Supplementary Materials

The following content was not necessarily subject to peer review.

Throughout the appendix, for a matrix $\boldsymbol{A}$, we shall define $\lambda_{max}\left(\boldsymbol{A}\right)$ and $\lambda_{min}\left(\boldsymbol{A}\right)$ as the maximum and minimum eigenvalue of $\boldsymbol{A}$ respectively. Further, the norm of a matrix $\boldsymbol{A}$ is defined as $\|\boldsymbol{A}\|_2^2 = \lambda_{max}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)$.

Without loss of generality, we also assume that $\kappa, K, d, R, S,$ and T are greater than 1 throughout the appendix.

7 Batched Multinomial Contextual Bandit Algorithm: B-MNL-CB

7.1 Notations

We first list a few matrices, vectors, and scalars that are commonly used throughout this section:

1. $\boldsymbol{V}_\beta = \lambda \boldsymbol{I}_{d\times d} + \sum\limits_{t\in\mathcal{T}_\beta} \boldsymbol{x}_t\boldsymbol{x}_t^\top$
2. $\tilde{\boldsymbol{V}}_\beta = \boldsymbol{I}_{K\times K} \otimes \boldsymbol{V}_\beta$
3. $\boldsymbol{A}(\boldsymbol{x},\boldsymbol{\theta}) = \mathrm{diag}(\boldsymbol{z}(\boldsymbol{x},\boldsymbol{\theta})) - \boldsymbol{z}(\boldsymbol{x},\boldsymbol{\theta})\boldsymbol{z}(\boldsymbol{x},\boldsymbol{\theta})^\top$
4. $\boldsymbol{M}(\boldsymbol{x},\boldsymbol{\theta}_1,\boldsymbol{\theta}_2) = \int\limits_0^1 \boldsymbol{A}(\boldsymbol{x}, v\boldsymbol{\theta}_1 + (1-v)\boldsymbol{\theta}_2)\ \mathrm{d}v$
5. $\boldsymbol{H}^\star_\beta := \lambda\boldsymbol{I}_{Kd\times Kd} + \sum\limits_{t\in\mathcal{T}_\beta} \boldsymbol{A}(\boldsymbol{x}_t,\boldsymbol{\theta}^\star)\otimes\boldsymbol{x}_t\boldsymbol{x}_t^\top$
6. $\gamma(\lambda) = 12S\sqrt{\log T + Kd} + 8S\lambda^{-1/2}(\log T + Kd) + 2S^{3/2}\lambda^{1/2}$
7. $B_\beta(\boldsymbol{x}) = \exp\left(\sqrt{6}\min\left\{\kappa^{1/2}\gamma(\delta)\left|\left|\boldsymbol{x}\right|\right|_{\boldsymbol{V}_\beta^{-1}}, 2S\right\}\right)$
8. $\boldsymbol{H}_\beta = \lambda\boldsymbol{I}_{Kd\times Kd} + \sum\limits_{t\in\mathcal{T}_\beta} \frac{\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_\beta)}{\boldsymbol{B}_\beta(\boldsymbol{x}_t)}\otimes\boldsymbol{x}_t\boldsymbol{x}_t^\top$
9. $\tilde{\boldsymbol{X}}_\beta = \frac{\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)^{1/2}}{\sqrt{B_\beta(\boldsymbol{x})}}\otimes\boldsymbol{x}$
10. $\tilde{\boldsymbol{x}}_\beta^{(i)} = \frac{\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)^{1/2}}{\sqrt{B_\beta(\boldsymbol{x})}}\boldsymbol{e}_i\otimes\boldsymbol{x}$
11. $\boldsymbol{m}_s = \left(\mathbb{1}\{y_s = 1\},\ldots,\mathbb{1}\{y_s = K\}\right)^\top$

We now present the regret upper bound for B-MNL-CB by restating Theorem 3.1:

Theorem 7.1. *(Regret of B-MNL-CB) With high probability, at the end of T rounds, the regret incurred by Algorithm 1 is bounded above by R_T where*

$$R_T \leq \tilde{\mathcal{O}}\left(RS^{5/4}K^{5/2}d\sqrt{T} + RS^{5/2}K^2d^2\kappa^{1/2}T^{1/4}\max\{e^{3S}K^{3/2}S^{-1}, \kappa^{1/2}d\}\right)$$

Proof. From Lemma 7.10, we have an upper bound for the regret incurred for any round $t\in\mathcal{T}_{\beta+1}$. Thus, the regret incurred in batch $\beta+1$ is given by:

$$R_{\beta+1} \leq 16RK^2\gamma(\lambda)\sqrt{d\log(Kd)}\left(\frac{\tau_{\beta+1}}{\sqrt{\tau_\beta}}\right) + 32RK\kappa^{1/2}d\gamma^2(\lambda)\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\}\left(\frac{\tau_{\beta+1}}{\tau_\beta}\right)$$

Choosing the batch lengths as $\tau_\beta = T^{1-2^{-\beta}}$ results in the following observation (Hanna et al., 2023; Gao et al., 2019):

$$\frac{\tau_{\beta+1}}{\sqrt{\tau_\beta}} \leq 2\sqrt{T} \qquad\qquad \frac{\tau_{\beta+1}}{\tau_\beta} \leq T^{\frac{1}{4}}$$

