K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{N_\ell^2 \epsilon^2}} \log\left(K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{D N_\ell}\right)$, we see that:

1. When $\alpha = 0$, we get the previously proven bound.
2. When $\alpha \in (0,1)$:

$$\begin{aligned}\mathbb{E}(R(T)) &\leq 2X_{\max} b_{\max}(2^L - 1) \\ &+ 24X_{\max}\sigma^2 \sum_{\ell \geq L} \Delta_*^{-\alpha} N_{\ell-1} \left[\left(\frac{s^2 \log d}{N_{\ell-1}}\right)^{\frac{1+\alpha}{2}} + \left(\frac{s^3 (\log d)^2 \log(\frac{1}{\delta}) \log \frac{\log N_{\ell-1}}{\delta}}{N_{\ell-1}^2 \epsilon^2}\right)^{\frac{1+\alpha}{2}} \right]\end{aligned}$$

   This implies that:

$$I_\alpha \leq \frac{s^{3(1+\alpha)} (\log d)^{\frac{(1+\alpha)}{2}}}{\Delta_*^\alpha} \left(\frac{T^{\frac{1-\alpha}{2}} - 1}{1-\alpha}\right) + \Psi_\alpha \Delta_*^{-\alpha} \epsilon^{-(1+\alpha)} \left\{\log(1/\delta)\right\}^{\frac{(1+\alpha)}{2}} \left(s^3 \log^2 d\right)^{\frac{1+\alpha}{2}}$$

   where, $\Psi_\alpha = \frac{1-T^{-\alpha}}{1-2^{-\alpha}} (\log T)^{2\alpha+2}$.

3. $\alpha = 1$: The previous bounds implies:

$$I_1 \leq \frac{s^6 (\log d)}{\Delta_*} \log T + \frac{\Psi_1}{\Delta_* \epsilon^2} s^3 \log d \log(\frac{1}{\delta})$$

4. $\alpha > 1$: In this case, we get:

$$I_\alpha \leq \frac{s^{3(1+\alpha)} (\log d)^{\frac{(1+\alpha)}{2}}}{\Delta_*^\alpha} \left(\frac{1 - T^{\frac{-(\alpha-1)}{2}}}{\alpha - 1}\right) + \Psi_\alpha \Delta_*^{-\alpha} \epsilon^{-(1+\alpha)} \left(\log(\frac{1}{\delta})\right)^{\frac{(1+\alpha)}{2}} \left(s^3 \log^2 d\right)^{\frac{(1+\alpha)}{2}}$$

□

G Proof of Theorem 4

Proof. Here, $\mathcal{G} \triangleq \cap_\ell \left\{ \max_b \mathbf{W}_b \leq K \frac{s^2 \log^2 d \left(\log \frac{B_\ell}{\delta}\right)}{N_\ell^2 \epsilon^2} \right\}$, which holds with probability $1-\delta$.

The regret decomposition is given by:

$$\begin{aligned}
\mathbb{E}\left[R(T)\right] &= \sum_{s=1}^{t} \mathbb{E}\left[X_{t,k_t^*}^\top \theta^* - X_{t,k_t}\theta^*\right] \\
&= 2X_{\max}\|\theta^*\|_1 \sum_{1\leq \ell \leq L} N_\ell + \sum_{s=t_L}^{T} \mathbb{E}\left[\left(X_{t,\theta^*}^\top \theta^* - X_{t,k_t^*}^\top \hat{\theta}_{\ell(t)}\right) + \left(X_{t,k_t^*}^\top \hat{\theta}_{\ell(t)} - X_{t,k_t}^\top \hat{\theta}_{\ell(t)}\right) + \left(X_{t,k_t}^\top \hat{\theta}_{\ell(t)} - X_{t,k_t}^\top\right.\right. \\
&= 2X_{\max} b_{\max}(2^L - 1) + \sum_{m\in[M]} \sum_{\ell \geq L} \sum_{t_\ell \leq t_{\ell+1}} \mathbb{E}\left[\left(X_{t,k_t^*}\theta^* - X_{t,k_t^*}\hat{\theta}_{\ell(t)}\right)\mathbf{1}_{\mathcal{G}} + \left(X_{t,k_t}\hat{\theta}_{\ell(t)} - X_{t,k_t}\theta^*\right)\mathbf{1}_{\mathcal{G}}\right] + X_{\max} \\
&\leq 2X_{\max} b_{\max}(2^L - 1) + 2X_{\max} \sum_{m\in[M]} \sum_{\ell \geq L} \sum_{t_\ell \leq t_{\ell+1}} \mathbb{E}\left[\|\theta^* - \hat{\theta}_{\ell(t)}\|_1 \mathbf{1}_{\mathcal{G}}\right] + X_{\max} b_{\max} \mathbb{P}(\mathbf{1}_{\overline{\mathcal{G}}}) \\
&\leq 2X_{\max} b_{\max}(2^L - 1) + 2X_{\max}\sigma^2\sqrt{s} \sum_{m\in[M]} \sum_{\ell \geq L} (t_{\ell+1} - t_\ell) \sqrt{K' \frac{(s^*)^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{M N_\ell}} \\
&+ 2X_{\max}\sigma^2\sqrt{s} \sum_{m\in[M]} \sum_{\ell \geq L} (t_\ell + 1 - t_\ell) \sqrt{K' \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{M^2 N_\ell^2 \epsilon^2}} \log\left(\frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{D N_\ell}\right) \\
&\leq 2X_{\max} b_{\max} \log T + 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log \log T}{\delta} \sum_{\ell \geq L} \sqrt{K' \cdot M N_\ell} \\
&+ 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log \log T}{\delta} \log\left(\frac{s^2 \log^2 d \log \frac{\log \log T}{\delta}}{D}\right) \sqrt{\frac{K'}{\epsilon^2}} \\
&\leq 2X_{\max} b_{\max} \log T + 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log \log T}{\delta} \sqrt{K' M T} \\
&+ 2X_{\max}\sigma^2 s^{3/2} (\log d) \log \frac{\log \log T}{\delta} \log\left(\frac{s^2 \log^2 d \log \frac{\log \log T}{\delta}}{D}\right) \sqrt{\frac{K'}{\epsilon^2}}
\end{aligned}$$

The per-step regret is given by:

$$\begin{aligned}
\Delta_{k_t}(t) &= x_{k_t^*}^\top \theta^* - x_{k_t}^\top \theta^* \\
&\leq x_{k_t^*}^\top \theta^* - x_{k_t^*}^\top \hat{\theta}_{\ell(t)} + x_{k_t^*}^\top \hat{\theta}_{\ell(t)} - x_{k_t}^\top \hat{\theta}_{\ell(t)} + x_{k_t}^\top \hat{\theta}_{\ell(t)} - x_{k_t}^\top \theta^* \\
&\leq \|x_{k_t^*}\|_\infty \|\theta^* - \hat{\theta}_{\ell(t)}\|_1 + \|x_{k_t}\|_\infty \|\theta^* - \hat{\theta}_{\ell(t)}\|_1 \\
&\leq 4\sigma X_{\max} s \sqrt{MT \log d} + 4\sigma X_{\max} \sqrt{\frac{s^3 (\log d)^2 \log\left(\frac{1}{\delta}\right) \log^6 T)}{\epsilon^2}}
\end{aligned}$$

where, the last-inequality follows from the previous lemma. Consider the event:

$$\mathcal{H}_t \triangleq \{x_{k_t^*}^\top \theta^* > \max_{k \neq k_t^*} x_k^\top \theta^* + p_t\}$$

Under $\mathcal{M}_\ell \cap \mathcal{A}_{\ell(t)}$, we have the following for any $k \neq k_t$:

$$\begin{aligned}
x_{k_t^*}^\top \hat{\theta}_{\ell(t)} - x_k^\top \hat{\theta}_{\ell(t)} &= \langle x_{k_t^*}, \hat{\theta} - \theta^* \rangle + \langle x_{k_t^*} - x_{k_t}, \theta^* \rangle + \langle x_{k_t}, \theta^* - \hat{\theta}_{\ell(t)} \rangle \\
&\geq -g_t + p_{t-1} - g_t
\end{aligned}$$

Choosing $p_t = 3g_{\ell(t)}$, with $g_{\ell(t)} = \sqrt{K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{N_\ell}} + \sqrt{K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{N_\ell^2 \epsilon^2}} \log\left(K' \cdot \frac{s^2 \log^2 d \log \frac{\log N_\ell}{\delta}}{D N_\ell}\right)$, we see that:

1. When $\alpha = 0$, we get the previously proven bound.
2. When $\alpha \in (0, 1)$:

$$\mathbb{E}(R(T)) \leq 2X_{\max}b_{\max}(2^L-1)+24X_{\max}\sigma^2 \sum_{\ell \geq L} \Delta_*^{-\alpha} N_{\ell-1} \left[\left(\frac{s^2 \log d}{N_{\ell-1}} \right)^{\frac{1+\alpha}{2}} + \left(\frac{s^3 (\log d)^2 \log(\frac{1}{\delta}) \log \frac{\log N_{\ell-1}}{N_{\ell-1}^2 \epsilon^2} \right)^{\frac{1+\alpha}{2}} \right]$$

   This implies that:

$$I_\alpha \leq \frac{s^{3(1+\alpha)} (\log d)^{\frac{(1+\alpha)}{2}}}{\Delta_*^\alpha} \left(\frac{(MT)^{\frac{1-\alpha}{2}} - 1}{1-\alpha} \right) + \Psi_\alpha \Delta_*^{-\alpha} \epsilon^{-(1+\alpha)} \left\{ \log(1/\delta) \right\}^{\frac{(1+\alpha)}{2}} \left(s^3 \log^2 d \right)^{\frac{1+\alpha}{2}}$$

   where, $\Psi_\alpha = \frac{1-(MT)^{-\alpha}}{1-2^{-\alpha}} (\log MT)^{2\alpha+2}$.
3. $\alpha = 1$: The previous bounds implies:

$$I_1 \leq \frac{s^6 (\log d)}{\Delta_*} \log T + \frac{\Psi_1}{\Delta_* \epsilon^2} s^3 \log d) \log(\frac{1}{\delta})$$

4. $\alpha > 1$: In this case, we get:

$$I_\alpha \leq \frac{s^{3(1+\alpha)} (\log d)^{\frac{(1+\alpha)}{2}}}{\Delta_*^\alpha} \left(\frac{1 - (MT)^{\frac{-(\alpha-1)}{2}}}{\alpha - 1} \right) + \Psi_\alpha \Delta_*^{-\alpha} \epsilon^{-(1+\alpha)} \left(\log(\frac{1}{\delta}) \right)^{\frac{(1+\alpha)}{2}} \left(s^3 \log^2 d \right)^{\frac{(1+\alpha)}{2}}$$

□

H Proof of Theorem 5

Definition 6 (Collinearly-dependent policies (Huang et al., 2021)). *Two clients iand jare called collinear if there exist an arm $k \in [K]$ and a subset $S \subset [M]$ such that the following conditions are satisfied: 1) $X_{i,k} \in \texttt{span}(\{X_{m,k} | m \in S\})$ and 2) $x_{i,k} \in \texttt{span}(\{X_{m,k} | m \in S\} \cup X_{j,k})$. A collinearly-dependent policy if for any two clients i and j that are not collinear, if the action of client is independent of the action of j under a policy π.*

Let $\mathcal{F}_{X,r}(\theta)$ denote the space of joint distributions of contexts and rewards upto time T, when the underlying parameter is θ. In particular,

$$\mathcal{F}_{X,r} = \otimes_{t \in [T]} \mathcal{F}_{X_t, r_t}$$

where, $\mathcal{F}_{x_t,r_t}$ denotes the joint distribution of contexts $\{x_1(t), x_2(t), \ldots, x_K(t)\} \in \mathbb{R}^{dK}$ and rewards $\{r_1(t), r_2(t), \ldots, r_K(t)\} \in \mathbb{R}^d$ at time t induced through $\{x_1(t), \ldots, x_K(t)\} \sim \mathbb{P}_X$ (where, the marginal for $x_i(t) \sim \mathbb{P}_i$ and the minimum and maximum eigenvalues of the marginal covariance are given by $\lambda_{\min}, \lambda_{\max}$). $r_i(t) = x_i(t)^\top \theta + \epsilon_i(t),\ \epsilon_i(t) \sim \mathbb{P}_\epsilon$ independently for all ϵ, where $\mathbb{P}_\epsilon$ is sub-Gaussian with variance proxy σ^2. For every choice of $\theta \in \mathbb{R}^d$, we get a bandit instance, $\mathbb{P}_{X,r} \in \mathcal{F}_{X,r}(\theta)$.

