Checklist

- 1. For all authors...
 - (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
 - (b) Did you describe the limitations of your work? [Yes] See Section 6.
 - (c) Did you discuss any potential negative societal impacts of your work? [No]
 - (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
- 2. If you are including theoretical results...
 - (a) Did you state the full set of assumptions of all theoretical results? [Yes]
 - (b) Did you include complete proofs of all theoretical results? [Yes]
- 3. If you ran experiments...
 - (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] See the supplemental material.
 - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See Section 5 and Appendix F
 - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] We perform 10 replications for each case and plot the mean and 0.5 standard deviation of their regrets. See Section 5
 - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No]
- 4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
 - (a) If your work uses existing assets, did you cite the creators? [Yes] We use the real data On-Line Auto Lending dataset CRPM-12-001 provided by Columbia University and we add their link in Section 5
 - (b) Did you mention the license of the assets? [No]
 - (c) Did you include any new assets either in the supplemental material or as a URL? [No]
 - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [No]
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [No]
- 5. If you used crowdsourcing or conducted research with human subjects...
 - (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
 - (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
 - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Randomness Condition

In this section, we show that a sub-gaussian random vector with bounded density satisfies Assumption 3:

We say a random vector x is σ^2 -sub-gaussian vector with bounded density, if for every $v \in S_1^{d-1}$, $v^T x$ is σ^2 -sub-gaussian and its density function exists and is bounded by γ for some $\gamma > 0$. For such kind of random vector, [30] shows that it satisfies Assumption 3 with $\kappa_l = \frac{2d}{3\gamma K}$ and $p_* = \frac{1}{3}$. In particular, [22] shows that when x follows $\mathcal{N}(0, \Sigma)$, with $\lambda_{\min}(\Sigma) \geq \frac{\kappa}{d}$, we can have $\kappa_l = \frac{c_1 \kappa}{d}$ and $p_* = c_2$ for constants c_1 and c_2 .

B Proof of Privacy Guarantee

B.1 Proof of Results in Section 3.1

Proof of Proposition 3.1. Since we assume that the features and rewards are bounded, $||x_{t,a}|| \leq C_B$, $||r_t|| \leq c_r$ for all $t \in [T]$ and $a \in [K]$, by Lemma 2.1, M_t is $(\varepsilon/2, \delta/2)$ -LDP and u_t is $(\varepsilon/2, \delta/2)$ -LDP. Thus Lemma 2.4 implies that ψ_t^{OLS} is (ε, δ) -LDP.

Proof of Proposition 3.2. Since we assume that the features and rewards are bounded, $||x_{t,a}|| \leq C_B$, $||r_t|| \leq c_r$ for all $t \in [T]$ and $a \in [K]$, we have $(\mu(x_{t,a_t}^T \hat{\theta}_{t-1}) - r_t)x_{t,a_t}$ bounded by $2c_r C_B$. Lemma 2.2 implies that ψ_t^{SGD} is ε -LDP.

B.2 Proof of Results in Section 4.1

Proof of Proposition 4.1. We simply denote ψ_t^{OLS} by ψ_t in this proof. At time t, for any two $x \neq x'$, without loss of generality assuming the action corresponding x and x' are $a_t = 1$ and $a_t = 2$, then the output corresponding x, x' is given by $(\psi_t(x, x^T\theta_1 + \epsilon_t), \psi_t(0, 0), \ldots, \psi_t(0, 0))$ and $(\psi_t(0, 0), \psi_t(x', x'^T\theta_2 + \epsilon_t), \ldots, \psi_t(0, 0))$. Since $\psi_t(0, 0)$ has the same distribution, we have for any subset $A_1 \times A_2 \times \cdots \times A_K \subset \mathbb{R}^{Kd}$ with A_i a Borel set in \mathbb{R}^d ,

$$\frac{\mathbb{P}(\psi_t(x, x^T \theta_1 + \epsilon_t) \in A_1, \psi_t(0, 0) \in A_2, \dots, \psi_t(0, 0) \in A_K)}{\mathbb{P}(\psi_t(0, 0) \in A_1, \psi_t(x', x'^T \theta_2 + \epsilon_t) \in A_2, \dots, \psi_t(0, 0) \in A_K)} = \frac{\mathbb{P}(\psi_t(x, x^T \theta_1 + \epsilon_t) \in A_1, \psi_t(0, 0) \in A_2)}{\mathbb{P}(\psi_t(x', x'^T \theta_2 + \epsilon_t) \in A_2, \psi_t(0, 0) \in A_1)}.$$
(8)

Set $\tilde{\psi}(v_1, v_2) := (\psi_t(v_1), \psi_t(v_2))$, and $(v_1, v_2) := (x, 0), (v'_1, v'_2) := (0, x')$, then we have (8) equals to $\tilde{\psi}(v_1, v_2)/\tilde{\psi}(v'_1, v'_2)$, thus applying Lemma 2.4 to it implies that (8) is upper bounded by $e^{\varepsilon} + \delta \mathbb{P}(\psi_t(x', x'^T \theta_2 + \epsilon_t) \in A_2, \psi_t(0, 0) \in A_1)^{-1}$, leading to the desired result. \Box

Proof of Proposition 4.2. That is nearly the same as the proof of Proposition 4.1, but replacing $e^{\varepsilon} + \delta \mathbb{P}(\psi_t(x', x'^T \theta_2 + \epsilon_t) \in A_2, \psi_t(0, 0) \in A_1)^{-1}$ by e^{ε} in the last step. \Box

C Proof of Results in Section 3.2

In the following analysis, without special explaination, all the c and C denote absolute constants. Sometimes we state the inequality of type $A_1 \leq C \log(A_2/\alpha)A_3$ holds with probability at least $1 - \alpha$ while in proof we derive the results hold with $1 - c\alpha$ for some constant c. In fact, they are equivalent by re-scaling α and changing C to some larger constant.

C.1 Proof of Worst-Case Bounds

Proof of Theorem 3.1. Since x_{t,a_t} is the greedy selection, we have $x_{t,a_t}^T \hat{\theta}_{t-1} \ge x_{t,a}^T \hat{\theta}_{t-1}$ for any time $t \in [T]$ and $a \in [K]$. Consequently we have the following upper bound for the instantaneous

regret at time t,

$$\max_{a \in [K]} (x_{t,a} - x_{t,a_t})^T \theta^* \leq \max_{a \in [K]} (x_{t,a} - x_{t,a_t})^T \left(\theta^* - \hat{\theta}_{t-1}\right)$$
$$\leq \max_{a,a' \in [K]} (x_{t,a} - x_{t,a'})^T \left(\theta^* - \hat{\theta}_{t-1}\right)$$
$$\leq 2 \max_{a \in [K]} \left| x_{t,a}^T \left(\theta^* - \hat{\theta}_{t-1}\right) \right|.$$

For any fixed $a \in [K]$, $x_{t,a}$ is independent of $\hat{\theta}_{t-1}$. By Assumption 3, conditioning on the historical information up to time t, $x_{t,a}^T(\theta^* - \hat{\theta}_{t-1})$ is a $\frac{\kappa_u}{d} \|\theta^* - \hat{\theta}_{t-1}\|^2$ -sub-gaussian random variable. Now by the maximal concentration inequality for a sub-gaussian sequence, we have with probability at least $1 - \frac{\alpha}{T}$,

$$\max_{a \in [K]} |x_{t,a}^T(\theta^* - \hat{\theta}_{t-1})| = O\left(\sqrt{\frac{\kappa_u \log(KT/\alpha)}{d}} \|\theta^* - \hat{\theta}_{t-1}\|\right).$$

To control the regret bound, we bound the estimation error $\|\theta^{\star} - \hat{\theta}_{t-1}\|$ in each time in the following lemma.

Lemma C.1 (Estimation Error for OLS). Using the private OLS update mechanism ψ_t^{OLS} and estimator φ_t^{OLS} , for any $8\frac{d\log 9 + \log(T/\alpha)}{p_*^2} < t \leq T$, we have with probability at least $1 - \frac{\alpha}{T}$,

$$\|\hat{\theta}_t - \theta^\star\|^2 \le C(C_B \sigma_\epsilon \sigma_{\epsilon,\delta} d)^2 \frac{d + \log(T/\alpha)}{\kappa_l^2 p_*^2 t},\tag{9}$$

for some C independent of d, K and T.

m

Lemma C.2 (Estimation Error for SGD). Using the private OLS update mechanism ψ_t^{SGD} and estimator φ_t^{SGD} , for any $3 \le t \le T$, we have with probability at least $1 - \frac{\alpha}{T}$,

$$\|\hat{\theta}_t - \theta^\star\|^2 \le \frac{(624\log(\log T/\alpha) + 1)r_{\varepsilon,d}^2 d^2}{4\kappa_l^2 \zeta^2 p_*^2 t}.$$
(10)

