## 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 Section5] 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 $\sigma^{2}$-sub-gaussian vector with bounded density, if for every $v \in S_{1}^{d-1}$, $v^{T} x$ is $\sigma^{2}$-sub-gaussian and its density function exists and is bounded by $\gamma$ for some $\gamma>0$. For such kind of random vector, [30] shows that it satisfies Assumption 3] with $\kappa_{l}=\frac{2 d}{3 \gamma K}$ 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, $\left\|x_{t, a}\right\| \leq$ $C_{B},\left\|r_{t}\right\| \leq c_{r}$ 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 $\psi_{t}^{O L S}$ is $(\varepsilon, \delta)$-LDP.

Proof of Proposition 3.2. Since we assume that the features and rewards are bounded, $\left\|x_{t, a}\right\| \leq$ $C_{B},\left\|r_{t}\right\| \leq c_{r}$ for all $t \in[T]$ and $a \in[K]$, we have $\left(\mu\left(x_{t, a_{t}}^{T} \hat{\theta}_{t-1}\right)-r_{t}\right) x_{t, a_{t}}$ bounded by $2 c_{r} C_{B}$. Lemma 2.2 implies that $\psi_{t}^{S G D}$ is $\varepsilon$-LDP.

## B. 2 Proof of Results in Section 4.1

Proof of Proposition 4.1. We simply denote $\psi_{t}^{O L S}$ by $\psi_{t}$ in this proof. At time $t$, for any two $x \neq x^{\prime}$, without loss of generality assuming the action corresponding $x$ and $x^{\prime}$ are $a_{t}=1$ and $a_{t}=2$, then the output corresponding $x, x^{\prime}$ is given by $\left(\psi_{t}\left(x, x^{T} \theta_{1}+\epsilon_{t}\right), \psi_{t}(0,0), \ldots, \psi_{t}(0,0)\right)$ and $\left(\psi_{t}(0,0), \psi_{t}\left(x^{\prime}, x^{T} \theta_{2}+\epsilon_{t}\right), \ldots, \psi_{t}(0,0)\right)$. Since $\psi_{t}(0,0)$ has the same distribution, we have for any subset $A_{1} \times A_{2} \times \cdots \times A_{K} \subset \mathbb{R}^{K d}$ with $A_{i}$ a Borel set in $\mathbb{R}^{d}$,

$$
\begin{align*}
& \frac{\mathbb{P}\left(\psi_{t}\left(x, x^{T} \theta_{1}+\epsilon_{t}\right) \in A_{1}, \psi_{t}(0,0) \in A_{2}, \ldots, \psi_{t}(0,0) \in A_{K}\right)}{\mathbb{P}\left(\psi_{t}(0,0) \in A_{1}, \psi_{t}\left(x^{\prime}, x^{\prime T} \theta_{2}+\epsilon_{t}\right) \in A_{2}, \ldots, \psi_{t}(0,0) \in A_{K}\right)} \\
= & \frac{\mathbb{P}\left(\psi_{t}\left(x, x^{T} \theta_{1}+\epsilon_{t}\right) \in A_{1}, \psi_{t}(0,0) \in A_{2}\right)}{\mathbb{P}\left(\psi_{t}\left(x^{\prime}, x^{\prime T} \theta_{2}+\epsilon_{t}\right) \in A_{2}, \psi_{t}(0,0) \in A_{1}\right)} \tag{8}
\end{align*}
$$

Set $\tilde{\psi}\left(v_{1}, v_{2}\right):=\left(\psi_{t}\left(v_{1}\right), \psi_{t}\left(v_{2}\right)\right)$, and $\left(v_{1}, v_{2}\right):=(x, 0),\left(v_{1}^{\prime}, v_{2}^{\prime}\right):=\left(0, x^{\prime}\right)$, then we have (8) equals to $\tilde{\psi}\left(v_{1}, v_{2}\right) / \tilde{\psi}\left(v_{1}^{\prime}, v_{2}^{\prime}\right)$, thus applying Lemma 2.4 to it implies that 8 is upper bounded by $e^{\varepsilon}+\delta \mathbb{P}\left(\psi_{t}\left(x^{\prime}, x^{T} \theta_{2}+\epsilon_{t}\right) \in A_{2}, \psi_{t}(0,0) \in A_{1}\right)^{-1}$, leading to the desired result.

Proof of Proposition 4.2. That is nearly the same as the proof of Proposition 4.1, but replacing $e^{\varepsilon}+\delta \mathbb{P}\left(\psi_{t}\left(x^{\prime}, x^{T} \theta_{2}+\epsilon_{t}\right) \in A_{2}, \psi_{t}(0,0) \in A_{1}\right)^{-1}$ by $e^{\varepsilon}$ in the last step.

## 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 $A_{1} \leq C \log \left(A_{2} / \alpha\right) A_{3}$ 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 $\alpha$ 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 $x_{t, a_{t}}^{T} \hat{\theta}_{t-1} \geq x_{t, a}^{T} \hat{\theta}_{t-1}$ for any time $t \in[T]$ and $a \in[K]$. Consequently we have the following upper bound for the instantaneous
regret at time $t$,

$$
\begin{aligned}
\max _{a \in[K]}\left(x_{t, a}-x_{t, a_{t}}\right)^{T} \theta^{\star} & \leq \max _{a \in[K]}\left(x_{t, a}-x_{t, a_{t}}\right)^{T}\left(\theta^{\star}-\hat{\theta}_{t-1}\right) \\
& \leq \max _{a, a^{\prime} \in[K]}\left(x_{t, a}-x_{t, a^{\prime}}\right)^{T}\left(\theta^{\star}-\hat{\theta}_{t-1}\right) \\
& \leq 2 \max _{a \in[K]}\left|x_{t, a}^{T}\left(\theta^{\star}-\hat{\theta}_{t-1}\right)\right| .
\end{aligned}
$$

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 $\mathrm{t}, x_{t, a}^{T}\left(\theta^{\star}-\hat{\theta}_{t-1}\right)$ is a $\frac{\kappa_{u}}{d}\left\|\theta^{\star}-\hat{\theta}_{t-1}\right\|^{2}$-sub-gaussian random variable. Now by the maximal concentration inequality for a sub-gaussian sequence, we have with probability at least $1-\frac{\alpha}{T}$,

$$
\max _{a \in[K]}\left|x_{t, a}^{T}\left(\theta^{\star}-\hat{\theta}_{t-1}\right)\right|=O\left(\sqrt{\frac{\kappa_{u} \log (K T / \alpha)}{d}}\left\|\theta^{\star}-\hat{\theta}_{t-1}\right\|\right)
$$

To control the regret bound, we bound the estimation error $\left\|\theta^{\star}-\hat{\theta}_{t-1}\right\|$ in each time in the following lemma.
Lemma C. 1 (Estimation Error for OLS). Using the private OLS update mechanism $\psi_{t}^{O L S}$ and estimator $\varphi_{t}^{O L S}$, for any $8 \frac{d \log 9+\log (T / \alpha)}{p_{*}^{2}}<t \leq T$, we have with probability at least $1-\frac{\alpha}{T}$,

$$
\begin{equation*}
\left\|\hat{\theta}_{t}-\theta^{\star}\right\|^{2} \leq C\left(C_{B} \sigma_{\epsilon} \sigma_{\varepsilon, \delta} d\right)^{2} \frac{d+\log (T / \alpha)}{\kappa_{l}^{2} p_{*}^{2} t} \tag{9}
\end{equation*}
$$

for some $C$ independent of $d, K$ and $T$.
Lemma C. 2 (Estimation Error for SGD). Using the private OLS update mechanism $\psi_{t}^{S G D}$ and estimator $\varphi_{t}^{S G D}$, for any $3 \leq t \leq T$, we have with probability at least $1-\frac{\alpha}{T}$,

$$
\begin{equation*}
\left\|\hat{\theta}_{t}-\theta^{\star}\right\|^{2} \leq \frac{(624 \log (\log T / \alpha)+1) r_{\varepsilon, d}^{2} d^{2}}{4 \kappa_{l}^{2} \zeta^{2} p_{*}^{2} t} \tag{10}
\end{equation*}
$$