Thus, the regret incurred in batch $\beta + 1$ is bounded by:

$$R_{\beta+1} \leq 32RK^2\gamma(\lambda)\sqrt{d\log(Kd)}\sqrt{T} + 32RK\kappa^{1/2}d\gamma^2(\lambda)\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\}T^{1/4}$$

We now trivially upper bound the regret for $\mathcal{T}_1$ as $R\tau_1 = R\sqrt{T}$. Thus, adding the regret incurred in each batch over all batches $\beta \in [1, \log\log T + 1]$ results in:

$$\begin{aligned} R_T \leq &\left(32RK^2\gamma(\lambda)\sqrt{d\log(Kd)} + R\right)\sqrt{T}\log\log T \\ &+ 32RK\kappa^{1/2}d\gamma^2(\lambda)\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\}T^{1/4}\log\log T \end{aligned}$$

From Lemma 7.1, setting $\lambda = S^{-1/2}Kd\log T$ along with the fact that $Kd + \log T \leq Kd\log T$ results in $\gamma(\lambda) \leq 22S^{5/4}\sqrt{Kd\log T}$. Substituting the value of $\gamma(\lambda)$ gives us:

$$\begin{aligned} R_T \leq &\left(704S^{5/4}RK^{5/2}d\sqrt{\log T\log(Kd)} + R\right)\sqrt{T}\log\log T \\ &+ 14784RS^{5/2}K^2d^2\kappa^{1/2}\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\}T^{1/4}\log^2 T\log\log T \end{aligned}$$

This concludes the proof. □

7.2 Supporting Lemmas for Theorem 7.1

Lemma 7.1. *For batch β, denoted by $\mathcal{T}_\beta$, let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\tau_\beta}\}$ be a set of i.i.d arms and $\{r_1, \ldots, r_{\tau_\beta}\}$ be the corresponding rewards associated with these arms, where $\tau_\beta = |\mathcal{T}_\beta|$. Define $\hat{\boldsymbol{\theta}}_\beta$ to be the MLE estimate for this batch, i.e*

$$\hat{\boldsymbol{\theta}}_\beta = \arg\min_{\boldsymbol{\theta}} \sum_{s\in\mathcal{T}_\beta}\sum_{i=1}^{K} \mathbb{1}\{y_s = i\}\log z_i(\boldsymbol{x}_s, \boldsymbol{\theta}) + \frac{\lambda}{2}\|\boldsymbol{\theta}\|_2^2$$

Let the optimal Hessian matrix for batch β, $\boldsymbol{H}^\star_\beta$, be defined as in Section 7.1. Then, with probability greater than $1 - \frac{1}{T^2}$, we have:

$$\left|\left|\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{H}^\star_\beta} \leq 12S\sqrt{\log T + Kd} + 8S\lambda^{-1/2}(\log T + Kd) + 2S^{3/2}\lambda^{1/2}$$

Proof. For a batch β, we define the following quantity:

$$\boldsymbol{G}_\beta(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \sum_{t\in\mathcal{T}_\beta} \boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \lambda\boldsymbol{I}_{Kd\times Kd}$$

Then,

$$
\begin{aligned}
\left|\left|\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{H}_\beta^\star} &\overset{(i)}{\leq} \sqrt{1+2S}\left|\left|\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{G}_\beta(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} \\
&\leq \sqrt{1+2S}\left|\left|\boldsymbol{G}_\beta(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)\left(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right)\right|\right|_{\boldsymbol{G}_\beta^{-1}(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} \\
&\leq \sqrt{1+2S}\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)\otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \boldsymbol{I}_{Kd\times Kd}\right]\left(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right)\right|\right|_{\boldsymbol{G}_\beta^{-1}(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} \\
&\overset{(ii)}{\leq} \sqrt{1+2S}\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)\otimes \boldsymbol{x}_t^\top\right]\left(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right)\otimes \boldsymbol{x}_t + \lambda\left(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right)\right|\right|_{\boldsymbol{G}_\beta^{-1}(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} \\
&\overset{(iii)}{\leq} \sqrt{1+2S}\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{z}(\boldsymbol{x}_t, \hat{\boldsymbol{\theta}}_\beta)\right]\otimes \boldsymbol{x}_t - \lambda\hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{G}_\beta^{-1}(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} + \lambda\sqrt{1+2S}\left|\left|\boldsymbol{\theta}^\star\right|\right|_{\boldsymbol{G}_\beta^{-1}(\boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\beta)} \\
&\overset{(iv)}{\leq} (1+2S)\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{z}(\boldsymbol{x}_t, \hat{\boldsymbol{\theta}}_\beta)\right]\otimes \boldsymbol{x}_t - \lambda\hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{H}_\beta^{\star\,-1}} + \sqrt{\lambda(1+2S)}\left|\left|\boldsymbol{\theta}^\star\right|\right|_2 \\
&\overset{(v)}{\leq} 3S\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right|\right|_{\boldsymbol{H}_\beta^{\star\,-1}} + \sqrt{3}\lambda^{1/2}S^{3/2}
\end{aligned}
$$

where (i) follows from Lemma 9.2 , (ii) follows from Mixed Product Property, (iii) follows from the Mean value Theorem and the triangle inequality, (iv) follows from the fact that $\boldsymbol{G}_\beta \succcurlyeq \lambda\boldsymbol{I}$ and Lemma 9.2, and (v) follows from Lemma 9.3 and the fact that $||\boldsymbol{\theta}^\star||_2 \leq S$.