Plugging OLS estimation error (9) into the regret bound, denote $t_1 := 8 \frac{d \log 9 + \log(T/\alpha)}{p_*^2}$, the following holds with probability at least $1 - \alpha$,

$$\sum_{t=1}^{T} \max_{a \in [K]} (x_{t,a} - x_{t,a_t})^T \theta^*$$

$$\leq t_1 c_r + \sum_{t=t_1+1}^{T} CC_B \sigma_\epsilon \sigma_{\varepsilon,\delta} d\sqrt{\frac{\kappa_u \log(KT/\alpha)}{d}} \frac{\sqrt{d + \log(T/\alpha)}}{\kappa_l p_* \sqrt{t}}$$

$$\leq 8 \frac{d \log 9 + \log(T/\alpha)}{p_*^2} + CC_B \sigma_{\varepsilon,\delta} \sigma_\epsilon \sqrt{d} \frac{\sqrt{d + \log(T/\alpha)}}{\kappa_l p_*} \sqrt{\kappa_u \log(KT/\alpha)} \sqrt{T}.$$
(11)

Plugging the SGD estimation error (10) into the regret bound, we have

$$\sum_{t=1}^{T} \max_{a \in [K]} (x_{t,a} - x_{t,a_t})^T \theta^*$$

$$\leq 2c_r + \sum_{t=3}^{T} \sqrt{\kappa_u \log(KT/\alpha)} \frac{\sqrt{(624 \log(\log T/\alpha) + 1)} r_{\varepsilon,d} \sqrt{d}}{2\kappa_l \zeta p_* \sqrt{t}}$$

$$\leq 2c_r + \frac{\sqrt{(624 \log(\log T/\alpha) + 1)} r_{\varepsilon,d} \sqrt{d}}{2\zeta \kappa_l p_*} \sqrt{\kappa_u \log(KT/\alpha)} \sqrt{T}.$$
(12)

So now it suffices to prove the Lemmas C.1 and C.2.

C.2 Proof of lemma C.1

Lemma C.3. As long as $t > 8 \frac{d \log 9 + \log(T/\alpha)}{p_*^2}$, the following lower bound

$$\lambda_{\min}\left(\sum_{i=1}^{t} x_{i,a_i} x_{i,a_i}^T\right) \ge C \cdot \frac{t\kappa_l p_*}{d},$$

holds with probability at least $1 - \frac{\alpha}{T}$, for some C independent of d and T.

Proof. Define \mathcal{F}_t^- as the filtration generated by $\{x_{i,a_i}\}_{i \in [t-1]}$, $\{\epsilon_i\}_{i \in [t-1]}$ and the randomness from $\{\psi_i^{OLS}\}_{i \in [t-1]}$. By greedy algorithm, in each time i, x_{i,a_i} is selected as $a_i = \operatorname{argmax}_{a \in [K]} x_{i,a}^T \hat{\theta}_{i-1}$. Thus by the Assumption 3, we have for any $0 < s < p_*$,

$$\begin{split} & \mathbb{P}(\sum_{i=1}^{t} (x_{i,a_{i}}^{T}v)^{2} < t\kappa_{l}(p_{*}-s)/d) \\ & \leq \mathbb{P}(\sum_{i=1}^{t} \mathbf{1}\{(x_{i,a_{i}}^{T}v)^{2} > \kappa_{l}/d\} < t(p_{*}-s)) \\ & \leq \mathbb{P}(\frac{1}{t}\sum_{i=1}^{t} (\mathbf{1}\{(x_{i,a_{i}}^{T}v)^{2} > \kappa_{l}/d\} - \mathbb{E}[\mathbf{1}\{(x_{i,a_{i}}^{T}v)^{2} > \kappa_{l}/d\} | \mathcal{F}_{i}^{-}])) < -s) \\ & \leq \exp(-\frac{s^{2}t}{2}), \end{split}$$

where in the last inequality we use the Azuma–Hoeffding's inequality for bounded martingaledifference sequence (see Corollary 2.20 in [42]).

For every $d \times d$ positive-definite matrix A, with an abuse of notation, we denote $\mathcal{N}_{\varepsilon}$ as the ε -net of S_1^{d-1} for some $\varepsilon > 0$ to be determined,

$$\lambda_{\max}(A) \le \frac{1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}_{\varepsilon}} x^T A x,$$

which then implies

$$\lambda_{\min}(A) = -\lambda_{\max}(-A) \ge \frac{-1}{1 - 2\varepsilon} \sup_{x \in \mathcal{N}_{\varepsilon}} x^T (-A) x = \frac{1}{1 - 2\varepsilon} \inf_{x \in \mathcal{N}_{\varepsilon}} x^T A x.$$

By choosing $\varepsilon = 1/4$, we can find an ε -net $\mathcal{N}_{\varepsilon}$ with cardinality $|\mathcal{N}_{\varepsilon}| \leq 9^d$. Therefore

$$\lambda_{\min}(A) \ge 2 \inf_{x \in \mathcal{N}_{\varepsilon}} x^T A x.$$

Note that

$$\mathbb{P}(\min_{\|v\|=1}\sum_{i=1}^{t} (x_{i,a_{i}}^{T}v)^{2} < 2t\kappa_{l}(p_{*}-s)/d) \leq \mathbb{P}(\sum_{i=1}^{t} (x_{i,a_{i}}^{T}v)^{2} < t\kappa_{l}(p_{*}-s)/d, \exists v \in \mathcal{N}_{\varepsilon})$$
$$\leq 9^{d} \exp(-\frac{s^{2}t}{2}).$$

By setting $s = \sqrt{\frac{2d \log 9 + 2 \log(T/\alpha)}{t}}$, we have when $t > 8 \frac{d \log 9 + \log(T/\alpha)}{p_*^2}$ with probability at least $1 - \frac{\alpha}{T}$,

$$\lambda_{\min}(\sum_{i=1}^{s} x_{i,a_i} x_{i,a_i}^T) = \min_{\|v\|=1} \sum_{i=1}^{s} \langle x_{i,a_i}, v \rangle^2 \ge \frac{\kappa_l p_* t}{d}.$$

Proof of Lemma C.1.

By lemma C.3 we know that with probability at least $1 - \frac{\alpha}{T}$,

$$\lambda_{\min}(\sum_{i=1}^t x_{i,a_i} x_{i,a_i}^T) \ge C_1 \kappa_l p_* t/d,$$

for some C_1 independent of d, K and T.

Since $\{W_i\}_{i \in [t]}$ are independent, therefore by concentration bounds for Wigner matrix we have with probability at least $1 - \frac{\alpha}{T}$,

$$\left\|\sum_{i=1}^{t} W_i\right\|^2 \le C_2 t \sigma_{\varepsilon,\delta}^2 (d + \log(T/\alpha)),$$

for some C_2 independent of d, K and T. However, it is important to note that the perturbation of privacy noise matrix $\sum_{i=1}^{t} W_i$ may destroy the positive definite property of the Gram matrix $\sum_{i=1}^{t} x_{i,a_i} x_{i,a_i}^T$ when t is still small. Therefore, we shift $\sum_{i=1}^{t} W_i$ by adding $\tilde{c}\sqrt{t}I_d$ where $\tilde{c} := C_2 \sigma_{\varepsilon,\delta}(\sqrt{d} + \sqrt{\log(T/\alpha)})$.

We denote $A_t \coloneqq \sum_{i=1}^t (x_{i,a_i} x_{i,a_i}^T + W_i) + \tilde{c}\sqrt{t}I$. Therefore, by Weyl's inequality we have with probability at least $1 - \frac{\alpha}{T}$,

$$\lambda_{\min}(A_t) = \lambda_{\min}\left(\sum_{i=1}^t (x_{i,at_i} x_{i,a_i}^T + W_i) + \tilde{c}\sqrt{t}I_d\right) \ge \lambda_{\min}\left(\sum_{i=1}^t x_{i,a_i} x_{i,a_i}^T\right) \ge C_1 \kappa_l p_* t/dt$$

So now we study the OLS estimator with x_{i,a_i} , ϵ_i given above and $r_i = x_{i,a_i}^T \theta^* + \epsilon_i$. In that case, the estimation error of the OLS estimator under LDP constraints at time t is given by

$$\hat{\theta}_{t} - \theta^{\star} = A_{t}^{-1} \sum_{i=1}^{t} (x_{i,a_{i}}r_{i} + \xi_{i}) - \theta^{\star}$$

$$= A_{t}^{-1} \sum_{i=1}^{t} (x_{i,a_{i}}x_{i,a_{i}}^{T}\theta^{\star} + x_{i,a_{i}}\epsilon_{i} + \xi_{i}) - \theta^{\star}$$

$$= A_{t}^{-1} (\sum_{i=1}^{t} x_{i,a_{i}}\epsilon_{i}) - A_{t}^{-1} \sum_{i=1}^{t} W_{i}\theta^{\star} + A_{t}^{-1} \sum_{i=1}^{t} \xi_{i} - \tilde{c}\sqrt{t}A_{t}^{-1}\theta^{\star}.$$

Define \mathcal{F}_t as the filtration generated by $\{x_{i,a_i}\}_{i \in [t]}, \{\epsilon_i\}_{i \in [t-1]}$ and the randomness from $\{\psi_i\}_{i \in [t-1]}$. Notice that for every unit vector u,

$$\mathbb{E}[\exp(\lambda \sum_{i=1}^{t} u^T x_{i,a_i} \epsilon_i)] = \mathbb{E}[\mathbb{E}[\exp(\lambda \sum_{i=1}^{t} u^T x_{i,a_i} \epsilon_i) |\mathcal{F}_t]]$$
$$= \mathbb{E}[\prod_{i=1}^{t-1} \exp(\lambda u^T x_{i,a_i} \epsilon_i) \mathbb{E}[\exp(\lambda u^T X_i \epsilon_i) |\mathcal{F}_t]]$$
$$\stackrel{(1)}{\leq} \exp(\frac{\lambda^2 C_B^2 \sigma_\epsilon^2}{2}) \mathbb{E}[\prod_{i=1}^{t-1} \exp(\lambda u^T x_{i,a_i} \epsilon_i)]$$
$$\stackrel{(2)}{\leq} \exp(\frac{\lambda^2 C_B^2 \sigma_\epsilon^2 t}{2}).$$