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 $1-\alpha$,

$$
\begin{align*}
& \sum_{t=1}^{T} \max _{a \in[K]}\left(x_{t, a}-x_{t, a_{t}}\right)^{T} \theta^{\star} \\
\leq & t_{1} c_{r}+\sum_{t=t_{1}+1}^{T} C C_{B} \sigma_{\epsilon} \sigma_{\varepsilon, \delta} d \sqrt{\frac{\kappa_{u} \log (K T / \alpha)}{d}} \frac{\sqrt{d+\log (T / \alpha)}}{\kappa_{l} p_{*} \sqrt{t}}  \tag{11}\\
\leq & 8 \frac{d \log 9+\log (T / \alpha)}{p_{*}^{2}}+C C_{B} \sigma_{\varepsilon, \delta} \sigma_{\epsilon} \sqrt{d} \frac{\sqrt{d+\log (T / \alpha)}}{\kappa_{l} p_{*}} \sqrt{\kappa_{u} \log (K T / \alpha)} \sqrt{T} .
\end{align*}
$$

Plugging the SGD estimation error $\sqrt{10}$ into the regret bound, we have

$$
\begin{align*}
& \sum_{t=1}^{T} \max _{a \in[K]}\left(x_{t, a}-x_{t, a_{t}}\right)^{T} \theta^{\star} \\
\leq & 2 c_{r}+\sum_{t=3}^{T} \sqrt{\kappa_{u} \log (K T / \alpha)} \frac{\sqrt{(624 \log (\log T / \alpha)+1)} r_{\varepsilon, d} \sqrt{d}}{2 \kappa_{l} \zeta p_{*} \sqrt{t}} \\
\leq & 2 c_{r}+\frac{\sqrt{(624 \log (\log T / \alpha)+1)} r_{\varepsilon, d} \sqrt{d}}{2 \zeta \kappa_{l} p_{*}} \sqrt{\kappa_{u} \log (K T / \alpha)} \sqrt{T} . \tag{12}
\end{align*}
$$

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

$$
\lambda_{\min }\left(\sum_{i=1}^{t} x_{i, a_{i}} x_{i, a_{i}}^{T}\right) \geq C \cdot \frac{t \kappa_{l} p_{*}}{d}
$$

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 $\left\{x_{i, a_{i}}\right\}_{i \in[t-1]},\left\{\epsilon_{i}\right\}_{i \in[t-1]}$ and the randomness from $\left\{\psi_{i}^{O L S}\right\}_{i \in[t-1]}$. 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 $0<s<p_{*}$,

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

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 $\varepsilon$-net of $S_{1}^{d-1}$ for some $\varepsilon>0$ to be determined,

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

which then implies

$$
\lambda_{\min }(A)=-\lambda_{\max }(-A) \geq \frac{-1}{1-2 \varepsilon} \sup _{x \in \mathcal{N}_{\varepsilon}} x^{T}(-A) x=\frac{1}{1-2 \varepsilon} \inf _{x \in \mathcal{N}_{\varepsilon}} x^{T} A x
$$

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

$$
\lambda_{\min }(A) \geq 2 \inf _{x \in \mathcal{N}_{\varepsilon}} x^{T} A x
$$

Note that

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

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

$$
\lambda_{\min }\left(\sum_{i=1}^{t} x_{i, a_{i}} x_{i, a_{i}}^{T}\right)=\min _{\|v\|=1} \sum_{i=1}^{t}\left\langle x_{i, a_{i}}, v\right\rangle^{2} \geq \frac{\kappa_{l} p_{*} t}{d}
$$

## Proof of Lemma C.1.

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

$$
\lambda_{\min }\left(\sum_{i=1}^{t} x_{i, a_{i}} x_{i, a_{i}}^{T}\right) \geq C_{1} \kappa_{l} p_{*} t / d
$$

for some $C_{1}$ independent of $d, K$ and $T$.
Since $\left\{W_{i}\right\}_{i \in[t]}$ are independent, therefore by concentration bounds for Wigner matrix we have with probability at least $1-\frac{\alpha}{T}$,

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

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 $A_{t}:=\sum_{i=1}^{t}\left(x_{i, a_{i}} x_{i, a_{i}}^{T}+W_{i}\right)+\tilde{c} \sqrt{t} I$. Therefore, by Weyl's inequality we have with probability at least $1-\frac{\alpha}{T}$,

$$
\lambda_{\min }\left(A_{t}\right)=\lambda_{\min }\left(\sum_{i=1}^{t}\left(x_{i, a t_{i}} x_{i, a_{i}}^{T}+W_{i}\right)+\tilde{c} \sqrt{t} I_{d}\right) \geq \lambda_{\min }\left(\sum_{i=1}^{t} x_{i, a_{i}} x_{i, a_{i}}^{T}\right) \geq C_{1} \kappa_{l} p_{*} t / d
$$

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

$$
\begin{aligned}
\hat{\theta}_{t}-\theta^{\star} & =A_{t}^{-1} \sum_{i=1}^{t}\left(x_{i, a_{i}} r_{i}+\xi_{i}\right)-\theta^{\star} \\
& =A_{t}^{-1} \sum_{i=1}^{t}\left(x_{i, a_{i}} x_{i, a_{i}}^{T} \theta^{\star}+x_{i, a_{i}} \epsilon_{i}+\xi_{i}\right)-\theta^{\star} \\
& =A_{t}^{-1}\left(\sum_{i=1}^{t} x_{i, a_{i}} \epsilon_{i}\right)-A_{t}^{-1} \sum_{i=1}^{t} W_{i} \theta^{\star}+A_{t}^{-1} \sum_{i=1}^{t} \xi_{i}-\tilde{c} \sqrt{t} A_{t}^{-1} \theta^{\star}
\end{aligned}
$$

Define $\mathcal{F}_{t}$ as the filtration generated by $\left\{x_{i, a_{i}}\right\}_{i \in[t]},\left\{\epsilon_{i}\right\}_{i \in[t-1]}$ and the randomness from $\left\{\psi_{i}\right\}_{i \in[t-1]}$. Notice that for every unit vector $u$,

$$
\begin{aligned}
\mathbb{E}\left[\exp \left(\lambda \sum_{i=1}^{t} u^{T} x_{i, a_{i}} \epsilon_{i}\right)\right] & =\mathbb{E}\left[\mathbb{E}\left[\exp \left(\lambda \sum_{i=1}^{t} u^{T} x_{i, a_{i}} \epsilon_{i}\right) \mid \mathcal{F}_{t}\right]\right] \\
& =\mathbb{E}\left[\prod_{i=1}^{t-1} \exp \left(\lambda u^{T} x_{i, a_{i}} \epsilon_{i}\right) \mathbb{E}\left[\exp \left(\lambda u^{T} X_{i} \epsilon_{i}\right) \mid \mathcal{F}_{t}\right]\right] \\
& \stackrel{(1)}{\leq} \exp \left(\frac{\lambda^{2} C_{B}^{2} \sigma_{\epsilon}^{2}}{2}\right) \mathbb{E}\left[\prod_{i=1}^{t-1} \exp \left(\lambda u^{T} x_{i, a_{i}} \epsilon_{i}\right)\right] \\
& \stackrel{(2)}{\leq} \exp \left(\frac{\lambda^{2} C_{B}^{2} \sigma_{\epsilon}^{2} t}{2}\right)
\end{aligned}
$$

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 $1-\frac{\alpha}{T}$,

$$
\left\|\sum_{i=1}^{t} x_{i, a_{i}} \epsilon_{i}\right\|^{2} \leq C_{3} \sigma_{\epsilon}^{2} C_{B}^{2} t(d+\log (T / \alpha))
$$

where $C_{3}$ is a positive constant independent of $d, K$ and $T$.
Therefore,

$$
\begin{align*}
\left\|A_{t}^{-1}\left(\sum_{i=1}^{t} x_{i, a_{i}} \epsilon_{i}\right)\right\|^{2} & \leq\left\|A_{t}^{-1}\right\|^{2}\left\|\left(\sum_{i=1}^{t} x_{i, a_{i}} \epsilon_{i}\right)\right\|^{2}  \tag{13}\\
& \leq \frac{C_{3} \sigma^{2} C_{B}^{2} d^{2} t(d+\log (T / \alpha))}{\left(C_{1} \kappa_{l} p_{*} t\right)^{2}}
\end{align*}
$$