Now, consider the following term:

$$
\begin{aligned}
\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right|\right|_{\boldsymbol{H}_\beta^{\star\,-1}} &= \left|\left|\sum_{t\in\mathcal{T}_\beta}\boldsymbol{H}_\beta^{\star\,-1/2}\left(\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right)\right|\right|_2 \\
&= \max_{\boldsymbol{y}\in\mathcal{B}_2(Kd)}\left\langle \boldsymbol{y}, \sum_{t\in\mathcal{T}_\beta}\boldsymbol{H}_\beta^{\star\,-1/2}\left(\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right)\right\rangle
\end{aligned}
$$

where $\mathcal{B}_2(Kd)$ represents the $Kd-$dimensional unit ball with respect to the ℓ_2 norm. We construct an $\epsilon-$net for this unit ball, denoted as C_ϵ. For any $\boldsymbol{y}\in\mathcal{B}_2(Kd)$, we define $\boldsymbol{y}_\epsilon = \arg\min_{\boldsymbol{x}\in C_\epsilon}||\boldsymbol{y}-\boldsymbol{x}||_2$, then,

$$
\begin{aligned}
\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right|\right|_{\boldsymbol{H}_\beta^{\star\,-1}} &= \max_{\boldsymbol{y}\in\mathcal{B}_2(Kd)}\left\langle \boldsymbol{y}, \sum_{t\in\mathcal{T}_\beta}\boldsymbol{H}_\beta^{\star\,-1/2}\left(\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right)\right\rangle \\
&= \max_{\boldsymbol{y}\in\mathcal{B}_2(Kd)}\left\langle (\boldsymbol{y}-\boldsymbol{y}_\epsilon) + \boldsymbol{y}_\epsilon, \sum_{t\in\mathcal{T}_\beta}\boldsymbol{H}_\beta^{\star\,-1/2}\left(\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right)\right\rangle
\end{aligned}
$$

Thus, an application of the Cauchy-Schwarz inequality along with the fact that $||\boldsymbol{y}-\boldsymbol{y}_\epsilon||_2 \leq \epsilon$ gives us

$$
\left|\left|\sum_{t\in\mathcal{T}_\beta}\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right|\right|_{\boldsymbol{H}_\beta^{\star\,-1}} \leq \frac{1}{1-\epsilon}\left\langle \boldsymbol{y}_\epsilon, \sum_{t\in\mathcal{T}_\beta}\boldsymbol{H}_\beta^{\star\,-\frac{1}{2}}\left(\left[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s\right]\otimes \boldsymbol{x}_t\right)\right\rangle
$$

The above term can be bounded using the Bernstein Inequality (Lemma 9.4), which has been done in Lemma 7.2. We note that $|C_\epsilon| \leq \left(\frac{2}{\epsilon}\right)^{Kd}$. We now set $\epsilon = 0.5$ and $\delta = (T^2|C_\epsilon|)^{-1}$ and then perform a union bound over C_ϵ. We get that with probability greater than $1 - \frac{1}{T^2}$, we have:

$$
\begin{aligned}
\left\|\sum_{t \in \mathcal{T}_\beta} [\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right\|_{\boldsymbol{H}_\beta^{\star\,-1}} &\leq 2\left(\sqrt{2\log\left(T^2 4^{Kd}\right)} + \frac{4}{3}\lambda^{-1/2}\log(T^2 4^{Kd})\right) \\
&\leq 4\sqrt{\log T + Kd} + \frac{8}{3}\lambda^{-1/2}(\log T + Kd)
\end{aligned}
$$

Substituting this into the original bound finishes the proof. □

Lemma 7.2. *Let $\boldsymbol{y}$ be a fixed vector with $||\boldsymbol{y}||_2 \leq 1$, then, with probability at least $1-\delta$*

$$
\sum_{t \in \mathcal{T}_\beta} \left[\boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} [\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right] \leq \sqrt{2\log\frac{1}{\delta}} + \frac{4}{3\sqrt{\lambda}}\log\frac{1}{\delta}
$$

Proof. Denote $\varphi_t = \boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left([\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right)$. From Lemma 9.5, we have that $\mathbb{E}[\varphi_t] = 0$.