Inequality (2) is due to the mathematical induction using the same technique in the equality (1). Thus $\sum_{i=1}^{t} x_{i,a_i} \epsilon_i$ is $\sigma^2 C_B^2 t$ -sub-gaussian vector, and by the concentration of norm for sub-gaussian vectors, we have then with probability at least $1 - \frac{\alpha}{T}$,

$$\|\sum_{i=1}^{l} x_{i,a_i} \epsilon_i\|^2 \le C_3 \sigma_{\epsilon}^2 C_B^2 t (d + \log(T/\alpha)),$$

where C_3 is a positive constant independent of d, K and T. Therefore,

$$\|A_t^{-1}(\sum_{i=1}^t x_{i,a_i}\epsilon_i)\|^2 \le \|A_t^{-1}\|^2 \|(\sum_{i=1}^t x_{i,a_i}\epsilon_i)\|^2 \le \frac{C_3\sigma^2 C_B^2 d^2 t (d+\log(T/\alpha))}{(C_1\kappa_l p_* t)^2}.$$
(13)

Moreover,

$$\|A_t^{-1}\sum_{i=1}^t W_i\theta^\star\|^2 \le \|A_t^{-1}\|^2 \|\sum_{i=1}^t W_i\|^2 \|\theta^\star\|^2 \le \|A_t^{-1}\|^2 \|\sum_{i=1}^t W_i\|^2 \le \frac{C_2 t \sigma_{\varepsilon,\delta}^2 (d + \log(T/\alpha))}{(C_1 \kappa_l p_* t)^2},$$
(14)

where the second inequality is from the assumption that $\|\theta^{\star}\| \leq 1$.

Third, Since ξ_i are random vector with independent, sub-gaussian coordinates that satisfy $\mathbb{E}\xi_{i,j}^2 = \sigma_{\varepsilon,\delta}^2$, $\sum_{i=1}^t \xi_i$ is a random vactor with independent sub-gaussian coordinates that satisfy $\mathbb{E}\sum_{i=1}^t \xi_{i,j}^2 = t\sigma_{\varepsilon,\delta}^2$. Therefore for all $t \in [T]$, with probability at least $1 - \frac{\alpha}{T}$,

$$\|\sum_{i=1}^{t} \xi_i\|^2 \le C_4 t \sigma_{\varepsilon,\delta}^2 (d + \log(T/\alpha)),$$

for some positive constant C_4 independent of d, K and T. Therefore,

$$\|A_t^{-1}\sum_{i=1}^t \xi_i\|^2 \le \frac{C_4 t \sigma_{\varepsilon,\delta}^2 d^2 (d + \log(T/\alpha))}{(C_2 \kappa_l p_* t)^2}.$$
(15)

Lastly,

$$\|\tilde{c}\sqrt{t}A_t^{-1}\theta^\star\|^2 \le \frac{\tilde{c}^2 t}{(C_2\kappa_l p_* t)^2},\tag{16}$$

holds with probability at least $1 - \frac{\alpha}{T}$. Plugging all bounds (13) (14) (15) and (16) together we get then with probability at least $1 - \frac{\alpha}{T}$,

$$\|\hat{\theta}_t - \theta^\star\|^2 \le C_5 \sigma_\epsilon^2 C_B^2 \sigma_{\varepsilon,\delta}^2 d^2 \frac{d + \log(T/\alpha)}{\kappa_l^2 p_*^2 t},$$

for some positive constant C_5 independent of d, K and T.

Proof of Lemma C.2. Denote g_t as the gradient at time t, $\hat{g}_t \coloneqq \Psi_{\varepsilon}[(\mu(x_{t,a_t}^T \hat{\theta}_t) - r_t)x_{t,a_t}]$ is the LDP private estimator of g_t and $\hat{z}_t = g_t - \hat{g}_t$. By the unbiasedness of Ψ_{ε} in Lemma 2.2 we have

$$\mathbb{E}[\Psi_{\varepsilon}((\mu(x_{t,a_{t}}^{T}\hat{\theta}_{t-1}) - r_{t})x_{t,a_{t}})^{T}(\hat{\theta}_{t-1} - \theta^{\star})|\mathcal{F}_{t-1}] \\ = \mathbb{E}[(\mu(x_{t,a_{t}}^{T}\hat{\theta}_{t-1}) - \mu(x_{t,a_{t}}^{T}\theta^{\star}))x_{t,a_{t}}^{T}(\hat{\theta}_{t-1} - \theta^{\star})|\mathcal{F}_{t-1}] \\ \geq \zeta \mathbb{E}[[x_{t,a_{t}}^{T}(\hat{\theta}_{t-1} - \theta^{\star})]^{2}|\mathcal{F}_{t-1}] \geq \zeta \kappa_{l} p_{\star}/d\|\hat{\theta}_{t-1} - \theta^{\star}\|^{2},$$

where the last inequality is from Lemma C.3 and Markov's inequality $\lambda_{\min}(\mathbb{E}_{x_{a_t}}[x_{a_t}x_{a_t}^T|\mathcal{F}_{t-1}]) \geq \kappa_l p_*/d$. Moreover, notice that $\|\hat{g}_t\| = r_{\varepsilon,\delta}$. Let $\lambda \coloneqq 2\kappa_l \zeta p_*/d$ and $\eta_t = \frac{1}{\lambda t}$,

$$\begin{split} \|\hat{\theta}_{t} - \theta^{\star}\|^{2} &= \|\hat{\theta}_{t-1} - \eta_{t}\hat{g}_{t} - \theta^{\star}\|^{2} \\ &= \|\hat{\theta}_{t-1} - \theta^{\star}\|^{2} - 2\eta_{t}\hat{g}_{t}^{T}(\hat{\theta}_{t-1} - \theta^{\star}) + \eta_{t}^{2}\|\hat{g}_{t}\|^{2} \\ &= \|\hat{\theta}_{t-1} - \theta^{\star}\|^{2} - 2\eta_{t}g_{t}^{T}(\hat{\theta}_{t-1} - \theta^{\star}) + 2\eta_{t}\hat{z}_{t}^{T}(\hat{\theta}_{t-1} - \theta^{\star}) + \eta_{t}^{2}\|\hat{g}_{t}\|^{2} \\ &\leq (1 - 2\lambda\eta_{t})\|\hat{\theta}_{t-1} - \theta^{\star}\|^{2} + 2\eta_{t}\hat{z}_{t}^{T}(\hat{\theta}_{t-1} - \theta^{\star}) + \eta_{t}^{2}\|\hat{g}_{t}\|^{2} \\ &\leq \left(1 - \frac{2}{t}\right)\|\hat{\theta}_{t-1} - \theta^{\star}\|^{2} + \frac{2}{\lambda t}\hat{z}_{t}(\hat{\theta}_{t-1} - \theta^{\star}) + \left(\frac{r_{\varepsilon,d}}{\lambda t}\right)^{2}. \end{split}$$

It follows from the same proof as in Proposition 1 in [28], we can obtain for any $0 < \alpha \leq \frac{1}{eT}$, $T \geq 4$ and for all $3 \leq t \leq T$, with probability at least $1 - \alpha$,

$$\|\hat{\theta}_t - \theta^{\star}\|^2 \le \frac{(624 \log(\log(T)/\alpha) + 1)r_{\varepsilon,d}^2 d^2}{4\kappa_l^2 \zeta^2 p_*^2 t}.$$

C.3 Proof of Problem-dependent Bound

To prove the problem-dependent bound, we need only combine Lemma C.1 and Lemma C.2 together with the following lemma.

Lemma C.4. Under the (β, γ) -margin condition, if we have $\|\hat{\theta}_t - \theta^*\| \leq \frac{U_0}{\sqrt{t}}$ holds uniformly for all $t_0 \leq t \leq T_0$ for some t_0 and U_0 with probability at least $1 - \alpha$, we have then with probability at least $1 - 2\alpha$,

$$\operatorname{Reg}(T) \le C \cdot \begin{cases} c_r t_0 + \gamma (LC_B U_0)^2 (\log T + o(1)), & \beta = 1\\ c_r t_0 + \frac{2\gamma}{1-\beta} (LC_B U_0)^{1+\beta} (T^{\frac{1-\beta}{2}} + o(1)), & 0 \le \beta < 1. \end{cases}$$

Proof. We have, with probability at least $1 - \alpha$,

$$\operatorname{Reg}(T) \leq 2c_r t_0 + (\mu(x_{t,a_t}^T \theta^*) - \mu(x_{t,a_t}^T \theta^*)) \mathbf{1}\{\|\hat{\theta}_t - \theta^*\| \leq \frac{U_0}{\sqrt{t}}, \triangle_t \leq \frac{2LC_B U_0}{\sqrt{t}}\} \\ \leq 2c_r t_0 + 2LC_B \frac{U_0}{\sqrt{t}} \mathbf{1}\{\triangle_t \leq \frac{2LC_B U_0}{\sqrt{t}}\}.$$

Denote $A_t := \frac{1}{\sqrt{t}} \mathbf{1} \{ \Delta_t \le \frac{2LC_B U_0}{\sqrt{t}} \}$, by Hoeffding's inequality we have with probability at least $1 - \alpha$,

$$\sum_{t} A_t < \sum_{t} \mathbb{E}[A_t] + \sqrt{\log T \log \frac{1}{\alpha}}.$$

Noting that $\mathbb{E}[\sum_t A_t] \leq 2\gamma L C_B U_0 \log T$ for $\beta = 1$ and $\mathbb{E}[\sum_t A_t] \leq \frac{2\gamma}{1-\beta} (L C_B U_0)^{\beta} T^{\frac{1-\beta}{2}}$ for $0 \leq \beta < 1$. Then the claim holds.