Moreover,

$$
\begin{align*}
\left\|A_{t}^{-1} \sum_{i=1}^{t} W_{i} \theta^{\star}\right\|^{2} & \leq\left\|A_{t}^{-1}\right\|^{2}\left\|\sum_{i=1}^{t} W_{i}\right\|^{2}\left\|\theta^{\star}\right\|^{2} \\
& \leq\left\|A_{t}^{-1}\right\|^{2}\left\|\sum_{i=1}^{t} W_{i}\right\|^{2}  \tag{14}\\
& \leq \frac{C_{2} t \sigma_{\varepsilon, \delta}^{2}(d+\log (T / \alpha))}{\left(C_{1} \kappa_{l} p_{*} t\right)^{2}}
\end{align*}
$$

where the second inequality is from the assumption that $\left\|\theta^{\star}\right\| \leq 1$.
Third, Since $\xi_{i}$ 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 $1-\frac{\alpha}{T}$,

$$
\left\|\sum_{i=1}^{t} \xi_{i}\right\|^{2} \leq C_{4} t \sigma_{\varepsilon, \delta}^{2}(d+\log (T / \alpha))
$$

for some positive constant $C_{4}$ independent of $d, K$ and $T$. Therefore,

$$
\begin{equation*}
\left\|A_{t}^{-1} \sum_{i=1}^{t} \xi_{i}\right\|^{2} \leq \frac{C_{4} t \sigma_{\varepsilon, \delta}^{2} d^{2}(d+\log (T / \alpha))}{\left(C_{2} \kappa_{l} p_{*} t\right)^{2}} \tag{15}
\end{equation*}
$$

Lastly,

$$
\begin{equation*}
\left\|\tilde{c} \sqrt{t} A_{t}^{-1} \theta^{\star}\right\|^{2} \leq \frac{\tilde{c}^{2} t}{\left(C_{2} \kappa_{l} p_{*} t\right)^{2}} \tag{16}
\end{equation*}
$$

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

$$
\left\|\hat{\theta}_{t}-\theta^{\star}\right\|^{2} \leq C_{5} \sigma_{\epsilon}^{2} C_{B}^{2} \sigma_{\varepsilon, \delta}^{2} d^{2} \frac{d+\log (T / \alpha)}{\kappa_{l}^{2} p_{*}^{2} t}
$$

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 $\mathrm{t}, \hat{g}_{t}:=\Psi_{\varepsilon}\left[\left(\mu\left(x_{t, a_{t}}^{T} \hat{\theta}_{t}\right)-r_{t}\right) x_{t, a_{t}}\right]$ is the LDP private estimator of $g_{t}$ and $\hat{z}_{t}=g_{t}-\hat{g}_{t}$. By the unbiasedness of $\Psi_{\varepsilon}$ in Lemma 2.2 we have

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

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

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

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

$$
\left\|\hat{\theta}_{t}-\theta^{\star}\right\|^{2} \leq \frac{(624 \log (\log (T) / \alpha)+1) r_{\varepsilon, d}^{2} d^{2}}{4 \kappa_{l}^{2} \zeta^{2} p_{*}^{2} t}
$$

## C. 3 Proof of Problem-dependent Bound

To prove the problem-dependent bound, we need only combine LemmaC. 1 and Lemma C. 2 together with the following lemma.
Lemma C.4. Under the $(\beta, \gamma)$-margin condition, if we have $\left\|\hat{\theta}_{t}-\theta^{\star}\right\| \leq \frac{U_{0}}{\sqrt{t}}$ 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 $1-2 \alpha$,

$$
\operatorname{Reg}(T) \leq C \cdot \begin{cases}c_{r} t_{0}+\gamma\left(L C_{B} U_{0}\right)^{2}(\log T+o(1)), & \beta=1 \\ c_{r} t_{0}+\frac{2 \gamma}{1-\beta}\left(L C_{B} U_{0}\right)^{1+\beta}\left(T^{\frac{1-\beta}{2}}+o(1)\right), & 0 \leq \beta<1\end{cases}
$$

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

$$
\begin{aligned}
\operatorname{Reg}(T) & \leq 2 c_{r} t_{0}+\left(\mu\left(x_{t, a_{t}^{*}}^{T} \theta^{\star}\right)-\mu\left(x_{t, a_{t}}^{T} \theta^{\star}\right)\right) \mathbf{1}\left\{\left\|\hat{\theta}_{t}-\theta^{\star}\right\| \leq \frac{U_{0}}{\sqrt{t}}, \triangle_{t} \leq \frac{2 L C_{B} U_{0}}{\sqrt{t}}\right\} \\
& \leq 2 c_{r} t_{0}+2 L C_{B} \frac{U_{0}}{\sqrt{t}} \mathbf{1}\left\{\triangle_{t} \leq \frac{2 L C_{B} U_{0}}{\sqrt{t}}\right\}
\end{aligned}
$$

Denote $A_{t}:=\frac{1}{\sqrt{t}} \mathbf{1}\left\{\triangle_{t} \leq \frac{2 L C_{B} U_{0}}{\sqrt{t}}\right\}$, by Hoeffding's inequality we have with probability at least $1-\alpha$,

$$
\sum_{t} A_{t}<\sum_{t} \mathbb{E}\left[A_{t}\right]+\sqrt{\log T \log \frac{1}{\alpha}}
$$

Noting that $\mathbb{E}\left[\sum_{t} A_{t}\right] \leq 2 \gamma L C_{B} U_{0} \log T$ for $\beta=1$ and $\mathbb{E}\left[\sum_{t} A_{t}\right] \leq \frac{2 \gamma}{1-\beta}\left(L C_{B} U_{0}\right)^{\beta} T^{\frac{1-\beta}{2}}$ 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 $\theta_{i}$ 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 $A_{1} \leq C \log \left(A_{2} / \alpha\right) A_{3}$ 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 $\alpha$ 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 $K s_{0}$, the estimator $\hat{\theta}_{K s_{0}, i}$ achieves the following error bound with probability at least $1-\alpha$,

$$
\sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\| \leq h_{0}:=\frac{h_{s u b}}{8 L C_{B}}
$$

With $h=h_{\text {sub }}$ in Algorithm 2 we have $\mathbb{P}\left\{a_{t}^{*} \in \hat{K}_{t}, \hat{K}_{t} \cap K_{\text {sub }}=\emptyset\right\} \geq 1-\alpha$ holds uniformly for all $K s_{0}<t \leq T$.

Proof. Firstly, to show $a_{t}^{*} \in \hat{K}_{t}$, without loss of generality we assume that $a_{t}^{*} \neq 1$, and $\operatorname{argmax}_{i \in[K]} \mu\left(X_{t}^{T} \hat{\theta}_{K s_{0}, i}\right)=1$. Then by the optimality of $\theta_{a_{t}^{*}}$, condition on $\sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\| \leq$ $h_{0}$,

$$
\begin{aligned}
\mathbb{P}\left(a_{t}^{*} \notin \hat{K}_{t}\right) & =\mathbb{P}\left(\mu\left(X_{t}^{T} \hat{\theta}_{K s_{0}, a_{t}^{*}}\right)<\mu\left(X_{t}^{T} \hat{\theta}_{K s_{0}, 1}\right)-h / 2\right) \\
& \leq \mathbb{P}\left(\mu\left(X_{t}^{T} \theta_{a_{t}^{*}}\right)-h / 8<\mu\left(X_{t}^{T} \theta_{1}\right)+h / 8-h / 2\right)=0 .
\end{aligned}
$$

Now for any $j \in K_{\text {sub }}$, we have condition on $\sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\| \leq h_{0}$,

$$
\begin{aligned}
\mathbb{P}\left(j \in \hat{K}_{t}\right) & \leq \mathbb{P}\left(\mu\left(X_{t}^{T} \hat{\theta}_{K s_{0}, a_{t}^{*}}\right)-h / 2<\mu\left(X_{t}^{T} \hat{\theta}_{K s_{0}, j}\right)\right) \\
& \leq \mathbb{P}\left(\mu\left(X_{t}^{T} \theta_{a_{t}^{*}}\right)-3 h / 4<\mu\left(X_{t}^{T} \theta_{j}\right)+h / 4\right)=0,
\end{aligned}
$$

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 $(\beta, \gamma)$-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 $K s_{0}, \sup _{i \in[K]}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq h_{0}$, and there exists some $t_{0}, U_{0}(\alpha)$ such that with probability at least $1-\alpha$,

$$
\sup _{i \in K_{o p t}}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \frac{U_{0}(\alpha)}{\sqrt{t}}, \quad \forall t_{0} \leq t \leq T
$$