Also,

$$
\begin{aligned}
\mathbb{V}[\varphi_t] &= \mathbb{E}\left[\varphi_t^2\right] - \mathbb{E}[\varphi_t]^2 \overset{(i)}{=} \mathbb{E}\left[\varphi_t \varphi_t^\top\right] \\
&= \mathbb{E}\left[\boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left([\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right)\left([\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right)^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \boldsymbol{y}\right] \\
&\overset{(ii)}{=} \boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \mathbb{E}\left[[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s][\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s]^\top \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top\right] \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \boldsymbol{y} \\
&= \boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left(\mathbb{E}\left[[\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s][\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s]^\top\right] \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top\right) \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \boldsymbol{y} \\
&\overset{(iii)}{=} \boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left(\boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top\right) \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \boldsymbol{y} \overset{(iv)}{=} \boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left(\boldsymbol{H}_\beta^\star - \lambda \boldsymbol{I}\right) \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \boldsymbol{y} \\
&\leq \boldsymbol{y}^\top \boldsymbol{y} \leq 1
\end{aligned}
$$

where (i) follows from the fact that φ_t is a scalar and $\mathbb{E}[\varphi_t] = 0$, (ii) follows from the fact that $(\boldsymbol{A} \otimes \boldsymbol{B})^\top = \boldsymbol{A}^\top \otimes \boldsymbol{B}^\top$ and the mixed-product property of the Kronecker Product, (iii) follows from Lemma 9.5, and (iv) follows from the definition of $\boldsymbol{H}_\beta^\star$.

Finally, we note that

$$
\begin{aligned}
|\varphi_t - \mathbb{E}[\varphi_t]| = |\varphi_t| &= \left|\boldsymbol{y}^\top \boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left([\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right)\right| \overset{(i)}{\leq} ||\boldsymbol{y}||_2 \left\|\boldsymbol{H}_\beta^{\star\,-\frac{1}{2}} \left([\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s] \otimes \boldsymbol{x}_t\right)\right\|_2 \\
&\overset{(ii)}{\leq} \left\|\boldsymbol{H}_\beta^{\star\,-\frac{1}{2}}\right\| ||(\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s) \otimes \boldsymbol{x}_t||_2 \overset{(iii)}{\leq} \frac{1}{\sqrt{\lambda}} ||\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star) - \boldsymbol{m}_s||_2 \, ||\boldsymbol{x}_t||_2 \\
&\overset{(iv)}{\leq} \frac{1}{\sqrt{\lambda}} \left(||\boldsymbol{z}(\boldsymbol{x_t}, \boldsymbol{\theta}^\star)||_2 + ||\boldsymbol{m}_s||_2\right) \overset{(v)}{\leq} \frac{2}{\sqrt{\lambda}}
\end{aligned}
$$

where (i) follows from Cauchy-Schwarz, (ii) follows from the fact that $||\boldsymbol{y}||_2 \leq 1$ and $||\boldsymbol{A}\boldsymbol{x}||_2 \leq ||\boldsymbol{A}|| \, ||\boldsymbol{x}||_2$, (iii) follows from $\boldsymbol{H}_\beta^\star \succcurlyeq \lambda \boldsymbol{I}$ and the fact that $||\boldsymbol{a} \otimes \boldsymbol{b}||_2 = ||\boldsymbol{a}||_2 \, ||\boldsymbol{b}||_2$, (iv) follows from $||\boldsymbol{x}||_2 \leq 1$ and uses the triangle inequality, and (v) follows from the fact $\|\boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})\|_2 \leq \|\boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})\|_1 \leq 1$.

Substituting $v = 1$ and $b = \frac{2}{\sqrt{\lambda}}$ in Lemma 9.4 finishes the proof. □

Lemma 7.3. *Let $\tilde{\boldsymbol{V}}_\beta$ and $\boldsymbol{H}^\star_\beta$ be the design and optimal Hessian matrices defined as in Section 7.1. Then, we have that*

$$\tilde{\boldsymbol{V}}_\beta \preccurlyeq \kappa \boldsymbol{H}^\star_\beta$$

Proof. From the definition of κ, we know that $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}) \succcurlyeq \frac{1}{\kappa}\boldsymbol{I}$.

Hence, using the fact that $\kappa > 1$, we can say that

$$\begin{aligned}\tilde{\boldsymbol{V}}_\beta = \boldsymbol{I}_{K\times K} \otimes \boldsymbol{V}_\beta = \boldsymbol{I}_{K\times K} \otimes \left(\lambda \boldsymbol{I}_{d\times d} + \sum_{t\in\mathcal{T}_\beta} \boldsymbol{x}_t\boldsymbol{x}_t^\top\right) = \lambda \boldsymbol{I}_{Kd\times Kd} + \boldsymbol{I}_{K\times K} \otimes \sum_{t\in\mathcal{T}_\beta} \boldsymbol{x}_t\boldsymbol{x}_t^\top \\ \preccurlyeq \lambda \boldsymbol{I}_{Kd\times Kd} + \kappa \sum_{t\in\mathcal{T}_\beta} \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top \preccurlyeq \kappa \boldsymbol{H}^\star_\beta\end{aligned}$$

□

Lemma 7.4. *Let $\boldsymbol{H}^\star_\beta$ and $\boldsymbol{H}_\beta$ be the optimal and proxy Hessian matrices in batch β as defined in Section 7.1. Then, we have that*

$$\boldsymbol{H}_\beta \preccurlyeq \boldsymbol{H}^\star_\beta$$

Proof. From Lemma 9.1, we have that

$$\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta) \preccurlyeq \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}^\star) \exp\left(\sqrt{6}\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2\right)$$