D Proof of Results in Section 4.2

To lighten the notation, in this section we denote θ_i the underlying parameter of arm i. In the following analysis, without special explaination, all the *c* and *C* denote absolute constants. Sometimes we state the inequality of type $A_1 \leq C \log(A_2/\alpha)A_3$ holds with probability at least $1 - \alpha$ while in proof we derive the results hold with $1 - c\alpha$ for some constant c. In fact, they are equivalent by re-scaling α and changing *C* to some larger constant.

D.1 Proof of Theorem 4.1

Lemma D.1. If after the warm up stage of length Ks_0 , the estimator $\hat{\theta}_{Ks_0,i}$ achieves the following error bound with probability at least $1 - \alpha$,

$$\sup_{i \in [K]} \|\hat{\theta}_{Ks_0, i} - \theta_i\| \le h_0 \coloneqq \frac{h_{sub}}{8LC_B}$$

With $h = h_{sub}$ in Algorithm 2, we have $\mathbb{P}\{a_t^* \in \hat{K}_t, \hat{K}_t \cap K_{sub} = \emptyset\} \ge 1 - \alpha$ holds uniformly for all $Ks_0 < t \le T$.

Proof. Firstly, to show $a_t^* \in \hat{K}_t$, without loss of generality we assume that $a_t^* \neq 1$, and $\underset{i \in [K]}{\operatorname{argmax}} \mu(X_t^T \hat{\theta}_{Ks_0,i}) = 1$. Then by the optimality of $\theta_{a_t^*}$, condition on $\sup_{i \in [K]} \|\hat{\theta}_{Ks_0,i} - \theta_i\| \leq h_0$,

$$\mathbb{P}(a_t^* \notin \hat{K}_t) = \mathbb{P}(\mu(X_t^T \hat{\theta}_{Ks_0, a_t^*}) < \mu(X_t^T \hat{\theta}_{Ks_0, 1}) - h/2) \\ \leq \mathbb{P}(\mu(X_t^T \theta_{a_t^*}) - h/8 < \mu(X_t^T \theta_1) + h/8 - h/2) = 0.$$

Now for any $j \in K_{sub}$, we have condition on $\sup_{i \in [K]} \|\hat{\theta}_{Ks_0,i} - \theta_i\| \le h_0$,

$$\mathbb{P}(j \in \hat{K}_t) \le \mathbb{P}(\mu(X_t^T \hat{\theta}_{Ks_0, a_t^*}) - h/2 < \mu(X_t^T \hat{\theta}_{Ks_0, j})) \\ \le \mathbb{P}(\mu(X_t^T \theta_{a_t^*}) - 3h/4 < \mu(X_t^T \theta_j) + h/4) = 0,$$

where the final equation is due to the sub-optimality gap assumed in Assumption 5.

Proof of Theorem 4.1. We first show the following lemma, which converts the regret bound under margin condition to the estimation error bound:

Lemma D.2. Under the (β, γ) -margin condition, given h_0 defined in Lemma D.1, suppose there exists some s_0 such that with a warm up stage of length Ks_0 , $\sup_{i \in [K]} \|\hat{\theta}_{t,i} - \theta_i\| \le h_0$, and there exists some $t_0, U_0(\alpha)$ such that with probability at least $1 - \alpha$,

$$\sup_{i \in K_{opt}} \|\hat{\theta}_{t,i} - \theta_i\| \le \frac{U_0(\alpha)}{\sqrt{t}}, \quad \forall t_0 \le t \le T.$$

Then, we have with probability at least $1 - 2\alpha$, for some constant C,

$$\operatorname{Reg}(T) \le C \cdot \begin{cases} c_r t_0 + \gamma (LC_B U_0(\alpha))^2 (\log T + o(1)), & \beta = 1\\ c_r t_0 + \frac{\gamma}{1 - \beta} (LC_B U_0(\alpha))^{1 + \beta} (T^{\frac{1 - \beta}{2}} + o(1)), & 0 < \beta < 1. \end{cases}$$

Proof of Lemma D.2. Denoting $E_t := \{\hat{K}_t \cap K_{sub} = \emptyset, a_t^* \in \hat{K}_t\}$, we have with probability at least $1 - \alpha$,

$$\begin{aligned} \operatorname{Reg}(T) &\leq 2c_r t_0 + L \sum_{t_0 < t \leq T} X_t^T (\theta_{a_t^*} - \theta_{a_t}) \\ &\leq 2c_r t_0 + L \sum_{t_0 < t \leq T} X_t^T (\theta_{a_t^*} - \theta_{a_t}) \mathbf{1} \{ \sup_{i \in K_{opt}} \| \hat{\theta}_{t,i} - \theta_i \| \leq \frac{U_0(\alpha)}{\sqrt{t}}, E_t \} \\ &\leq 2c_r t_0 + L \sum_{t_0 < t \leq T} X_t^T (\theta_{a_t^*} - \theta_{a_t}) \mathbf{1} \{ \sup_{i \in K_{opt}} \| \hat{\theta}_{t,i} - \theta_i \| \leq \frac{U_0(\alpha)}{\sqrt{t}}, \Delta_t \leq \frac{2LC_B U_0(\alpha)}{\sqrt{t}}, E_t \} \\ &\leq 2c_r t_0 + L \sum_{t_0 < t \leq T} \frac{2C_B U_0(\alpha)}{\sqrt{t}} \mathbf{1} \{ \Delta_t \leq \frac{2LC_B U_0(\alpha)}{\sqrt{t}} \}. \end{aligned}$$

Let $A_t = \mathbf{1}\{\Delta_t < \frac{2LC_B U_0(\alpha)}{\sqrt{t}}\}$. Then A_t is a sequence of independent 0-1 valued random variable such that $\mathbb{P}(A_t = 1) \le \gamma(\frac{2LC_B U_0(\alpha)}{\sqrt{t}})^{\beta}$. Then Hoeffding's inequality implies with probability at

least $1 - \alpha$,

$$\sum_{t_0 \le t \le T} \frac{1}{\sqrt{t}} A_t \le \mathbb{E}\left[\sum_{1 \le t \le T} \frac{1}{\sqrt{t}} A_t\right] + \sqrt{\log T \cdot \log(\frac{1}{\alpha})}.$$

Notice that $\mathbb{E}\left[\sum_{1 \le t \le T} \frac{1}{\sqrt{t}} A_t\right] \le CLC_B \gamma U_0(\alpha) \log T$ when $\beta = 1$ and $\mathbb{E}\left[\sum_{1 \le t \le T} \frac{1}{\sqrt{t}} A_t\right] \le C \frac{\gamma}{1-\beta} (LC_B U_0(\alpha))^{\beta} T^{\frac{1-\beta}{2}}$ when $0 < \beta < 1$. This completes the proof. \Box

Given Lemma D.2, we need only show that for both the private OLS estimator and the private SGD estimator, we can find the corresponding s_0, t_0 and $U_0(\alpha)$.

Lemma D.3 (Result of OLS estimator). Given $h_0 = \frac{h_{sub}}{8LC_B}$ and $\lambda_0 = (2LC_B)^{-1} (\frac{p'}{2\gamma})^{1/\beta}$, $r_{opt} := |K_{opt}|/K$, under the (β, γ) -margin condition,

$$s_0 = CK(\frac{C_B\sigma_{\epsilon} + \sigma_{\varepsilon,\delta}}{\min\{\lambda_0, h_0\}p'\kappa_l r_{opt}})^2 (d + \log(TK/\alpha))$$
$$t_0 = 2Ks_0,$$
$$U_0(\alpha) = \frac{K(C_B\sigma_{\epsilon} + \sigma_{\varepsilon,\delta})\sqrt{d + \log(TK/\alpha)}}{\kappa_l p'}.$$

satisfy the requirements in Lemma D.2.

Lemma D.4 (Result of SGD estimator). Given $h_0 = \frac{h_{sub}}{8LC_B}$ and $\lambda_0 = (2LC_B)^{-1} (\frac{p'}{2\gamma})^{1/\beta}$, $r_{opt} := |K_{opt}|/K$, under the (β, γ) -margin condition,

$$s_{0} = C \left(\frac{Kr_{\varepsilon,d}}{\zeta \kappa_{l} p' r_{opt} \min\{\lambda_{0}, h_{0}\}} \right)^{2} \log(KT \log(KT)/\alpha)$$
$$t_{0} = Ks_{0} + 1,$$
$$U_{0}(\alpha) = C \frac{K \sqrt{\log((KT \log KT)/\alpha)} r_{\varepsilon,d}}{\zeta \kappa_{l} p'},$$

satisfy the requirements in Lemma D.2.