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

$$
\operatorname{Reg}(T) \leq C \cdot \begin{cases}c_{r} t_{0}+\gamma\left(L C_{B} U_{0}(\alpha)\right)^{2}(\log T+o(1)), & \beta=1 \\ c_{r} t_{0}+\frac{\gamma}{1-\beta}\left(L C_{B} U_{0}(\alpha)\right)^{1+\beta}\left(T^{\frac{1-\beta}{2}}+o(1)\right), & 0<\beta<1\end{cases}
$$

Proof of Lemma D.2. Denoting $E_{t}:=\left\{\hat{K}_{t} \cap K_{s u b}=\emptyset, a_{t}^{*} \in \hat{K}_{t}\right\}$, we have with probability at least $1-\alpha$,

$$
\begin{aligned}
\operatorname{Reg}(T) & \leq 2 c_{r} t_{0}+L \sum_{t_{0}<t \leq T} X_{t}^{T}\left(\theta_{a_{t}^{*}}-\theta_{a_{t}}\right) \\
& \leq 2 c_{r} t_{0}+L \sum_{t_{0}<t \leq T} X_{t}^{T}\left(\theta_{a_{t}^{*}}-\theta_{a_{t}}\right) \mathbf{1}\left\{\sup _{i \in K_{\text {opt }}}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \frac{U_{0}(\alpha)}{\sqrt{t}}, E_{t}\right\} \\
& \leq 2 c_{r} t_{0}+L \sum_{t_{0}<t \leq T} X_{t}^{T}\left(\theta_{a_{t}^{*}}-\theta_{a_{t}}\right) \mathbf{1}\left\{\sup _{i \in K_{o p t}}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \frac{U_{0}(\alpha)}{\sqrt{t}}, \triangle_{t} \leq \frac{2 L C_{B} U_{0}(\alpha)}{\sqrt{t}}, E_{t}\right\} \\
& \leq 2 c_{r} t_{0}+L \sum_{t_{0}<t \leq T} \frac{2 C_{B} U_{0}(\alpha)}{\sqrt{t}} \mathbf{1}\left\{\triangle_{t} \leq \frac{2 L C_{B} U_{0}(\alpha)}{\sqrt{t}}\right\} .
\end{aligned}
$$

Let $A_{t}=\mathbf{1}\left\{\triangle_{t}<\frac{2 L C_{B} U_{0}(\alpha)}{\sqrt{t}}\right\}$. Then $A_{t}$ is a sequence of independent $0-1$ valued random variable such that $\mathbb{P}\left(A_{t}=1\right) \leq \gamma\left(\frac{2 L C_{B} U_{0}(\alpha)}{\sqrt{t}}\right)^{\beta}$. Then Hoeffding's inequality implies with probability at
least $1-\alpha$,

$$
\sum_{t_{0} \leq t \leq T} \frac{1}{\sqrt{t}} A_{t} \leq \mathbb{E}\left[\sum_{1 \leq t \leq T} \frac{1}{\sqrt{t}} A_{t}\right]+\sqrt{\log T \cdot \log \left(\frac{1}{\alpha}\right)}
$$

Notice that $\mathbb{E}\left[\sum_{1 \leq t \leq T} \frac{1}{\sqrt{t}} A_{t}\right] \leq C L C_{B} \gamma U_{0}(\alpha) \log T$ when $\beta=1$ and $\mathbb{E}\left[\sum_{1 \leq t \leq T} \frac{1}{\sqrt{t}} A_{t}\right] \leq$ $C \frac{\gamma}{1-\beta}\left(L C_{B} U_{0}(\alpha)\right)^{\beta} T^{\frac{1-\beta}{2}}$ when $0<\beta<1$. This completes the proof.

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 $h_{0}=\frac{h_{\text {sub }}}{8 L C_{B}}$ and $\lambda_{0}=\left(2 L C_{B}\right)^{-1}\left(\frac{p^{\prime}}{2 \gamma}\right)^{1 / \beta}, r_{\text {opt }}:=$ $\left|K_{o p t}\right| / K$, under the $(\beta, \gamma)$-margin condition ,

$$
\begin{aligned}
s_{0} & =C K\left(\frac{C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}}{\min \left\{\lambda_{0}, h_{0}\right\} p^{\prime} \kappa_{l} r_{o p t}}\right)^{2}(d+\log (T K / \alpha)) \\
t_{0} & =2 K s_{0} \\
U_{0}(\alpha) & =\frac{K\left(C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}\right) \sqrt{d+\log (T K / \alpha)}}{\kappa_{l} p^{\prime}}
\end{aligned}
$$

satisfy the requirements in Lemma D. 2
Lemma D. 4 (Result of SGD estimator). Given $h_{0}=\frac{h_{\text {sub }}}{8 L C_{B}}$ and $\lambda_{0}=\left(2 L C_{B}\right)^{-1}\left(\frac{p^{\prime}}{2 \gamma}\right)^{1 / \beta}, r_{\text {opt }}:=$ $\left|K_{\text {opt }}\right| / K$, under the $(\beta, \gamma)$-margin condition,

$$
\begin{aligned}
s_{0} & =C\left(\frac{K r_{\varepsilon, d}}{\zeta \kappa_{l} p^{\prime} r_{o p t} \min \left\{\lambda_{0}, h_{0}\right\}}\right)^{2} \log (K T \log (K T) / \alpha), \\
t_{0} & =K s_{0}+1, \\
U_{0}(\alpha) & =C \frac{K \sqrt{\log ((K T \log K T) / \alpha)} r_{\varepsilon, d}}{\zeta \kappa_{l} p^{\prime}},
\end{aligned}
$$

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 $r_{o p t}$. That is because of our assumption $\mathbb{P}\left(\left(v^{T} X\right)^{2} 1\left\{X_{t} \in U_{i}\right\}>\kappa_{l} / K\right)>p^{\prime}$. In fact, a more natural assumption should be $\mathbb{P}\left(\left(v^{T} X\right)^{2} \mathbf{1}\left\{X_{t} \in U_{i}\right\}>\kappa_{l} /\left|K_{o p t}\right|\right)>p^{\prime}$. In that case, we have $r_{o p t}=1$, 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 $h_{0}=\frac{h_{\text {sub }}}{8 L C_{B}}, \lambda_{0}=\left(2 L C_{B}\right)^{-1}\left(\frac{p^{\prime}}{2 \gamma}\right)^{1 / \beta}, A_{t}:=\left\{\sup _{i \in K_{o p t}}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \lambda_{0}\right\}, H_{0}:=$ $\left\{\sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\| \leq h_{0}\right\}$.
Lemma D.5. Define $\mathcal{F}_{t}$ the filtration generated by $\left\{X_{i}\right\}_{i \in[t]},\left\{\epsilon_{i}\right\}_{i \in[t]}$ together with all randomness from $\left\{\psi_{i}\right\}_{i \in[t]}$. Then we have:

$$
\lambda_{\min }\left(\mathbb{E}\left[X_{t} X_{t} \mathbf{1}\left\{a_{t}=i\right\} \mid \mathcal{F}_{t-1}\right]\right) \geq \frac{p^{\prime} \kappa_{l}}{2 K} \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_{0}}, \quad \forall i \in K_{o p t}
$$

Proof. We have for every unit vector $v$

$$
\begin{aligned}
& \mathbb{E}\left[v^{T} X_{t} X_{t}^{T} v \mathbf{1}\left\{a_{t}=i\right\} \mid \mathcal{F}_{t-1}\right] \\
& \geq \mathbf{1}_{H_{0}} \frac{\kappa_{l}}{K} \mathbb{E}\left[\mathbf{1}\left\{\left|v^{T} X \mathbf{1}\left\{X_{t} \in U_{i}\right\}\right|^{2} \geq \kappa_{l} / K, a_{t}=i, A_{t-1}\right\} \mid \mathcal{F}_{t-1}\right] \\
& \geq \mathbf{1}_{H_{0}} \mathbf{1}_{A_{t-1}} \frac{\kappa_{l}}{K} \mathbb{E}\left[\mathbf{1}\left\{\left|v^{T} X \mathbf{1}\left\{X_{t} \in U_{i}\right\}\right|^{2} \geq \kappa_{l} / K\right\}-\mathbf{1}\left\{a_{t} \neq i, X_{t} \in U_{i}, A_{t-1}\right\} \mid \mathcal{F}_{t-1}\right] \\
& \geq \mathbf{1}_{H_{0}} \mathbf{1}_{A_{t-1}} \frac{\kappa_{l}}{K}\left[p^{\prime}-\mathbb{P}\left(\left\{a_{t} \neq i, X_{t}\right.\right.\right.\left.\left.\left.\in U_{i}\right\} \cap H_{0} \cap A_{t-1} \mid \mathcal{F}_{t-1}\right)\right] . \\
& \mathbb{P}\left(\left\{a_{t} \neq i, X_{t} \in U_{i}\right\} \cap H_{0} \cap A_{t-1} \mid \mathcal{F}_{t-1}\right)=\mathbf{1}_{H_{0}} \mathbf{1}_{A_{t-1}} \mathbb{P}\left(\left\{a_{t} \neq i, X_{t} \in U_{i}\right\} \cap E_{t} \cap A_{t-1} \mid \mathcal{F}_{t-1}\right) \\
& \leq \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_{0}} \mathbb{P}\left(\triangle_{t}<2 L C_{B} \lambda_{0}\right) \\
& \leq \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_{0}} \gamma\left(2 L C_{B} \lambda_{0}\right)^{\beta} \\
& \leq \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_{0}} \frac{p^{\prime}}{2},
\end{aligned}
$$