We can bound $\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2$ as follows:

$$\begin{aligned}\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2 &\overset{(i)}{\leq} 2S\left|\left|\boldsymbol{I}\otimes\boldsymbol{x}^\top\right|\right|_2 \overset{(ii)}{=} 2S\sqrt{\lambda_{max}\left((\boldsymbol{I}\otimes\boldsymbol{x})(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\right)} \\ &\overset{(iii)}{=} 2S\sqrt{\lambda_{max}\left(\boldsymbol{I}\otimes\boldsymbol{x}\boldsymbol{x}^\top\right)} \overset{(iv)}{\leq} 2S\end{aligned}$$

where (i) uses the sub-multiplicativity of the norm, a triangle inequality, and the fact that $||\boldsymbol{\theta}^\star||_2 \leq S$, (ii) uses the definition of the norm, i.e., $||\boldsymbol{A}||_2 = \sqrt{\lambda_{max}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)}$, (iii) follows from the Mixed-Product property of Kronecker Products, and (iv) follows from the fact that $\lambda\left(\boldsymbol{A}\otimes\boldsymbol{B}\right) = \lambda(\boldsymbol{A})\lambda(\boldsymbol{B})$ and since $\boldsymbol{x}\boldsymbol{x}^\top$ is a rank-one matrix, the only eigenvalues are $\|\boldsymbol{x}\|_2^2$ and 0, and $0 \leq ||\boldsymbol{x}||_2 \leq 1$.

We can also bound $\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}_\beta})\right|\right|_2$ as follows:

$$\begin{aligned}\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2 &= \left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\boldsymbol{H}_\beta^{\star\,-1/2}\boldsymbol{H}_\beta^{\star\,1/2}(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2 \overset{(i)}{\leq} \left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\boldsymbol{H}_\beta^{\star\,-1/2}\right|\right|_2\left|\left|\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta\right|\right|_{\boldsymbol{H}^\star_\beta} \\ &\overset{(ii)}{\leq} \kappa^{1/2}\gamma(\lambda)\left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\kappa^{1/2}\tilde{\boldsymbol{V}}_\beta^{-1/2}\right|\right|_2 \overset{(iii)}{=} \kappa^{1/2}\gamma(\lambda)\sqrt{\lambda_{max}\left(\tilde{\boldsymbol{V}}_\beta^{-1/2}(\boldsymbol{I}\otimes\boldsymbol{x})(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\tilde{\boldsymbol{V}}_\beta^{-1/2}\right)} \\ &\overset{(iv)}{=} \kappa^{1/2}\gamma(\lambda)\sqrt{\lambda_{max}\left((\boldsymbol{I}\otimes\boldsymbol{V}_\beta^{-1/2})(\boldsymbol{I}\otimes\boldsymbol{x})(\boldsymbol{I}\otimes\boldsymbol{x}^\top)(\boldsymbol{I}\otimes\boldsymbol{V}_\beta^{-1/2})\right)} \\ &\overset{(v)}{=} \kappa^{1/2}\gamma(\lambda)\sqrt{\lambda_{max}\left(\boldsymbol{I}\otimes\boldsymbol{V}_\beta^{-1/2}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{V}_\beta^{-1/2}\right)} \overset{(vi)}{=} \kappa^{1/2}\gamma(\lambda)\left|\left|\boldsymbol{x}\right|\right|_{\boldsymbol{V}_\beta^{-1}}\end{aligned}$$

where (i) follows from the sub-multiplicativity of the norm, (ii) follows from Lemma 7.1 and Lemma 7.3, (iii) follows from the definition of the norm, (iv) follows from the definition of $\tilde{\boldsymbol{V}}_\beta$ and the fact that $(\boldsymbol{A} \otimes \boldsymbol{B})^n = \boldsymbol{A}^n \otimes \boldsymbol{B}^n$, (v) follows from the Mixed-Product property, and (vi) follows from $\lambda(\boldsymbol{A} \otimes \boldsymbol{B}) = \lambda(\boldsymbol{A})\lambda(\boldsymbol{B})$.

Thus, we can say that $\left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\beta)\right|\right|_2 \leq \min\left\{\gamma(\lambda)\kappa^{1/2}\, ||\boldsymbol{x}||_{\boldsymbol{V}_\beta^{-1}}, 2S\right\}$.

Define $B_\beta(\boldsymbol{x}) = \exp\left(\sqrt{6}\min\left\{\gamma(\lambda)\kappa^{1/2}\, ||\boldsymbol{x}||_{\boldsymbol{V}_\beta^{-1}}, 2S\right\}\right)$. Then, $\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta) \preccurlyeq \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}^\star) B_\beta(\boldsymbol{x})$. Hence, we can say,

$$\boldsymbol{H}_\beta = \lambda \boldsymbol{I}_{Kd\times Kd} + \sum_{t\in\beta} \frac{\boldsymbol{A}(\boldsymbol{x}_t, \hat{\boldsymbol{\theta}}_\beta)}{\boldsymbol{B}_\beta(\boldsymbol{x}_t)} \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top \preccurlyeq \lambda \boldsymbol{I}_{Kd\times Kd} + \sum_{t\in\beta} \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top = \boldsymbol{H}_\beta^\star$$