Then Theorem 4.1 follows from combining Lemma D.2, D.3 and D.4.

Remark. Notice that in the statement of Lemma D.3 and Lemma D.4, there exists a term r_{opt} . That is because of our assumption $\mathbb{P}((v^T X)^2 \mathbf{1}\{X_t \in U_i\} > \kappa_l/K) > p'$. In fact, a more natural assumption should be $\mathbb{P}((v^T X)^2 \mathbf{1}\{X_t \in U_i\} > \kappa_l/|K_{opt}|) > p'$. In that case, we have $r_{opt} = 1$, which leads to more refined results.

The proof of Lemma D.3 and Lemma D.4 needs the following result: For a fixed $\beta \in (0,1]$, we define $h_0 = \frac{h_{sub}}{8LC_B}, \lambda_0 = (2LC_B)^{-1}(\frac{p'}{2\gamma})^{1/\beta}, A_t := \{\sup_{i \in K_{opt}} \|\hat{\theta}_{t,i} - \theta_i\| \le \lambda_0\}, H_0 := \{\sup_{i \in [K]} \|\hat{\theta}_{Ks_0,i} - \theta_i\| \le h_0\}.$

Lemma D.5. Define \mathcal{F}_t the filtration generated by $\{X_i\}_{i \in [t]}, \{\epsilon_i\}_{i \in [t]}$ together with all randomness from $\{\psi_i\}_{i \in [t]}$. Then we have:

$$\lambda_{\min}(\mathbb{E}[X_t X_t \mathbf{1}\{a_t = i\} | \mathcal{F}_{t-1}]) \ge \frac{p' \kappa_l}{2K} \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_0}, \quad \forall i \in K_{opt}.$$

Proof. We have for every unit vector v

$$\begin{split} \mathbb{E}[v^{T}X_{t}X_{t}^{T}v\mathbf{1}\{a_{t}=i\}|\mathcal{F}_{t-1}] \\ \geq & \mathbf{1}_{H_{0}}\frac{\kappa_{l}}{K}\mathbb{E}[\mathbf{1}\{|v^{T}X\mathbf{1}\{X_{t}\in U_{i}\}|^{2}\geq \kappa_{l}/K, a_{t}=i, A_{t-1}\}|\mathcal{F}_{t-1}] \\ \geq & \mathbf{1}_{H_{0}}\mathbf{1}_{A_{t-1}}\frac{\kappa_{l}}{K}\mathbb{E}[\mathbf{1}\{|v^{T}X\mathbf{1}\{X_{t}\in U_{i}\}|^{2}\geq \kappa_{l}/K\}-\mathbf{1}\{a_{t}\neq i, X_{t}\in U_{i}, A_{t-1}\}|\mathcal{F}_{t-1}] \\ \geq & \mathbf{1}_{H_{0}}\mathbf{1}_{A_{t-1}}\frac{\kappa_{l}}{K}[p'-\mathbb{P}(\{a_{t}\neq i, X_{t}\in U_{i}\}\cap H_{0}\cap A_{t-1}|\mathcal{F}_{t-1})]. \\ \mathbb{P}(\{a_{t}\neq i, X_{t}\in U_{i}\}\cap H_{0}\cap A_{t-1}|\mathcal{F}_{t-1})=\mathbf{1}_{H_{0}}\mathbf{1}_{A_{t-1}}\mathbb{P}(\{a_{t}\neq i, X_{t}\in U_{i}\}\cap E_{t}\cap A_{t-1}|\mathcal{F}_{t-1}) \\ \leq & \mathbf{1}_{A_{t-1}}\mathbf{1}_{H_{0}}\mathbb{P}(\Delta_{t}<2LC_{B}\lambda_{0}) \\ \leq & \mathbf{1}_{A_{t-1}}\mathbf{1}_{H_{0}}\frac{p'}{2}, \end{split}$$

where the last inequality is by the choice of λ_0 . Then the proof is finished.

D.2 Proof of Lemma D.3

We first establish the lower bound of the sample-covariance matrix sampled by the greedy action based on the following matrix-martingale concentration result:

Lemma D.6 (Theorem 3.1 in [40]). Let z^1, \ldots, z^t be a sequence of random, positive-semidefinite $d \times d$ matrices adapted to a filtration \mathcal{F}'_t , let $Z_t := \sum_{i=1}^t z^i$ and $\tilde{Z}_t := \sum_{i=1}^t \mathbb{E}[z^i|\mathcal{F}'_{i-1}]$. Suppose that $\lambda_{\max}(z^i) \leq R^2$ almost surely for all *i*, then for any μ and $\alpha \in (0, 1)$,

$$\mathbb{P}[\lambda_{\min}(Z_t) \le (1-\alpha)\mu, \lambda_{\min}(\tilde{Z}_t) \ge \mu] \le d(\frac{1}{e^{\alpha}(1-\alpha)^{1-\alpha}})^{\mu/R^2}.$$

Now we can show the following result:

Lemma D.7. For $t_1 < t_2 \in \mathbb{N}$ such that $(t_2 - t_1) \cdot \frac{\kappa_l p'}{8K} > 10C_B^2 \log(d/\alpha')$, for a fixed $i \in [K]$ we have

$$\mathbb{P}(\lambda_{\min}(\sum_{t=t_1}^{t_2} X_t X_t \mathbf{1}\{a_t=i\}) \le \frac{t_2 - t_1}{8K} \kappa_l p', \sup_{t_1 \le t \le t_2, i \in K_{opt}} \|\hat{\theta}_{t,i} - \theta_i\| \le \lambda_0, H_0) \le \alpha'.$$

Proof. Denote $S_{t_1,t_2} \coloneqq \bigcap_{t_1 \leq t \leq t_2} A_t$, by Lemma D.5 we have

$$\lambda_{\min}(\sum_{t=t_1}^{t_2} \mathbb{E}[X_t X_t^T \mathbf{1}\{a_t=i\} | \mathcal{F}_{t-1}]) \ge \sum_{t=t_1}^{t_2} \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_0} \frac{\kappa_l p'}{2K}.$$

That implies

$$\mathbb{P}(\lambda_{\min}(\sum_{t=t_{1}}^{t_{2}}X_{t}X_{t}^{T}\mathbf{1}\{a_{t}=i\}) \leq \frac{t_{1}-t_{2}}{4K}\kappa_{l}p', S_{t_{1},t_{2}}, H_{0})$$

$$\leq \mathbb{P}(\lambda_{\min}(\sum_{t=t_{1}}^{t_{2}}X_{t}X_{t}^{T}\mathbf{1}\{a_{t}=i\}) \leq \frac{t_{1}-t_{2}}{4K}\kappa_{l}p', \mathbb{E}[X_{t}X_{t}^{T}\mathbf{1}\{a_{t}=i\}|\mathcal{F}_{t-1}]) \geq (t_{2}-t_{1})\frac{\kappa_{l}p'}{2K}).$$

Then selecting $\alpha = 1/2$ and $\mu = (t_2 - t_1) \cdot \frac{\kappa_l p'}{4K}$ in Lemma D.6, we have

$$\mathbb{P}(\lambda_{\min}(\sum_{t=t_1}^{t_2} X_t X_t^T \mathbf{1}\{a_t=i\}) \le (t_2 - t_1) \frac{\kappa_l p'}{8K}, S_{t_1, t_2}, H_0) \le d(\frac{1}{\sqrt{e/2}})^{10\log(\frac{d}{\alpha'})} \le \alpha'.$$

That leads to the claim.

In warm up stage, we have the following lemma.

Lemma D.8. As long as $s_0 \ge C(r_{opt}\kappa_l p')^{-2} \max\{\log \frac{1}{\alpha}, d\}$ for some absolute constant C, we have with probability at least $1 - \alpha$,

$$\lambda_{\min} (\sum_{t=1}^{Ks_0} \mathbf{1}\{a_t = i\} X_t X_t^T)^{-1} \le \frac{2}{s_0 p' r_{opt} \kappa_l}, \quad \forall i \in [K].$$

Proof. Since X_t are i.i.d. for $(i-1)s_0 + 1 \le t \le is_0$, using classical concentration results for i.i.d. sub-gaussian covariance matrix result (e.g. Theorem 6.5 in [42]), we have when $s_0 > C(r_{opt}\kappa_l p')^{-2} \max\{\log \frac{1}{\alpha}, d\}$, with probability at least $1 - \alpha$,

$$\begin{aligned} \|\frac{1}{s_0} \sum_{t=1}^{Ks_0} \mathbf{1}\{a_t = i\} X_t X_t^T - \mathbb{E}[X_1 X_1^T] \| &\leq c_1 (\sqrt{\frac{d}{s_0}} + \frac{d}{s_0}) + c_2 \max\{\sqrt{\frac{\log 1/\alpha}{s_0}}, \frac{\log 1/\alpha}{s_0}\} \\ &\leq c_3 (\sqrt{\frac{d}{s_0}} + \sqrt{\frac{\log(1/\alpha)}{s_0}}) \\ &\leq r_{out} p' \kappa_t / 2. \end{aligned}$$

On the other hand, we have by Markov's inequality

$$\lambda_{\min} \mathbb{E}[X_1 X_1^T] \ge \sum_{i \in K_{opt}} \lambda_{\min} \mathbb{E}[X_1 X_1^T \mathbf{1}\{X_1 \in U_i\}] \ge r_{opt} \kappa_l p'$$