where the last inequality is by the choice of $\lambda_{0}$. 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 $\mathcal{F}_{t}^{\prime}$, let $Z_{t}:=\sum_{i=1}^{t} z^{i}$ and $\tilde{Z}_{t}:=\sum_{i=1}^{t} \mathbb{E}\left[z^{i} \mid \mathcal{F}_{i-1}^{\prime}\right]$. Suppose that $\lambda_{\max }\left(z^{i}\right) \leq R^{2}$ almost surely for all $i$, then for any $\mu$ and $\alpha \in(0,1)$,

$$
\mathbb{P}\left[\lambda_{\min }\left(Z_{t}\right) \leq(1-\alpha) \mu, \lambda_{\min }\left(\tilde{Z}_{t}\right) \geq \mu\right] \leq d\left(\frac{1}{e^{\alpha}(1-\alpha)^{1-\alpha}}\right)^{\mu / R^{2}}
$$

Now we can show the following result:
Lemma D.7. For $t_{1}<t_{2} \in \mathbb{N}$ such that $\left(t_{2}-t_{1}\right) \cdot \frac{\kappa_{l} p^{\prime}}{8 K}>10 C_{B}^{2} \log \left(d / \alpha^{\prime}\right)$, for a fixed $i \in[K]$ we have

$$
\mathbb{P}\left(\lambda_{\min }\left(\sum_{t=t_{1}}^{t_{2}} X_{t} X_{t} \mathbf{1}\left\{a_{t}=i\right\}\right) \leq \frac{t_{2}-t_{1}}{8 K} \kappa_{l} p^{\prime}, \sup _{t_{1} \leq t \leq t_{2}, i \in K_{o p t}}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \lambda_{0}, H_{0}\right) \leq \alpha^{\prime}
$$

Proof. Denote $S_{t_{1}, t_{2}}:=\cap_{t_{1} \leq t \leq t_{2}} A_{t}$, by Lemma D.5 we have

$$
\lambda_{\min }\left(\sum_{t=t_{1}}^{t_{2}} \mathbb{E}\left[X_{t} X_{t}^{T} \mathbf{1}\left\{a_{t}=i\right\} \mid \mathcal{F}_{t-1}\right]\right) \geq \sum_{t=t_{1}}^{t_{2}} \mathbf{1}_{A_{t-1}} \mathbf{1}_{H_{0}} \frac{\kappa_{l} p^{\prime}}{2 K}
$$

That implies

$$
\begin{aligned}
& \mathbb{P}\left(\lambda_{\min }\left(\sum_{t=t_{1}}^{t_{2}} X_{t} X_{t}^{T} \mathbf{1}\left\{a_{t}=i\right\}\right) \leq \frac{t_{1}-t_{2}}{4 K} \kappa_{l} p^{\prime}, S_{t_{1}, t_{2}}, H_{0}\right) \\
\leq & \left.\mathbb{P}\left(\lambda_{\min }\left(\sum_{t=t_{1}}^{t_{2}} X_{t} X_{t}^{T} \mathbf{1}\left\{a_{t}=i\right\}\right) \leq \frac{t_{1}-t_{2}}{4 K} \kappa_{l} p^{\prime}, \mathbb{E}\left[X_{t} X_{t}^{T} \mathbf{1}\left\{a_{t}=i\right\} \mid \mathcal{F}_{t-1}\right]\right) \geq\left(t_{2}-t_{1}\right) \frac{\kappa_{l} p^{\prime}}{2 K}\right)
\end{aligned}
$$

Then selecting $\alpha=1 / 2$ and $\mu=\left(t_{2}-t_{1}\right) \cdot \frac{\kappa_{l} p^{\prime}}{4 K}$ in Lemma D.6, we have

$$
\mathbb{P}\left(\lambda_{\min }\left(\sum_{t=t_{1}}^{t_{2}} X_{t} X_{t}^{T} \mathbf{1}\left\{a_{t}=i\right\}\right) \leq\left(t_{2}-t_{1}\right) \frac{\kappa_{l} p^{\prime}}{8 K}, S_{t_{1}, t_{2}}, H_{0}\right) \leq d\left(\frac{1}{\sqrt{e / 2}}\right)^{10 \log \left(\frac{d}{\alpha^{\prime}}\right)} \leq \alpha^{\prime}
$$

That leads to the claim.

In warm up stage, we have the following lemma.
Lemma D.8. As long as $s_{0} \geq C\left(r_{\text {opt }} \kappa_{l} p^{\prime}\right)^{-2} \max \left\{\log \frac{1}{\alpha}\right.$, $\left.d\right\}$ for some absolute constant $C$, we have with probability at least $1-\alpha$,

$$
\lambda_{\min }\left(\sum_{t=1}^{K s_{0}} \mathbf{1}\left\{a_{t}=i\right\} X_{t} X_{t}^{T}\right)^{-1} \leq \frac{2}{s_{0} p^{\prime} r_{o p t} \kappa_{l}}, \quad \forall i \in[K] .
$$

Proof. Since $X_{t}$ are i.i.d. for $(i-1) s_{0}+1 \leq t \leq i s_{0}$, using classical concentration results for i.i.d. sub-gaussian covariance matrix result (e.g. Theorem 6.5 in [42] ), we have when $s_{0}>$ $C\left(r_{o p t} \kappa_{l} p^{\prime}\right)^{-2} \max \left\{\log \frac{1}{\alpha}, d\right\}$, with probability at least $1-\alpha$,

$$
\begin{aligned}
\left\|\frac{1}{s_{0}} \sum_{t=1}^{K s_{0}} \mathbf{1}\left\{a_{t}=i\right\} X_{t} X_{t}^{T}-\mathbb{E}\left[X_{1} X_{1}^{T}\right]\right\| & \leq c_{1}\left(\sqrt{\frac{d}{s_{0}}}+\frac{d}{s_{0}}\right)+c_{2} \max \left\{\sqrt{\frac{\log 1 / \alpha}{s_{0}}}, \frac{\log 1 / \alpha}{s_{0}}\right\} \\
& \leq c_{3}\left(\sqrt{\frac{d}{s_{0}}}+\sqrt{\frac{\log (1 / \alpha)}{s_{0}}}\right) \\
& \leq r_{o p t} p^{\prime} \kappa_{l} / 2
\end{aligned}
$$

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

$$
\lambda_{\min } \mathbb{E}\left[X_{1} X_{1}^{T}\right] \geq \sum_{i \in K_{o p t}} \lambda_{\min } \mathbb{E}\left[X_{1} X_{1}^{T} \mathbf{1}\left\{X_{1} \in U_{i}\right\}\right] \geq r_{o p t} \kappa_{l} p^{\prime}
$$

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

$$
\lambda_{\min }\left(\sum_{t=1}^{K s_{0}} X_{t} X_{t}^{T}\right) \geq s_{0} r_{o p t} p^{\prime} \kappa_{l} / 2
$$

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 $K s_{0}$, we have for any $\alpha>0$, with probability at least $1-\alpha$,

$$
\sup _{i \in[K]}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \frac{\left(4 C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}\right) \sqrt{t\left(\log \left(\frac{T K}{\alpha}\right)+d\right)}}{s_{0} p^{\prime} r_{o p t} \kappa_{l}} \quad \text { holds for all } K s_{0} \leq t \leq T
$$