□

Lemma 7.5. *(Proposition 1 , Zhang & Sugiyama (2023)) For any arm $\boldsymbol{x}$, we have that,*

$$\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}_j)\right| \leq \epsilon_1(j, \boldsymbol{x}, \lambda) + \epsilon_2(j, \boldsymbol{x}, \lambda)$$

where

$$\epsilon_1(j, \boldsymbol{x}, \lambda) = \gamma(\lambda)\left|\left|\boldsymbol{H}_j^{-1/2}(\boldsymbol{I} \otimes \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)\boldsymbol{\rho}\right|\right|_2 \text{ and } \epsilon_2(j, \boldsymbol{x}, \lambda) = 3R\gamma(\lambda)^2\left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{-1/2}\right|\right|_2^2$$

Proof. We provide the proof for the sake of completeness:

$$\begin{aligned}
\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}_j)\right| &= \left|\sum_{i=1}^K \rho_i \left[z_i(\boldsymbol{x}, \boldsymbol{\theta}^\star) - z_i(\boldsymbol{x}, \boldsymbol{\theta}_j)\right]\right| \\
&= \left|\sum_{i=1}^K \rho_i \boldsymbol{\nabla} z_i(\boldsymbol{x}, \boldsymbol{\theta}_j)^\top \left[(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right] + \sum_{i=1}^K \rho_i \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right|\right|_{\boldsymbol{Z}_i}\right| \\
&\leq \left|\boldsymbol{\rho}^\top \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right| + \left|\sum_{i=1}^K \rho_i \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right|\right|_{\boldsymbol{Z}_i}^2\right|
\end{aligned}$$

where

$$\boldsymbol{Z}_i = \int_0^1 (1-v)\nabla^2 z_i(\boldsymbol{x}, v\boldsymbol{\theta}^\star + (1-v)\boldsymbol{\theta}_j)\ \mathrm{d}v$$

Beginning with the first term :

$$\begin{aligned}
\left|\boldsymbol{\rho}^\top \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right| &= \left|\boldsymbol{\rho}^\top \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{\star\,-1/2}\boldsymbol{H}_j^{\star\,1/2}(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right| \\
&\overset{(i)}{\leq} ||\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j||_{\boldsymbol{H}_j^\star}\left|\left|\boldsymbol{\rho}^\top \boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{\star\,-1/2}\right|\right|_2 \\
&\leq \gamma(\lambda)\left|\left|\boldsymbol{H}_j^{\star\,-1/2}(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)\boldsymbol{\rho}\right|\right|_2 \\
&\overset{(ii)}{\leq} \gamma(\lambda)\left|\left|\boldsymbol{H}_j^{-1/2}(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}_j)\boldsymbol{\rho}\right|\right|_2
\end{aligned}$$

where (i) follows from the sub-multiplicativity of the norm and (ii) is due to Lemma 7.4.

For the second term, for some $k \in [1, K]$, we make the following observation:

$$\boldsymbol{Z}_k = \int_0^1 (1-v)\nabla^2 z_k(\boldsymbol{x}, v\boldsymbol{\theta}^\star + (1-v)\boldsymbol{\theta}_j) \ \mathrm{d}v \preccurlyeq 3\boldsymbol{I} \int_0^1 (1-v) \ \mathrm{d}v \preccurlyeq 3\boldsymbol{I}$$

Thus, we have:

$$\begin{aligned}\left|\sum_{i=1}^{K} \rho_i \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right|\right|_{\boldsymbol{Z}_i}^2\right| &\leq \left|\sum_{i=1}^{K} 3\rho_i \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right|\right|_2^2\right| \\
&\leq 3R \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{\star\ -1/2}\boldsymbol{H}_j^{\star\ 1/2}(\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j)\right|\right|_2^2 \\
&\leq 3R \left|\left|\boldsymbol{\theta}^\star - \boldsymbol{\theta}_j\right|\right|_{\boldsymbol{H}_j^\star}^2 \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{\star\ -1/2}\right|\right|_2^2 \\
&\leq 3R\gamma(\lambda)^2 \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{\star\ -1/2}\right|\right|_2^2 \\
&\leq 3R\gamma(\lambda)^2 \left|\left|(\boldsymbol{I}_{K\times K} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_j^{-1/2}\right|\right|_2^2\end{aligned}$$

□

Lemma 7.6. *Let $\boldsymbol{x}_t^\star$ be the optimal arm at round t, i.e $\boldsymbol{x}_t^\star = \arg\max_{\boldsymbol{x}\in\mathcal{X}_t} \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}^\star)$. Then, the optimal arm never gets eliminated in any round.*

Proof. From Lemma 7.5, we know that

$$\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}_j)\right| \leq \epsilon_1(j, \boldsymbol{x}, \lambda) + \epsilon_2(j, \boldsymbol{x}, \lambda)$$

Also, from Algorithm 1, we have the definitions of $\text{UCB}(j, \boldsymbol{x}, \lambda)$ and $\text{LCB}(j, \boldsymbol{x}, \lambda)$ as:

$$\begin{aligned}\text{UCB}(j, \boldsymbol{x}, \lambda) &= \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}_j) + \epsilon_1(j, \boldsymbol{x}, \lambda) + \epsilon_2(j, \boldsymbol{x}, \lambda) \\
\text{LCB}(j, \boldsymbol{x}, \lambda) &= \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}_j) - \epsilon_1(j, \boldsymbol{x}, \lambda) - \epsilon_2(j, \boldsymbol{x}, \lambda)\end{aligned}$$