A policy π is a sequence of (randomized) maps $\pi_t : \mathcal{H}_{t-1}^\pi \times X_t \mapsto \Delta(K)$, where $\mathcal{H}_{t-1} = \{(x_{a_s}, r_{a_s}) : 1 \le s \le t\}$ denotes the history upto time t and $\Delta([K])$ denotes the probability simplex over the set of arms. The minmax regret in the private setting is defined as:

$$R^{\text{minmax}}_{(\epsilon,\delta)}(T) = \inf_{\pi \in \Pi(\epsilon,\delta)} \sup_{P_{X,r} \in \mathcal{F}_{X,r}} \mathbb{E}\left[R(T)\right]$$

where, the expectation is taken wrt randomness due to interaction of the bandit instance $P_{X,r}$ and policy π; the infimum is over all possible history-adapted (ϵ, δ)-JDP policies and the supremum is wrt all problem instances with sparsity s.

1. **Constructing the hard instance:**
   We construct several hard instances as follows: consider the case of 2-armed problem with $\theta \in \tilde{\Theta} = \{\theta \in \mathbb{R}^d : \theta_i \in \{-r_{\min}, 0, r_{\max}\}, \|\theta\|_0 = s\}$, where s is the true sparsity parameter. Then, $|\tilde{\Theta}| = \binom{d}{s} 2^s$. Let

   $$\delta(t) = \{\delta : \|\theta - \theta'\|_2 \ge \delta', \theta, \theta' \in \tilde{\Theta},\ d_H(\theta, \theta') \ge t\}$$

   where, d_H is the Hamming distance. Considering a t-packing of $\tilde{\Theta}$ in the Hamming distance, this leads to $\delta(t)$ packing of $\tilde{\Theta}$ in the ℓ_2-distance Duchi and Wainwright (2013). With our construction, this leads to $\delta(t) > \max\{1, \sqrt{t}\}r$. Let $\Theta^* = \{\theta_1, \theta_2, \ldots, \theta_M\}$ be the elements of this packing with $t >$ and $M > cs \log \frac{d}{s}$ for some absolute constant c. For $\theta, \theta' \in \tilde{\Theta},\ \|\theta - \theta'\| \ge \alpha$.
2. **Bounds between the reward-context distribution:**
   For $\theta, \theta' \in \tilde{\Theta}$, we bound the KL-divergence between the context and reward distribution, which is bounded as:

   $$\begin{aligned} \text{KL}\left(P(\cdot|\theta) \| P(\cdot|\theta')\right) &= MKt\text{KL}\left(\mathcal{N}(x^\top\theta, \sigma^2), \mathcal{N}(x^\top\theta', \sigma^2)\right) \\ &= \frac{MKt\sigma_x^2 \|\theta - \theta'\|}{2\sigma^2} \\ &\le \frac{MKt\sigma_X^2 sr^2}{\sigma^2} \end{aligned}$$

   Using, Pinsker's inequality, the Total variation distance can be bounded as:

   $$\begin{aligned} \text{TV}\left(P(\cdot|\theta) \| P(\cdot|\theta')\right) &\le MKt\sqrt{\frac{1}{2}\mathbb{E}\left[\text{KL}\left(\mathcal{N}(x^\top\theta, \sigma^2), \mathcal{N}(x^\top\theta', \sigma^2)\right)\right]} \\ &\le MKt\frac{\sigma^2}{\sigma_x^2} r\sqrt{s} \end{aligned}$$

3. **Reduction to Estimation Problem**
Given a policy π and its associated estimator $\hat{\theta}_t$ be the maximizer of

$$\mathbb{P}(k_t^\pi = 1, z_t^\top \theta \geq 0 | H_{t-1}^\pi) + \mathbb{P}(k_t^\pi = 2, z_t^\top \theta \leq 0 | H_{t-1}^\pi)$$

where, the max is over all possible random maps from history induced by π which are (ϵ, δ)-DP. Let $\hat{\Theta}_t^{\epsilon,\delta}$ denote this space of random (ϵ, δ)-DP maps and π be an (ϵ, δ)-JDP policy. Therefore, from Theorem **??**, we can conclude that:

$$\inf_{\pi \in \Pi} \sup_{P_{X,r} \in \mathcal{F}} \mathbb{E}\left[R^\pi(T)\right] \geq \inf_{\pi \in \Pi} \sup_{P_{X,r} \in \mathcal{F}-construct} \mathbb{E}\left[R^\pi(T)\right] = \Omega\left(\frac{1}{r\sqrt{s}} \sum_{t=1}^{T} \inf_{\hat{\theta}_t \in \hat{\Theta}_t^{\epsilon,\delta}} \mathbb{E}_\nu\left[\|\theta - \hat{\theta}_t\|_2^2\right]\right)$$