Proof. Denote $U_{t}=\sum_{s=1}^{t}\left(\mathbf{1}\left\{a_{s}=i\right\} X_{s} X_{s}^{T}+\left(\mathbf{1}\left\{a_{s}=i, s \leq K s_{0}\right\}+\mathbf{1}\left\{s>K s_{0}\right\}\right) W_{s}\right)+\tilde{c} \sqrt{t} I_{d}$, we have

$$
\begin{aligned}
\hat{\theta}_{t, i} & =U_{t}^{-1}\left(\sum_{s=1}^{t} \mathbf{1}\left\{a_{s}=i\right\} X_{s} y_{s}+\left(\mathbf{1}\left\{a_{s}=i, s \leq K s_{0}\right\}+\mathbf{1}\left\{s>K s_{0}\right\}\right) \xi_{s}\right) \\
& =U_{t}^{-1}\left(\sum_{s=1}^{t} \mathbf{1}\left\{a_{s}=i\right\}\left[X_{s} X_{s}^{T} \theta_{i}+X_{s} \epsilon_{s}\right]+\left(\mathbf{1}\left\{a_{s}=i, s \leq K s_{0}\right\}+\mathbf{1}\left\{s>K s_{0}\right\}\right) \xi_{s}\right) \\
& =\theta_{i}+U_{t}^{-1}\left(\sum_{s=1}^{t}\left(\mathbf{1}\left\{a_{s}=i\right\} X_{s} \epsilon_{s}+\left(\mathbf{1}\left\{a_{s}=i, s \leq K s_{0}\right\}+\mathbf{1}\left\{s>K s_{0}\right\}\right)\left(\xi_{s}-W_{s} \theta_{i}\right)\right)-\tilde{c} \sqrt{t} I_{d} \theta_{i}\right) .
\end{aligned}
$$

By $\left\|\sum_{s=1}^{K s_{0}} \mathbf{1}\left\{a_{s}=i\right\} W_{s}+\sum_{s=K s_{0}+1}^{t} W_{s}\right\| \leq \tilde{c} \sqrt{t}, \forall K s_{0} \leq t \leq T, i \in[K]$ with probability at least $1-\alpha$, we have with probability at least $1-2 \alpha$,

$$
\left[\lambda_{\min }(U)\right]^{-1} \leq \lambda_{\min }\left(\sum_{s=1}^{K s_{0}} \mathbf{1}\left\{a_{s}=i\right\} X_{s} X_{s}^{T}\right)^{-1} \leq \frac{2}{s_{0} r_{o p t} p^{\prime} \kappa_{l}}, \quad \forall K s_{0} \leq t \leq T
$$

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

$$
\begin{align*}
& \left\|\sum_{s=1}^{t} \mathbf{1}\left\{a_{s}=i\right\} X_{s} \epsilon_{s}\right\| \leq C C_{B} \sigma_{\epsilon} \sqrt{t(d+\log (T K / \alpha))}  \tag{17}\\
& \left\|\sum_{s=1}^{t}\left(\mathbf{1}\left\{a_{s}=i, s \leq K s_{0}\right\}+\mathbf{1}\left\{s>K s_{0}\right\}\right) \xi_{s}\right\| \leq C \sigma_{\varepsilon, \delta} \sqrt{t(d+\log (T K / \alpha))},  \tag{18}\\
& \left\|\sum_{s=1}^{K s_{0}} \mathbf{1}\left\{a_{s}=i\right\} W_{s} \theta_{i}+\sum_{s=K s_{0}+1}^{t} W_{s} \theta_{i}\right\| \leq \tilde{c} \sqrt{t}\left\|\theta_{i}\right\| \leq C \sigma_{\varepsilon, \delta} \sqrt{t(d+\log (T K / \alpha))} . \tag{19}
\end{align*}
$$

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

$$
\sup _{i \in[K]}\left\|\hat{\theta}_{t, i}-\theta_{i}\right\| \leq \frac{2 C}{s_{0} p^{\prime} r_{o p t} \kappa_{l}}\left(C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}\right) \sqrt{t(\log (T K / \alpha)+d)} .
$$

That finishes the proof.
Lemma D.10. As long as

$$
s_{0} \geq C K\left(\frac{C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}}{\min \left\{\lambda_{0}, h_{0}\right\} r_{o p t} p^{\prime} \kappa_{l}}\right)^{2}(d+\log (T K / \alpha)),
$$

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

$$
\begin{align*}
& \sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\|_{2} \leq \min \left\{\lambda_{0}, h_{0}\right\},  \tag{20}\\
& \sup _{i \in K_{o p t}}\left\|\hat{\theta}_{s, i}-\theta_{i}\right\|_{2} \leq \lambda_{0} \text { holds uniformly for } K s_{0} \leq s \leq(K+1) s_{0}  \tag{21}\\
& C \frac{K\left(C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}\right) \sqrt{d+\log (T K / \alpha)}}{\sqrt{t-K s_{0}} \kappa_{l} p^{\prime}} \leq \lambda_{0} \text { holds for all } t \geq 2 K s_{0} \tag{22}
\end{align*}
$$

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 $t>K s_{0}$, denote $t^{\prime}=t-K s_{0}$ and $\tilde{t}_{0}=2 K s_{0}$, we have if

$$
H_{0} \text { holds and }\left\|\hat{\theta}_{t, i}-\theta_{i}\right\|_{2} \leq \min \left\{\tilde{U}_{s}(\alpha), \lambda_{0}\right\} \text { holds uniformly for } i \in K_{o p t}, \tilde{t}_{0} \leq s \leq t
$$

with probability at least $1-\sum_{j=1}^{t^{\prime}} \frac{2}{j^{2}} \alpha$, then
$H_{0}$ holds and $\left\|\hat{\theta}_{t, i}-\theta_{i}\right\|_{2} \leq \min \left\{\tilde{U}_{s}(\alpha), \lambda_{0}\right\}$ holds uniformly for $i \in K_{o p t}, \tilde{t}_{0} \leq s \leq t+1$, with probability at least $1-\sum_{j=1}^{t^{\prime}+1} \frac{2}{j^{2}} \alpha$, where

$$
\tilde{U}_{s}(\alpha)=C \frac{K\left(C_{B} \sigma_{\epsilon}+\sigma_{\varepsilon, \delta}\right) \sqrt{d+\log (T K / \alpha)}}{\sqrt{s} \kappa_{l} p^{\prime}}
$$

Proof. Denote $S_{\tilde{t}_{0}, t}=\left\{\left\|\hat{\theta}_{s, i}-\theta_{i}\right\| \leq \min \left\{\tilde{U}_{s}(\alpha), \lambda_{0}\right\}, \forall K \in K_{o p t}, \forall \tilde{t}_{0} \leq s \leq t\right\}, \tilde{A}_{t}=$ $\left\{\sup _{i \in K_{o p t}}\left\|\hat{\theta}_{i, t}-\theta_{i}\right\| \leq \tilde{U}_{t}(\alpha)\right\}$, we have by Lemma D. 7

$$
\mathbb{P}\left(S_{\tilde{t}_{0}, t}, H_{0}, \lambda_{\min }\left(\sum_{s=1}^{t} X_{s} X_{s} \mathbf{1}\left\{a_{s}=i\right\}\right)>\frac{t^{\prime} \kappa_{l} p^{\prime}}{8 K}\right) \geq 1-\frac{\alpha}{2 K T^{2}}
$$

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

$$
\begin{aligned}
\mathbb{P}\left(H_{0}, S_{\tilde{t}_{0}, t}, A_{t+1}\right) & \geq 1-\sum_{j=1}^{t^{\prime}} \frac{2}{j^{2}} \alpha-\frac{3 \alpha}{2 T^{2}}-\sum_{i \in K_{o p t}} \mathbb{P}\left(H_{0}, S_{\tilde{t}_{0}, t}, \tilde{A}_{t+1}, \lambda_{\min }\left(\sum_{s=1}^{t} X_{s} X_{s} \mathbf{1}\left\{a_{s}=i\right\}\right) \leq \frac{t^{\prime} \kappa_{l} p^{\prime}}{4}\right) \\
& \geq 1-\sum_{j=1}^{t^{\prime}} \frac{2}{j^{2}} \alpha-\frac{2 \alpha}{T^{2}} \\
& \geq 1-2 \sum_{j=1}^{t^{\prime}+1} \frac{1}{j^{2}} \alpha
\end{aligned}
$$

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, $\mathbb{P}\left(H_{0}, S_{\tilde{t}_{0}, t+1}\right)=$ $\mathbb{P}\left(H_{0}, S_{\tilde{t}_{0}, t}, \tilde{A}_{t+1}\right)$. Thus the claim holds.