From Algorithm 1, we know that an arm $\boldsymbol{x} \in \mathcal{X}_t$ gets eliminated if $\text{UCB}(j, \boldsymbol{x}, \lambda) \leq \max_{\boldsymbol{y}\in\mathcal{X}_t} \text{LCB}(j, \boldsymbol{y}, \lambda)$. Thus, showing that $\text{UCB}(j, \boldsymbol{x}_t^\star, \lambda) \geq \max_{\boldsymbol{y}\in\mathcal{X}_t} \text{LCB}(j, \boldsymbol{y}, \lambda)$ accounts to showing that $\boldsymbol{x}_t^\star$ never gets eliminated.

We assume that $\arg\max_{\boldsymbol{y}\in\mathcal{X}_t} \text{LCB}(j, \boldsymbol{y}, \lambda) = \boldsymbol{y}$. Then, for any arm $\boldsymbol{x} \in \mathcal{X}_t$, we have that

$$\begin{aligned}\text{LCB}(j, \boldsymbol{x}, \lambda) &\leq \max_{\boldsymbol{y}\in\mathcal{X}_t} \text{LCB}(j, \boldsymbol{y}, \lambda) \\
&= \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{y}, \boldsymbol{\theta}_j) - \epsilon_1(j, \boldsymbol{y}, \lambda) - \epsilon_2(j, \boldsymbol{y}, \lambda) \\
&\overset{(i)}{\leq} \left[\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{y}, \boldsymbol{\theta}^\star) + \epsilon_1(j, \boldsymbol{y}, \lambda) + \epsilon_2(j, \boldsymbol{y}, \lambda)\right] - \epsilon_1(j, \boldsymbol{y}, \lambda) - \epsilon_2(j, \boldsymbol{y}, \lambda) \\
&= \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{y}, \boldsymbol{\theta}^\star) \\
&\overset{(ii)}{\leq} \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}^\star) \\
&\overset{(iii)}{\leq} \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}_j) + \epsilon_1(j, \boldsymbol{x}_t^\star, \lambda) + \epsilon_2(j, \boldsymbol{x}_t^\star, \lambda) \\
&= \text{UCB}(j, \boldsymbol{x}_t^\star, \lambda)\end{aligned}$$

where (i) follows from Lemma 7.5, (ii) follows from the fact that $\boldsymbol{x}_t^\star = \underset{\boldsymbol{y}\in\mathcal{X}_t}{\arg\max}\, \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{y}, \boldsymbol{\theta}^\star)$, and (iii) again follows from Lemma 7.5. □

Lemma 7.7. *Let $B_\beta(\boldsymbol{x})$ be as defined in Section 7.1. Then, we have that*

$$\sqrt{B_\beta(\boldsymbol{x})} \leq \frac{1}{2}e^{3S}\gamma(\lambda)\kappa^{1/2}S^{-1}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}} + 1$$

Proof.

$$\sqrt{B_\beta(\boldsymbol{x})} = \exp\left(\sqrt{6}\min\left\{S, \frac{1}{2}\gamma(\lambda)\kappa^{1/2}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}}\right\}\right) \leq \frac{1}{2}e^{3S}\gamma(\lambda)\kappa^{1/2}S^{-1}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}} + 1$$

where the inequality follows from Lemma 9.6 by choosing $\min\{2S, \gamma(\lambda)\kappa^{1/2}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}}\} = \gamma(\lambda)\kappa^{1/2}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}}$ and $M = \sqrt{6}S$.

□

Lemma 7.8. *Let $\epsilon_1(\beta, \boldsymbol{x}, \lambda)$ be as defined in Lemma 7.5. Then, we have*

$$\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\epsilon_1(\beta, \boldsymbol{x}, \lambda)\right] \leq \frac{8R\kappa^{1/2}K^{5/2}de^{3S}\gamma(\lambda)^2S^{-1}\sqrt{\log Kd\log d}}{\tau_\beta} + \frac{4RK^2d^{1/2}\gamma(\lambda)\sqrt{\log(Kd)}}{\sqrt{\tau_\beta}}$$