Thus we have with probability at least $1 - \alpha$,

$$\lambda_{\min}(\sum_{t=1}^{Ks_0} X_t X_t^T) \ge s_0 r_{opt} p' \kappa_l / 2.$$

Now we can claim our first result about the private OLS-estimator in the warm up stage:

Lemma D.9. Selecting s_0 as in Lemma D.8. For the warm up stage with private-OLS-estimator and length Ks_0 , we have for any $\alpha > 0$, with probability at least $1 - \alpha$,

$$\sup_{i \in [K]} \|\hat{\theta}_{t,i} - \theta_i\| \leq \frac{(4C_B\sigma_\epsilon + \sigma_{\varepsilon,\delta})\sqrt{t(\log(\frac{TK}{\alpha}) + d)}}{s_0p'r_{opt}\kappa_l} \quad holds for all \ Ks_0 \leq t \leq T.$$

Proof. Denote $U_t = \sum_{s=1}^t (\mathbf{1}\{a_s = i\}X_sX_s^T + (\mathbf{1}\{a_s = i, s \le Ks_0\} + \mathbf{1}\{s > Ks_0\})W_s) + \tilde{c}\sqrt{t}I_d$, we have

$$\begin{aligned} \hat{\theta}_{t,i} &= U_t^{-1} (\sum_{s=1}^t \mathbf{1}\{a_s = i\} X_s y_s + (\mathbf{1}\{a_s = i, s \le K s_0\} + \mathbf{1}\{s > K s_0\}) \xi_s) \\ &= U_t^{-1} (\sum_{s=1}^t \mathbf{1}\{a_s = i\} [X_s X_s^T \theta_i + X_s \epsilon_s] + (\mathbf{1}\{a_s = i, s \le K s_0\} + \mathbf{1}\{s > K s_0\}) \xi_s) \\ &= \theta_i + U_t^{-1} (\sum_{s=1}^t (\mathbf{1}\{a_s = i\} X_s \epsilon_s + (\mathbf{1}\{a_s = i, s \le K s_0\} + \mathbf{1}\{s > K s_0\}) (\xi_s - W_s \theta_i)) - \tilde{c} \sqrt{t} I_d \theta_i) \end{aligned}$$

By $\|\sum_{s=1}^{Ks_0} \mathbf{1}\{a_s = i\}W_s + \sum_{s=Ks_0+1}^t W_s\| \leq \tilde{c}\sqrt{t}, \forall Ks_0 \leq t \leq T, i \in [K]$ with probability at least $1 - \alpha$, we have with probability at least $1 - 2\alpha$,

$$[\lambda_{\min}(U)]^{-1} \le \lambda_{\min}(\sum_{s=1}^{Ks_0} \mathbf{1}\{a_s = i\}X_s X_s^T)^{-1} \le \frac{2}{s_0 r_{opt} p' \kappa_l}, \quad \forall Ks_0 \le t \le T$$

On the other hand, we have by the concentration of sub-gaussian random vector, the following bounds hold with probability at least $1 - \alpha/(T^2K)$:

$$\left\|\sum_{s=1}^{\iota} \mathbf{1}\{a_s = i\} X_s \epsilon_s \right\| \le C C_B \sigma_\epsilon \sqrt{t(d + \log(TK/\alpha))},\tag{17}$$

$$\|\sum_{s=1}^{l} (\mathbf{1}\{a_s = i, s \le Ks_0\} + \mathbf{1}\{s > Ks_0\})\xi_s\| \le C\sigma_{\varepsilon,\delta}\sqrt{t(d + \log(TK/\alpha))},$$
(18)

$$\left\|\sum_{s=1}^{K_{s_0}} \mathbf{1}\{a_s=i\}W_s\theta_i + \sum_{s=Ks_0+1}^t W_s\theta_i\right\| \le \tilde{c}\sqrt{t}\|\theta_i\| \le C\sigma_{\varepsilon,\delta}\sqrt{t(d+\log(TK/\alpha))}.$$
 (19)

Gathering all bounds together, we have with probability at least $1 - (2 + \frac{1}{\tau^2})\alpha$,

$$\sup_{i \in [K]} \|\hat{\theta}_{t,i} - \theta_i\| \le \frac{2C}{s_0 p' r_{opt} \kappa_l} (C_B \sigma_\epsilon + \sigma_{\varepsilon,\delta}) \sqrt{t(\log(TK/\alpha) + d)}.$$

That finishes the proof.

Lemma D.10. As long as

i

$$s_0 \ge CK(\frac{C_B \sigma_{\epsilon} + \sigma_{\varepsilon,\delta}}{\min\{\lambda_0, h_0\} r_{opt} p' \kappa_l})^2 (d + \log(TK/\alpha)),$$

we have with probability at least $1 - \alpha$,

$$\sup_{i \in [K]} \|\hat{\theta}_{Ks_0, i} - \theta_i\|_2 \le \min\{\lambda_0, h_0\},\tag{20}$$

$$\sup_{i \in K_{opt}} \|\hat{\theta}_{s,i} - \theta_i\|_2 \le \lambda_0 \text{ holds uniformly for } Ks_0 \le s \le (K+1)s_0,$$
(21)

$$C\frac{K(C_B\sigma_\epsilon + \sigma_{\varepsilon,\delta})\sqrt{d + \log(TK/\alpha)}}{\sqrt{t - Ks_0}\kappa_l p'} \le \lambda_0 \text{ holds for all } t \ge 2Ks_0.$$
(22)

Proof. To show (20),(21), we can just plug the value of s_0 into the upper bound in Lemma D.9. (22) comes directly from the value of s_0 .

Now, we can show the following result:

Lemma D.11. With the choice of s_0 same as in Lemma D.10, for $t > Ks_0$, denote $t' = t - Ks_0$ and $\tilde{t}_0 = 2Ks_0$, we have if

 H_0 holds and $\|\hat{\theta}_{t,i} - \theta_i\|_2 \leq \min\{\tilde{U}_s(\alpha), \lambda_0\}$ holds uniformly for $i \in K_{opt}, \tilde{t}_0 \leq s \leq t$,

with probability at least $1 - \sum_{j=1}^{t'} \frac{2}{j^2} \alpha$, then

 H_0 holds and $\|\hat{\theta}_{t,i} - \theta_i\|_2 \le \min\{\tilde{U}_s(\alpha), \lambda_0\}$ holds uniformly for $i \in K_{opt}, \tilde{t}_0 \le s \le t+1$,

with probability at least $1 - \sum_{j=1}^{t'+1} \frac{2}{j^2} \alpha$, where

$$\tilde{U}_s(\alpha) = C \frac{K(C_B \sigma_\epsilon + \sigma_{\varepsilon, \delta}) \sqrt{d + \log(TK/\alpha)}}{\sqrt{s} \kappa_l p'}$$

 $\textit{Proof. Denote } S_{\tilde{t}_0,t} = \{ \| \hat{\theta}_{s,i} - \theta_i \| \leq \min\{\tilde{U}_s(\alpha), \lambda_0\}, \forall K \in K_{opt}, \forall \tilde{t}_0 \leq s \leq t \}, \tilde{A}_t = \{ \| \hat{\theta}_{s,i} - \theta_i \| \leq t \} \}$ $\{\sup_{i\in K_{opt}}\|\hat{ heta}_{i,t}- heta_i\|\leq ilde{U}_t(lpha)\}$, we have by Lemma D.7

$$\mathbb{P}(S_{\tilde{t}_{0},t}, H_{0}, \lambda_{\min}(\sum_{s=1}^{t} X_{s} X_{s} \mathbf{1}\{a_{s}=i\}) > \frac{t' \kappa_{l} p'}{8K}) \ge 1 - \frac{\alpha}{2KT^{2}}.$$

Applying the inequalities (17),(18) (19), we have

$$\mathbb{P}(H_0, S_{\tilde{t}_0, t}, A_{t+1}) \ge 1 - \sum_{j=1}^{t'} \frac{2}{j^2} \alpha - \frac{3\alpha}{2T^2} - \sum_{i \in K_{opt}} \mathbb{P}(H_0, S_{\tilde{t}_0, t}, \tilde{A}_{t+1}, \lambda_{\min}(\sum_{s=1}^t X_s X_s \mathbf{1}\{a_s = i\}) \le \frac{t' \kappa_l p'}{4})$$
$$\ge 1 - \sum_{j=1}^{t'} \frac{2}{j^2} \alpha - \frac{2\alpha}{T^2}$$
$$\ge 1 - 2\sum_{j=1}^{t'+1} \frac{1}{j^2} \alpha.$$

By the selection of s_0 , we have $\tilde{U}_s(\alpha) \leq \lambda_0$ for $\tilde{t}_0 \leq s \leq t+1$, and as a result, $\mathbb{P}(H_0, S_{\tilde{t}_0, t+1}) = \mathbb{P}(H_0, S_{\tilde{t}_0, t}, \tilde{A}_{t+1})$. Thus the claim holds.

Proof of Lemma D.3. Lemma D.3 is implied directly by Lemma D.11 and Lemma D.10.