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

## D. 3 Proof of Lemma D. 4

Proof. For the estimator $\hat{\theta}_{K s_{0}, i}$ at the end of warm up stage, since the action is independent of the contexts, every $\hat{\theta}_{K s_{0}, i}$ 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 $\hat{\theta}_{K s_{0}, 1}$ (notice that by the sampling strategy in the warm up stage, we have $\hat{\theta}_{K s_{0}, 1}=\hat{\theta}_{s_{0}, 1}$ ). The result for other $\hat{\theta}_{K s_{0}, i}$ can be established using the same argument. For $2 \leq t \leq s_{0}$,

$$
\begin{aligned}
\left\|\hat{\theta}_{t, i}-\theta_{i}\right\|^{2} & =\left\|\hat{\theta}_{t-1, i}-\eta_{t} \hat{g}_{t}-\theta_{i}\right\|^{2} \\
& =\left\|\hat{\theta}_{t-1, i}-\theta_{i}\right\|^{2}-2 \eta_{t} \hat{g}_{t}^{T}\left(\hat{\theta}_{t-1, i}-\theta_{i}\right)+2 \eta_{t}^{2}\left\|\hat{g}_{t}\right\|^{2}
\end{aligned}
$$

Here $\hat{g}_{t}:=\Psi_{\varepsilon}\left[\left(\mu\left(X_{t}^{T} \hat{\theta}_{t, i}\right)-r_{t}\right) X_{t}\right]$, by the unbiasedness of $\Psi_{\varepsilon}$ in Lemma 2.2 we have

$$
\begin{aligned}
& \mathbb{E}\left[\Psi_{\varepsilon}\left(\left(\mu\left(X_{t}^{T} \hat{\theta}_{t-1, i}\right)-r_{t}\right) X_{t}\right)^{T}\left(\hat{\theta}_{t-1, i}-\theta_{i}\right) \mid \mathcal{F}_{t-1}\right] \\
& =\mathbb{E}\left[\left(\mu\left(X_{t}^{T} \hat{\theta}_{t-1, i}\right)-\mu\left(X_{t}^{T} \theta_{i}\right)\right) X_{t}^{T}\left(\hat{\theta}_{t-1, i}-\theta_{i}\right) \mid \mathcal{F}_{t-1}\right] \\
& \geq \zeta \mathbb{E}\left[\left[X_{t}^{T}\left(\hat{\theta}_{t-1, i}-\theta_{i}\right)\right]^{2} \mid \mathcal{F}_{t-1}\right] \\
& \geq \zeta \kappa_{l} r_{o p t} p^{\prime}\left\|\hat{\theta}_{t-1, i}-\theta_{i}\right\|^{2} .
\end{aligned}
$$

We get

$$
\left\|\hat{\theta}_{t, i}-\theta_{i}\right\|^{2} \leq\left(1-2 \zeta r_{o p t} \kappa_{l} p^{\prime} \eta_{t}\right)\left\|\hat{\theta}_{t-1, i}-\theta_{i}\right\|^{2}+2 \eta_{t}\left(\mathbb{E}\left[\hat{g}_{t} \mid \mathcal{F}_{t-1}\right]-\hat{g}_{t}\right)^{T}\left(\hat{\theta}_{t-1, i}-\theta_{i}\right)+2 \eta_{t}^{2}\left\|\hat{g}_{t}\right\|^{2}
$$

Notice $\left\|\hat{g}_{t}\right\|_{2}^{2}$ 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 $\eta_{t}=1 /\left(r_{o p t} \zeta \kappa_{l} p^{\prime} t\right)$ in the warm up stage, then with probability at least $1-\alpha$,

$$
\begin{equation*}
\sup _{i \in[K]}\left\|\hat{\theta}_{K s_{0}, i}-\theta_{i}\right\|^{2} \leq C \frac{(\log (\log (K T) / \delta)+1) r_{\varepsilon, \delta}^{2}}{\zeta^{2} \kappa_{l}^{2} r_{o p t}^{2} p^{\prime 2} s_{0}} \tag{23}
\end{equation*}
$$

Notice that in our algorithm, when $t>K s_{0}$, for any $i \in K_{o p t}$, the private gradient descent formula is given by

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

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

$$
\begin{aligned}
\mathbb{E}\left[\tilde{g}^{T}\left(\hat{\theta}_{t-1,1}-\theta_{1}\right) \mid \mathcal{F}_{t-1}\right] & =\mathbb{E}\left[\mathbf{1}\left\{a_{t}=i\right\} \hat{g}^{T}\left(\hat{\theta}_{t-1,1}-\theta_{1}\right) \mid \mathcal{F}_{t-1}\right] \\
& =\mathbb{E}\left[\mathbf{1}\left\{a_{t}=i\right\}\left(\mu\left(X_{t}^{T} \hat{\theta}_{t-1,1}\right)-\mu\left(X_{t}^{T} \theta_{i}\right)\right) X_{t}^{T}\left(\hat{\theta}_{t-1,1}-\theta_{1}\right) \mid \mathcal{F}_{t-1}\right] \\
& \geq \zeta \mathbb{E}\left[\mathbf{1}\left\{a_{t}=i\right\}\left[X_{t}^{T}\left(\hat{\theta}_{t-1,1}-\theta_{1}\right)\right]^{2} \mid \mathcal{F}_{t-1}\right] \\
& \geq \mathbf{1}_{A_{t}, H_{0}} \zeta \kappa_{l} p^{\prime} \eta_{t}\left\|\hat{\theta}_{t-1,1}-\theta_{1}\right\|^{2} / K
\end{aligned}
$$

select $\eta_{t}:=K /\left(\zeta \kappa_{l} p^{\prime} t^{\prime}\right)$, with $t^{\prime}=t-(K-1) s_{0}$ we have then

$$
\left\|\hat{\theta}_{t, 1}-\theta_{1}\right\|_{2}^{2} \leq\left(1-\frac{2}{t^{\prime}} \mathbf{1}_{A_{t}, H_{0}}\right)\left\|\hat{\theta}_{t-1,1}-\theta_{1}\right\|^{2}+\frac{2 K}{\zeta \kappa_{l} p^{\prime} t^{\prime}}\left(\mathbb{E}\left[\tilde{g}_{t} \mid \mathcal{F}_{t-1}\right]-\tilde{g}_{t}\right)^{T}\left(\hat{\theta}_{t-1,1}-\theta_{1}\right)+2\left(\frac{K r_{\varepsilon, d}}{\zeta \kappa_{l} p^{\prime} t^{\prime}}\right)^{2}
$$

If we denote $S_{t}:=\cap_{s=K s_{0}}^{t} A_{s}$, then using the above inequality recursively until $t=K s_{0}+1$ (i.e. until $t^{\prime}=s_{0}+1$ ), we have

$$
\begin{aligned}
\mathbf{1}_{S_{t-1}, H_{0}}\left\|\hat{\theta}_{t, 1}-\theta_{1}\right\|^{2} & \leq \frac{s_{0}\left(s_{0}-1\right)}{t^{\prime}\left(t^{\prime}-1\right)}\left\|\hat{\theta}_{K s_{0}, 1}-\theta_{1}\right\|^{2}+2\left(\frac{K r_{\varepsilon, d}}{\zeta \kappa_{l} p^{\prime} t^{\prime}}\right)^{2} \\
& +\frac{2 K}{\left(t^{\prime}-1\right) t^{\prime} \zeta \kappa_{l} p^{\prime}} \sum_{s=K s_{0}+1}^{t}\left(\mathbb{E}\left[\tilde{g}_{s} \mid \mathcal{F}_{t-1}\right]-\tilde{g}_{s}\right)^{T}\left(\hat{\theta}_{s-1,1}-\theta_{1}\right)
\end{aligned}
$$

Then it follows from the same proof as in Proposition 1 in [28] that for any fixed $K s_{0}<t \leq T$, we have with probability at least $1-\alpha / T$,

$$
\begin{equation*}
\mathbf{1}_{S_{t-1}, H_{0}}\left\|\hat{\theta}_{t, 1}-\theta_{1}\right\|^{2} \leq \frac{s_{0}\left(s_{0}-1\right)}{t^{\prime}\left(t^{\prime}-1\right)}\left\|\hat{\theta}_{K s_{0}, 1}-\theta_{1}\right\|^{2}+C \frac{K^{2}(\log (T K \log (T K) / \alpha)+1) r_{\varepsilon, d}^{2}}{\zeta^{2} \kappa_{l}^{2} p^{\prime 2} t^{\prime}} \tag{24}
\end{equation*}
$$