4. We divide this step into two parts, where for the ϵ-DP case we use Fano's inequality (Acharya et al., 2021) and the connection between ρ-zCDP and (ϵ, δ)-DP (Bun et al., 2018) and ρ-zCDP Fano's inequality (Kamath et al., 2022) to get the lower bound for any (ϵ, δ)-JDP.
   (a) Analysis under ϵ-DP. For the non-private lower bound, we have:

$$\mathbb{E}\left[\|\theta_t - \theta^\star\|_2^2\right] \quad \geq \quad \frac{\max\{1, s/4\} r^2}{2}\left(1 - \frac{\frac{2MKt\sigma_x^2 sr^2}{\sigma^2} + \log 2}{cs\log(d/s)}\right)$$

   For the private lower-bound, we have:

$$\mathbb{E}\left[\|\theta_t - \theta^\star\|_2^2\right] \quad \geq 0.4 \max\{1, s/4\} r^2 \min\left\{1, \exp\left(cs\log(d/s) - 10\epsilon t MK \tfrac{\sigma_x}{\sigma} r\sqrt{s}\right)\right\}$$

   For $s > 4$ and $\epsilon > 0$, we have:

$$\mathbb{E}\left[\|\theta_t - \theta^\star\|_2^2\right] \geq \frac{sr^2}{4} \max\left\{0.5\left(1 - \frac{2t\frac{\sigma_x^2}{\sigma^2} sr^2 MK}{cs\log(d/s)}\right), 0.4\exp\left(cs\log(d/s) - 10\epsilon t\sqrt{s} r \frac{\sigma_x}{\sigma} MK\right)\right\}$$

   Hence, for $K = 2$, the minimax-regret is given by:

$$\inf_{\pi \in \Pi} \sup_{P_{X,r} \in \mathcal{F}} \mathbb{E}[R(T)] \geq \frac{r\sqrt{s}MT}{4} \max\left\{1 - \frac{\frac{4MTsr^2\sigma_x^2}{\sigma^2} + \log 2}{cs\log(d/s)}, \exp\left(cs\log(d/s) - 20\epsilon MT\frac{\sigma_x}{\sigma} r\sqrt{s}\right), \right\}$$

   Setting $r^2 = \frac{\sigma^2 \log(d/s)}{8MT}$, the non-private lower bound becomes:

$$\inf_{\pi \in \Pi} \sup_{P_{X,r} \in \mathcal{F}} \mathbb{E}[R(T)] \geq \frac{1}{8\sqrt{2}} \sigma \sqrt{MTs\log(d/s)}\left(1 - \frac{\sigma_x^2}{2c} - \frac{\log 2}{cs\log(d/s)}\right)$$

   Noting that $\sigma_x^2 = (\log d)^{-1}, d > 9$ and $s\log(d/s) > 4c^{-1}\log 2$, our non-private lower bound is $\Omega\left(\sqrt{sMT\log(d/s)}\right)$.

   For the private part, we set $r = \frac{c\sigma\sqrt{s}\log(d/s)}{20\epsilon\sigma_x MT}$, to get:

$$\inf_{\pi \in \Pi} \sup_{P_{X,r} \in \mathcal{F}} \mathbb{E}[R(T)] \geq \sigma \frac{s\log(d/s)\log^{1/2}(d)}{\epsilon} \geq \sigma \frac{\log^{3/2}(d/s)}{\epsilon}$$

   Combining these results, we get:

$$R_\epsilon^{\text{minimax}}(T) = \Omega\left(\max\{s\log^{3/2}(d/s)\epsilon^{-1}, \sqrt{sMT\log(d/s)}\right)$$

   (b) For (ϵ, δ)-JDP, we construct $\rho(\epsilon, \delta)$ such that any (ϵ, δ)-DP algorithm is ρ-zCDP and thus $\Theta_t^{\epsilon,\delta} \subset \Theta_t^\rho$ where the latter is the space of all ρ-zCDP estimators constructed using the entire history H_{t-1}. Other arguments are as before.