D.3 Proof of Lemma D.4

Proof. For the estimator $\hat{\theta}_{Ks_0,i}$ at the end of warm up stage, since the action is independent of the contexts, every $\hat{\theta}_{Ks_0,i}$ can be seen as an output of performing private gradient descent over s_0 i.i.d. samples. Without loss of generality, we perform the analysis for the parameter of the first arm $\hat{\theta}_{Ks_0,1}$ (notice that by the sampling strategy in the warm up stage, we have $\hat{\theta}_{Ks_0,1} = \hat{\theta}_{s_0,1}$). The result for other $\hat{\theta}_{Ks_0,i}$ can be established using the same argument. For $2 \le t \le s_0$,

$$\begin{aligned} \|\hat{\theta}_{t,i} - \theta_i\|^2 &= \|\hat{\theta}_{t-1,i} - \eta_t \hat{g}_t - \theta_i\|^2 \\ &= \|\hat{\theta}_{t-1,i} - \theta_i\|^2 - 2\eta_t \hat{g}_t^T (\hat{\theta}_{t-1,i} - \theta_i) + 2\eta_t^2 \|\hat{g}_t\|^2 \end{aligned}$$

Here $\hat{g}_t := \Psi_{\varepsilon}[(\mu(X_t^T \hat{\theta}_{t,i}) - r_t)X_t]$, by the unbiasedness of Ψ_{ε} in Lemma 2.2 we have

$$\begin{split} & \mathbb{E}[\Psi_{\varepsilon}((\mu(X_t^T\hat{\theta}_{t-1,i}) - r_t)X_t)^T(\hat{\theta}_{t-1,i} - \theta_i)|\mathcal{F}_{t-1}] \\ &= \mathbb{E}[(\mu(X_t^T\hat{\theta}_{t-1,i}) - \mu(X_t^T\theta_i))X_t^T(\hat{\theta}_{t-1,i} - \theta_i)|\mathcal{F}_{t-1}] \\ &\geq \zeta \mathbb{E}[[X_t^T(\hat{\theta}_{t-1,i} - \theta_i)]^2|\mathcal{F}_{t-1}] \\ &\geq \zeta \kappa_l r_{opt} p' \|\hat{\theta}_{t-1,i} - \theta_i\|^2. \end{split}$$

We get

$$\|\hat{\theta}_{t,i} - \theta_i\|^2 \le (1 - 2\zeta r_{opt}\kappa_l p'\eta_t) \|\hat{\theta}_{t-1,i} - \theta_i\|^2 + 2\eta_t (\mathbb{E}[\hat{g}_t|\mathcal{F}_{t-1}] - \hat{g}_t)^T (\hat{\theta}_{t-1,i} - \theta_i) + 2\eta_t^2 \|\hat{g}_t\|^2.$$

Notice $\|\hat{g}_t\|_2^2$ is upper bounded by $r_{\varepsilon,d}^2$. Now using the same argument as in the proof of Proposition 1 of [28] leads to the following result:

Lemma D.12. If we pick $\eta_t = 1/(r_{opt}\zeta\kappa_l p't)$ in the warm up stage, then with probability at least $1 - \alpha$,

$$\sup_{i \in [K]} \|\hat{\theta}_{Ks_0, i} - \theta_i\|^2 \le C \frac{(\log(\log(KT)/\delta) + 1)r_{\varepsilon, \delta}^2}{\zeta^2 \kappa_l^2 r_{opt}^2 p'^2 s_0}.$$
(23)

Notice that in our algorithm, when $t > Ks_0$, for any $i \in K_{opt}$, the private gradient descent formula is given by

$$\hat{\theta}_{t,i} = \hat{\theta}_{t-1,i} - \eta_t \tilde{g}_t,$$

with $\tilde{g}_t = \mathbf{1}\{a_t = i\}\hat{g}_t + \mathbf{1}\{a_t \neq i\}\Psi_{\varepsilon}(0)$. Again without loss of generality we assume that $1 \in K_{opt}$, and we provide the analysis for i = 1, the argument is same for other $i \in K_{opt}$:

$$\begin{split} \mathbb{E}[\tilde{g}^{T}(\hat{\theta}_{t-1,1} - \theta_{1})|\mathcal{F}_{t-1}] &= \mathbb{E}[\mathbf{1}\{a_{t} = i\}\hat{g}^{T}(\hat{\theta}_{t-1,1} - \theta_{1})|\mathcal{F}_{t-1}] \\ &= \mathbb{E}[\mathbf{1}\{a_{t} = i\}(\mu(X_{t}^{T}\hat{\theta}_{t-1,1}) - \mu(X_{t}^{T}\theta_{i}))X_{t}^{T}(\hat{\theta}_{t-1,1} - \theta_{1})|\mathcal{F}_{t-1}] \\ &\geq \zeta \mathbb{E}[\mathbf{1}\{a_{t} = i\}[X_{t}^{T}(\hat{\theta}_{t-1,1} - \theta_{1})]^{2}|\mathcal{F}_{t-1}] \\ &\geq \mathbf{1}_{A_{t},H_{0}}\zeta \kappa_{l}p'\eta_{t}\|\hat{\theta}_{t-1,1} - \theta_{1}\|^{2}/K \end{split}$$

select $\eta_t \coloneqq K/(\zeta \kappa_l p't')$, with $t' = t - (K-1)s_0$ we have then

$$\|\hat{\theta}_{t,1} - \theta_1\|_2^2 \le (1 - \frac{2}{t'} \mathbf{1}_{A_t,H_0}) \|\hat{\theta}_{t-1,1} - \theta_1\|^2 + \frac{2K}{\zeta \kappa_l p' t'} (\mathbb{E}[\tilde{g}_t | \mathcal{F}_{t-1}] - \tilde{g}_t)^T (\hat{\theta}_{t-1,1} - \theta_1) + 2(\frac{Kr_{\varepsilon,d}}{\zeta \kappa_l p' t'})^2 (\tilde{\theta}_{t-1,1} - \theta_1) + 2(\frac{Kr_{\varepsilon,d}}{\zeta \kappa_l p' t'})^$$

If we denote $S_t := \bigcap_{s=Ks_0}^t A_s$, then using the above inequality recursively until $t = Ks_0 + 1$ (i.e. until $t' = s_0 + 1$), we have

$$\begin{aligned} \mathbf{1}_{S_{t-1},H_0} \|\hat{\theta}_{t,1} - \theta_1\|^2 &\leq \frac{s_0(s_0 - 1)}{t'(t' - 1)} \|\hat{\theta}_{Ks_0,1} - \theta_1\|^2 + 2(\frac{Kr_{\varepsilon,d}}{\zeta\kappa_l p't'})^2 \\ &+ \frac{2K}{(t' - 1)t'\zeta\kappa_l p'} \sum_{s=Ks_0+1}^t (\mathbb{E}[\tilde{g}_s|\mathcal{F}_{t-1}] - \tilde{g}_s)^T (\hat{\theta}_{s-1,1} - \theta_1). \end{aligned}$$

Then it follows from the same proof as in Proposition 1 in [28] that for any fixed $Ks_0 < t \le T$, we have with probability at least $1 - \alpha/T$,

$$\mathbf{1}_{S_{t-1},H_0} \|\hat{\theta}_{t,1} - \theta_1\|^2 \le \frac{s_0(s_0 - 1)}{t'(t' - 1)} \|\hat{\theta}_{Ks_0,1} - \theta_1\|^2 + C \frac{K^2 (\log(TK\log(TK)/\alpha) + 1)r_{\varepsilon,d}^2}{\zeta^2 \kappa_l^2 p'^2 t'},$$
(24)

Now choose $s_0 \geq 2C \frac{K^2(\log(TK\log(TK)/\alpha) + 1)r_{\varepsilon,d}^2}{r_{opt}^2 \zeta^2 \kappa_l^2 p'^2 \min\{\lambda_0, h_0\}^2}$, so that the second term in (24) is less or equal to $\lambda_0/2$, we have $\mathbb{P}(S_{Ks_0+1}, H_0) \geq 1 - 2\alpha$ by (23). And by calling (24) recursively we can get $\mathbb{P}(S_{t-1}, H_0) > 1 - 2\alpha - \frac{t - Ks_0}{T} \alpha \geq 1 - 3\alpha, \forall Ks_0 < t \leq T$. Then with probability at least $1 - 3\alpha$, we have

$$\|\hat{\theta}_{t,1} - \theta_1\|^2 \le C \frac{K^2 (\log(3TK\log(TK)/\alpha)) r_{\varepsilon,d}^2}{\zeta^2 \kappa_l^2 p'^2 (t - (K-1)s_0)}, \quad \forall Ks_0 < t \le T.$$

The above inequality is because the term $\frac{s_0(s_0-1)}{t'(t'-1)}\|\hat{\theta}_{Ks_0,1}-\theta_1\|^2 \leq \frac{s_0(s_0-1)}{t'(t'-1)}\frac{\min\{\lambda_0,h_0\}}{2},$ which can be absorbed into the constant C.

E Proof of Theorem 3.2

In this section, we would give a proof on the Theorem 3.2 by combining the argument in [22] and the divergence contraction inequality in [15].