Now choose $s_{0} \geq 2 C \frac{K^{2}(\log (T K \log (T K) / \alpha)+1) r_{\varepsilon, d}^{2}}{r_{o p t}^{2} \zeta^{2} \kappa_{l}^{2} p^{2} \min \left\{\lambda_{0}, h_{0}\right\}^{2}}$, so that the second term in (24) is less or equal to $\lambda_{0} / 2$, we have $\mathbb{P}\left(S_{K s_{0}+1}, H_{0}\right) \geq 1-2 \alpha$ by (23). And by calling (24) recursively we can get $\mathbb{P}\left(S_{t-1}, H_{0}\right)>1-2 \alpha-\frac{t-K s_{0}}{T} \alpha \geq 1-3 \alpha, \forall K s_{0}<t \leq T$. Then with probability at least $1-3 \alpha$, we have

$$
\left\|\hat{\theta}_{t, 1}-\theta_{1}\right\|^{2} \leq C \frac{K^{2}(\log (3 T K \log (T K) / \alpha)) r_{\varepsilon, d}^{2}}{\zeta^{2} \kappa_{l}^{2} p^{2}\left(t-(K-1) s_{0}\right)}, \quad \forall K s_{0}<t \leq T
$$

The above inequality is because the term $\frac{s_{0}\left(s_{0}-1\right)}{t^{\prime}\left(t^{\prime}-1\right)}\left\|\hat{\theta}_{K s_{0}, 1}-\theta_{1}\right\|^{2} \leq \frac{s_{0}\left(s_{0}-1\right)}{t^{\prime}\left(t^{\prime}-1\right)} \frac{\min \left\{\lambda_{0}, h_{0}\right\}}{2}$, 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 $2, x_{t, i} \sim \mathcal{N}\left(0, \frac{1}{d} I_{d}\right)$ 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 $\varepsilon$-LDP bandit algorithm $\pi$ 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 $\left(x_{1, a_{1}}, y_{1}, x_{2, a_{2}}, y_{2}, \ldots, x_{t-1, a_{t-1}}, y_{t-1}\right)$. Since the algorithm is under the $\varepsilon$-LDP constraint, each $a_{t}$ can only access $S_{t}:=\left(M_{1}\left(x_{1, a_{1}}, y_{1}\right), M_{2}\left(x_{2, a_{2}}, y_{2}\right), \ldots, M_{t-1}\left(x_{t-1, a_{t-1}}, y_{t-1}\right)\right)$ with $M_{1}, \ldots, M_{t-1}$ a sequence of $\varepsilon$-LDP mechanisms. We denote the distribution of $S_{t}$ by $Q_{\theta}^{t}$, and we have

$$
\begin{align*}
& \mathbb{E}_{\theta \sim Q_{0}}\left[\mathbb{E}_{Q_{\theta}}^{t}\left[\left(x_{t, a_{t}^{*}}-x_{t, a_{t}}\right)^{T} \theta \mid x_{t, 1}, x_{t, 2}\right]\right] \\
= & \mathbb{E}_{\theta \sim Q_{0}}\left[\left(\left(x_{t, 1}-x_{t, 2}\right)^{T} \theta\right)+Q_{\theta}^{t}\left(a_{t}\left(S_{t}, x_{t}\right)=2\right)+\left(\left(x_{t, 2}-x_{t, 1}\right)^{T} \theta\right)_{+} Q_{\theta}^{t}\left(a_{t}\left(S_{t}, x_{t}\right)=1\right)\right], \tag{25}
\end{align*}
$$

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

$$
\frac{d Q_{1}}{d Q_{0}}:=\frac{\left(\left(x_{t, 1}-x_{t, 2}\right)^{T} \theta\right)_{+}}{Z_{0}}, \quad \frac{d Q_{2}}{d Q_{0}}:=\frac{\left(\left(x_{t, 2}-x_{t, 1}\right)^{T} \theta\right)_{+}}{Z_{0}}
$$

where $Z_{0}=\mathbb{E}_{Q_{0}}\left[\left(\left(x_{t, 1}-x_{t, 2}\right)^{T} \theta\right)_{+}\right]=\mathbb{E}_{Q_{0}}\left[\left(\left(x_{t, 2}-x_{t, 1}\right)^{T} \theta\right)_{+}\right]$is the normalization factor. Denote $r_{t}=\left\|x_{t, 1}-x_{t, 2}\right\|, u_{t}=r_{t}^{-1}\left(x_{t, 1}-x_{t, 2}\right)$, then the right hand side of 25 is lower bounded by

$$
\begin{align*}
& =Z_{0}\left(Q_{1} \circ Q_{\theta}^{t}\left(a_{t}\left(S_{t}, x_{t}\right)=2\right)+Q_{2} \circ Q_{\theta}^{t}\left(a_{t}\left(S_{t}, x_{t}\right)=1\right)\right) \\
& \geq_{(a)} Z_{0}\left(1-\operatorname{TV}\left(Q_{1} \circ Q_{\theta}^{t}, Q_{2} \circ Q_{\theta}^{t}\right)\right) \\
& \geq_{(b)} \frac{Z_{0}}{2} \exp \left(-D_{K L}\left(Q_{1} \circ Q_{\theta}^{t} \| Q_{2} \circ Q_{\theta}^{t}\right)\right) \\
& ={ }_{(c)} \frac{Z_{0}}{2} \exp \left(-D_{K L}\left(Q_{1} \circ Q_{\theta}^{t} \| Q_{1} \circ Q_{\theta-2\left(u_{t}^{t} \theta\right) u_{t}}^{t}\right)\right) \\
& \geq_{(d)} \frac{Z_{0}}{2} \exp \left(-\mathbb{E}_{Q_{1}}\left[D_{K L}\left(Q_{\theta}^{t} \| Q_{\theta-2\left(u_{t}^{T} \theta\right) u_{t}}^{t}\right)\right]\right) \tag{F.1}
\end{align*}
$$

where $D_{K L}(\cdot \| \cdot)$ denote the KL-divergence, $T V(\cdot, \cdot)$ denote the total variation distance and $Q_{i} \circ Q_{\theta}^{t}$ means $\mathbb{E}_{\theta \sim Q_{i}}\left[Q_{\theta}^{t}\right]$. 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

$$
\begin{aligned}
D_{K L}\left(Q_{\theta}^{t} \| Q_{\theta-2\left(u_{t}^{T} \theta\right) u_{t}}^{t}\right) & =\sum_{s=1}^{t-1} \mathbb{E}_{Q_{\theta}^{s-1}}\left[D_{K L}\left(P_{\theta}^{t}\left(\cdot \mid S_{s-1}\right) \| P_{\theta-2\left(u_{t}^{T} \theta\right) u_{t}}^{t}\left(\cdot \mid S_{s-1}\right)\right)\right] \\
& \leq \sum_{s=1}^{t-1} \frac{c}{2}\left(e^{\varepsilon}-1\right)^{2}\left(2\left(u_{t}^{T} \theta\right)^{2}\left\|u_{t}\right\|^{2}\right)
\end{aligned}
$$

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

$$
\frac{r_{t} \triangle}{C \sqrt{d}} \exp \left(-C \frac{\left(e^{\varepsilon}-1\right)^{2} \triangle^{2}}{d+1} u_{t}^{T}\left(\sum_{s=1}^{t-1} x_{t, a_{t}} x_{t, a_{t}}^{T}\right) u_{t}\right)
$$

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

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

Selecting $\triangle \asymp \frac{d}{\left(e^{\varepsilon}-1\right) \sqrt{t}}$ and taking summation over $1 \leq t \leq T$ leads to $\Omega\left(\sqrt{T d} /\left(e^{\varepsilon}-1\right)\right)$ lower bound, finally noticing $e^{\varepsilon}-1 \leq C \varepsilon$ for $\varepsilon \leq 1$ leads to the desired lower bound when $\varepsilon \leq 1$.

## 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 $\eta_{t}=c_{1} d /\left(\kappa_{l} \zeta p_{*} t\right)$ to $\eta_{t}=c_{2} d /\left(\kappa_{l} \zeta p_{*} \sqrt{t}\right)$ 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.,

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

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 $v=(x, p x)^{T} \theta^{\star}$. 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 $\theta^{\star}$. 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 $\left\{p_{i}\right\}_{i \in[25]}$. 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.