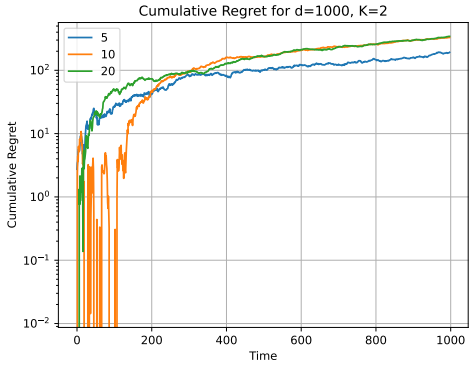

Figure 1: Regret in central setting. The synthetic instances are constructed as in Ariu et al. (2020) and the sparsity parameter is varied. We see that the regret behaves well across different values of sparsity.

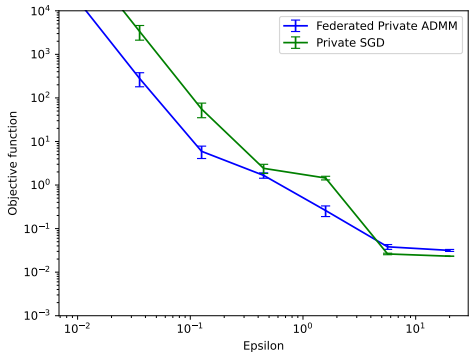

Figure 2: Decrement in estimation error for distributed LASSO.

I Experimental Evaluation

The experiments are done on a MacBook Air with an Apple M1 chip, 16 GB memory and 10 core CPU. All codes are written in Python3 using several open source packages. The running time for all experiments ranges from less than a minute to a few hours. The code is available at: this link.

I.1 Simulations with non-bandit feedback

We illustrate the performance of Algorithm 3 on a distributed LASSO problem, i.e., solving:

$$\min_x \frac{1}{2n}\|Ax - b\|^2 + \kappa\|x\|_1$$

where, $A \in \mathbb{R}^{n\times p}$ and $b \in \mathbb{R}^p$ is a vector of regression targets. We generate synthetic data by drawing A as random vectors from p-dimensional sphere and x from uniform distribution with support size 10. Labels are then obtained by taking $b = Ax + \eta$ where $\eta \sim \mathcal{N}(0, 0.01)$ with $n = 500$ and $p = 32$.

The benchmark comparison is performed with `scikit-learn` and we obtain the best parameter using cross-validation. We also compare our algorithm with ADMM with soft-thresholding (see (Cyffers et al., 2023)) and DP-SGD where noise is added to the gradients. For both approaches, we tune step-size and clipping threshold using grid search. These parameters are tuned on the smallest privacy budget and used for all other values.

We report the objective function value on test set at the end of training and convert them to (ϵ, δ)-DP for the sake of comparison with $\delta = 1e-6$. Each user is randomly sampled with probability 10%.

J Auxiliary Results

In this appendix, we provide a collection of existing results used in our proofs.

Theorem 6 (Fixed-point iteration Combettes and Pesquet (2021))**.** *Assume that R is a τ-contractive operator with fixed point u_* for $\tau \in [0, 1)$. Then there exists a learning rate $\lambda_b \in (0, 1]$ such that the iterates of Algorithm 1 satisfy:*

$$\mathbb{E}\left[\|u_{k+1} - u^*\|^2 | \mathcal{F}_0\right] \leq \left(1 - \frac{(1-\tau)}{8}\right)^B D$$

where $D = \max \|u_0 - u^*\|$ is the diameter of the domain, p is the dimension of $u, \sigma^2 > 1 - \tau$ is the variance of variance of gaussian noise and $\mathbb{E}[\|e_k\|^2] \leq \xi^2$ for some $\xi \geq 0$.

Theorem 7 (Laplace concentration Cai et al. (2020))**.** *Consider $w \in \mathbb{R}^k$ with $w_1, w_2, \ldots, w \overset{iid}{\sim}$ Laplace(λ). For every $C > 1$, we have:*

$$\begin{aligned} \mathbb{P}(\|w\|_2^2 > kC^2\lambda^2) &\leq k \exp(-C) \\ \mathbb{P}(\|w\|_\infty > C^2\lambda^2 \log^2 k) &\leq \exp(-(C-1)\log k) \end{aligned}$$