Proof of Theorem 3.2. Consider the two-arm stochastic contextual bandit environment: for each d-dimensional context i = 1 or 2, $x_{t,i} \sim \mathcal{N}(0, \frac{1}{d}I_d)$ independently. If choosing action a_t at time t, the reward y_t is generated via $y_t = x_{t,a_t}^T \theta + \epsilon_t$ with $\epsilon_t \sim_{i.i.d.} \mathcal{N}(0, 1)$. Given any fixed ε -LDP bandit algorithm π with $\varepsilon \leq 1$, we denote its decision at t-th step by a_t , by definition a_t can be seen as a function of current contextual $x_{t,1}, x_{t,2}$ and all history outputs $(x_{1,a_1}, y_1, x_{2,a_2}, y_2, \ldots, x_{t-1,a_{t-1}}, y_{t-1})$. Since the algorithm is under the ε -LDP constraint, each a_t can only access $S_t := (M_1(x_{1,a_1}, y_1), M_2(x_{2,a_2}, y_2), \ldots, M_{t-1}(x_{t-1,a_{t-1}}, y_{t-1}))$ with M_1, \ldots, M_{t-1} a sequence of ε -LDP mechanisms. We denote the distribution of S_t by Q_{θ}^t , and we have

$$\mathbb{E}_{\theta \sim Q_0} [\mathbb{E}_{Q_\theta}^t [(x_{t,a_t^*} - x_{t,a_t})^T \theta | x_{t,1}, x_{t,2}]] \\ = \mathbb{E}_{\theta \sim Q_0} [((x_{t,1} - x_{t,2})^T \theta)_+ Q_\theta^t (a_t(S_t, x_t) = 2) + ((x_{t,2} - x_{t,1})^T \theta)_+ Q_\theta^t (a_t(S_t, x_t) = 1)],$$
⁽²⁵⁾

where $(x)_+$ denote $\max\{x, 0\}$ and Q_0 denote the uniform distribution over $\triangle S_1^{d-1}$ with $\triangle > 0$ some positive number to be determined, we define Q_1, Q_2 as

$$\frac{dQ_1}{dQ_0} := \frac{((x_{t,1} - x_{t,2})^T \theta)_+}{Z_0}, \quad \frac{dQ_2}{dQ_0} := \frac{((x_{t,2} - x_{t,1})^T \theta)_+}{Z_0}$$

where $Z_0 = \mathbb{E}_{Q_0}[((x_{t,1} - x_{t,2})^T \theta)_+] = \mathbb{E}_{Q_0}[((x_{t,2} - x_{t,1})^T \theta)_+]$ is the normalization factor. Denote $r_t = ||x_{t,1} - x_{t,2}||, u_t = r_t^{-1}(x_{t,1} - x_{t,2})$, then the right hand side of (25) is lower bounded by

$$= Z_{0}(Q_{1} \circ Q_{\theta}^{t}(a_{t}(S_{t}, x_{t}) = 2) + Q_{2} \circ Q_{\theta}^{t}(a_{t}(S_{t}, x_{t}) = 1))$$

$$\geq_{(a)} Z_{0}(1 - \mathrm{TV}(Q_{1} \circ Q_{\theta}^{t}, Q_{2} \circ Q_{\theta}^{t}))$$

$$\geq_{(b)} \frac{Z_{0}}{2} \exp(-D_{KL}(Q_{1} \circ Q_{\theta}^{t} || Q_{2} \circ Q_{\theta}^{t}))$$

$$=_{(c)} \frac{Z_{0}}{2} \exp(-D_{KL}(Q_{1} \circ Q_{\theta}^{t} || Q_{1} \circ Q_{\theta-2(u_{t}^{T}\theta)u_{t}}^{t})))$$

$$\geq_{(d)} \frac{Z_{0}}{2} \exp(-\mathbb{E}_{Q_{1}}[D_{KL}(Q_{\theta}^{t} || Q_{\theta-2(u_{t}^{T}\theta)u_{t}}^{t})]), \quad (F.1)$$

where $D_{KL}(\cdot \| \cdot)$ denote the KL-divergence, $TV(\cdot, \cdot)$ denote the total variation distance and $Q_i \circ Q_{\theta}^t$ means $\mathbb{E}_{\theta \sim Q_i}[Q_{\theta}^t]$. The (a) inequality comes from the fundamental limit of two-point testing (see e.g. Section 15.2 in [42]), and the (b) inequality comes from Lemma 2.6 of [41], the (c) equality comes from Lemma 8 in [22] and the (d) inequality comes from the strongly-convexity of KL-divergence. Now by chain rule of KL-divergence, the divergence contraction inequality in Theorem 1 of [15] and the formula of KL-divergence between Gaussian distributions, we have

$$D_{KL}(Q_{\theta}^{t} \| Q_{\theta-2(u_{t}^{T}\theta)u_{t}}^{t}) = \sum_{s=1}^{t-1} \mathbb{E}_{Q_{\theta}^{s-1}}[D_{KL}(P_{\theta}^{t}(\cdot|S_{s-1}) \| P_{\theta-2(u_{t}^{T}\theta)u_{t}}^{t}(\cdot|S_{s-1}))]$$
$$\leq \sum_{s=1}^{t-1} \frac{c}{2}(e^{\varepsilon} - 1)^{2}(2(u_{t}^{T}\theta)^{2} \| u_{t} \|^{2})$$

By the argument of in [22], we have (F.1) is lower bounded by

$$\frac{r_t \Delta}{C\sqrt{d}} \exp(-C \frac{(e^{\varepsilon} - 1)^2 \Delta^2}{d+1} u_t^T (\sum_{s=1}^{t-1} x_{t,a_t} x_{t,a_t}^T) u_t).$$

Now taking expectation over $x_{t,1}, x_{t,2}$, and using the convexity of function $f(x) = \exp(-x)$ we get

$$\mathbb{E}_{x}\mathbb{E}_{\theta}\mathbb{E}_{Q_{\theta}^{t}}[x_{t,a_{t}^{*}}-x_{t,a_{t}}^{T}] \geq \frac{\triangle}{C\sqrt{d}}\exp(-\frac{C(e^{\varepsilon}-1)^{2}\triangle^{2}t}{d^{2}}).$$

Selecting $\Delta \approx \frac{d}{(e^{\varepsilon} - 1)\sqrt{t}}$ and taking summation over $1 \le t \le T$ leads to $\Omega(\sqrt{Td}/(e^{\varepsilon} - 1))$ lower bound, finally noticing $e^{\varepsilon} - 1 \le C\varepsilon$ for $\varepsilon \le 1$ leads to the desired lower bound when $\varepsilon \le 1$. \Box

F Additional Experiments

In this section, We evaluate all the four methods on two different privacy levels $\varepsilon = 0.5$ and 1 in a larger scale scheme. To be specific, for single-param setting we increase dimension d to 20 and increase the number of arms K to 20; for multi-param setting we increase dimension d to 10 and increase the number of the arms K to 10.

In this simulation we change the learning rate scheme of LDP-SGD from $\eta_t = c_1 d/(\kappa_l \zeta p_* t)$ to $\eta_t = c_2 d/(\kappa_l \zeta p_* \sqrt{t})$ for some $c_2 > 1$ for its better empirical performance, while other details in data generation process are the same as in Section 5. The first and second columns in Figure 3 are for single-param and multi-param settings, respectively, which are simulation studies on linear bandits. As we can see the proposed LDP-OLS and LDP-SGD algorithms can still achieve better performance against their competitors under different privacy constraints.



Figure 3: Simulation study in larger-scale scheme. We perform 10 replications for each case and plot the mean and 0.5 standard deviation of their regrets.

G Auto Loan Experiment Details

We use the same features selection as in [1, 11] in the dataset and select FICO score, the term of contract, the loan amount approved, prime rate, the type of car, and the competitor's rate as the feature vector for each customer. Note that a description of the data set (with descriptive statistics on the demand and available features) is available in [1]. The objective is to offer a personalized lending price (from a range of choices) based on personal information such as FICO score to a customer who will either accept or reject it. In contrast to linear bandits, the binary reward is non-linear. Therefore we leave LDP-UCB and LDP-OLS out of considerations. To formulate a bandit environment, first we need to recover the underlying true parameter. Since the lender's decision, i.e., the price for each customer, is not presented in the dataset, we follow [1, 11] and impute it by using the net-present value of futher payment minus the loan amount, i.e.,

$$p =$$
 Monthly Payment $\times \sum_{\tau=1}^{\text{ferm}} (1 + \text{Rate})^{-\tau} - \text{Loan Amount}$

After imputing the loan prices, to represent customers' binary loan choices, we employ the logit demand model. To be specific, given a price p and a context $x \in \mathbb{R}^d$, the binary variable *apply* takes value of 1 with probability $\frac{\exp(v)}{1+\exp(v)}$ and takes value of 0 with probability $\frac{1}{1+\exp(v)}$ where the linear predictor $v = (x, px)^T \theta^*$. We conduct one-hot encoding for categorical features in the dataset and use the python package sklearn [25] for the estimation of the underlying parameter θ^* . We use the interval [0, 25000] as the feasible region of the prices, which covers the lending prices computed from the dataset, and we discrete the feasible region uniformly into 25 options $\{p_i\}_{i \in [25]}$. We use LDP-SGD and LDP-GLOC to sequentially compute the loan prices for the 100,000 with

randomly selected customers in the dataset , and compute the company's expected regret based on the population model mentioned above under two privacy constraints scheme $\varepsilon = 0.5$ and $\varepsilon = 1$.



Figure 4: We perform 10 replications for each case and plot the mean and 0.5 standard deviation of their regrets.