Proof.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\epsilon_1(\beta, \boldsymbol{x}, \lambda)\right] &= \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\gamma(\lambda)\left|\left|\boldsymbol{H}_\beta^{-1/2}(\boldsymbol{I}\otimes\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)\boldsymbol{\rho}\right|\right|_2\right] \\
&\overset{(i)}{\leq} \gamma(\lambda)\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\boldsymbol{H}_\beta^{-1/2}(\boldsymbol{I}\otimes\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)^{1/2}\right|\right|_2||\boldsymbol{\rho}||_{\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)}\right] \\
&\overset{(ii)}{\leq} R\gamma(\lambda)\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\boldsymbol{A}(\boldsymbol{x},\hat{\boldsymbol{\theta}}_\beta)^{1/2}(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\boldsymbol{H}_\beta^{-1/2}\right|\right|_2\right] \\
&\overset{(iii)}{\leq} R\gamma(\lambda)\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\sqrt{B_\beta(\boldsymbol{x})}\tilde{\boldsymbol{X}}_\beta^\top\boldsymbol{H}_\beta^{-1/2}\right|\right|_2\right] \\
&\overset{(iv)}{\leq} 4R\gamma(\lambda)K^2\sqrt{\frac{d\log Kd}{\tau_\beta}}\left\{\frac{1}{2}e^{3S}\gamma(\lambda)\kappa^{1/2}S^{-1}\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\|\boldsymbol{x}\|_{\boldsymbol{V}_\beta^{-1}}\right] + 1\right\} \\
&\overset{(v)}{\leq} 4R\gamma(\lambda)K^2\sqrt{\frac{d\log Kd}{\tau_\beta}}\left\{2e^{3S}\gamma(\lambda)\kappa^{1/2}S^{-1}\sqrt{\frac{Kd\log d}{\tau_\beta}} + 1\right\} \\
&\leq \frac{8R\kappa^{1/2}K^{5/2}de^{3S}\gamma(\lambda)^2S^{-1}\sqrt{\log Kd\log d}}{\tau_\beta} + \frac{4RK^2d^{1/2}\gamma(\lambda)\sqrt{\log(Kd)}}{\sqrt{\tau_\beta}}
\end{aligned}$$

where (i) follows from $||\boldsymbol{A}\boldsymbol{x}||_2 \leq ||\boldsymbol{A}||_2\,||\boldsymbol{x}||_2$, (ii) follows from the fact that $\boldsymbol{A}(\boldsymbol{x},\boldsymbol{\theta}) \preccurlyeq \boldsymbol{I}$, (iii) follows from the defintion of $\tilde{\boldsymbol{X}}$, (iv) follows from Lemma 7.7, the fact that $\max\{ab\} \leq \max\{a\}\max\{b\}$, and Lemma 7.18, and (v) follows from Lemma 7.17.

□

Lemma 7.9. *Let $\epsilon_2(\beta, \boldsymbol{x}, \lambda)$ be as defined in Lemma 7.5. Then, we have*

$$\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\epsilon_2(\beta, \boldsymbol{x}, \lambda)\right] \leq \frac{96R\kappa\gamma(\lambda)^2}{\tau_\beta}Kd^2$$

Proof. Recall from Lemma 7.5, in one of the intermediate steps, we have that

$$\epsilon_2(\beta, \boldsymbol{x}, \lambda) = 3R\gamma(\lambda)^2 \left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_\beta^{\star\ -1/2}\right|\right|_2^2$$

Thus, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}[\epsilon_2(\beta, \boldsymbol{x}, \lambda)] &= \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} 3R\gamma(\lambda)^2 \left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_\beta^{\star\ -1/2}\right|\right|_2^2\right] \\
&= 3R\gamma(\lambda)^2 \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} \left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_\beta^{\star\ -1/2}\right|\right|_2^2\right] \\
&\overset{(i)}{\leq} 3R\kappa\gamma(\lambda)^2 \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} \left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\tilde{\boldsymbol{V}}_\beta^{-1/2}\right|\right|_2^2\right] \\
&\overset{(ii)}{\leq} 3R\kappa\gamma(\lambda)^2 \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} \left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{I} \otimes \boldsymbol{V}_\beta^{-1/2})\right|\right|_2^2\right] \\
&\overset{(iii)}{\leq} 3R\kappa\gamma(\lambda)^2 \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||_{\boldsymbol{V}_\beta^{-1}}^2\right] \overset{(iv)}{\leq} \frac{48R\kappa\gamma(\lambda)^2}{\tau_\beta}(K+1)d^2 \leq \frac{96R\kappa\gamma(\lambda)^2}{\tau_\beta}Kd^2
\end{aligned}$$

where (i) follows from Lemma 7.3, (ii) follows from the definition of $\tilde{\boldsymbol{V}}_\beta$, (iii) follows from the Mixed-Product Property and the fact that $\lambda(\boldsymbol{A} \otimes \boldsymbol{B}) = \lambda(\boldsymbol{A})\lambda(\boldsymbol{B})$, and (iv) follows from Lemma 7.16.

□

Lemma 7.10. *Let t be a time round in batch $\beta+1$, i.e $t \in \mathcal{T}_\beta$. Then, the expected regret incurred at round t, denoted as R_t can be bounded as:*

$$R_t \leq \frac{32RK\kappa^{1/2}d\gamma(\lambda)^2}{\tau_\beta}\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\} + \frac{16RK^2d^{1/2}\gamma(\lambda)\sqrt{\log(Kd)}}{\sqrt{\tau_\beta}}$$

Proof. Using Lemma 7.5,

$$\boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \leq \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}_\beta) - \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}_\beta) + \epsilon_1(\beta, \boldsymbol{x}_t^\star, \lambda) + \epsilon_2(\beta, \boldsymbol{x}_t^\star, \lambda) + \epsilon_1(\beta, \boldsymbol{x}_t, \lambda) + \epsilon_2(\beta, \boldsymbol{x}_t, \lambda)$$

Since $\boldsymbol{x}_t$ was not eliminated, we have $\text{UCB}(\beta, \boldsymbol{x}_t, \lambda) \geq \max_{\boldsymbol{y}\in\mathcal{X}} \text{LCB}(\beta, \boldsymbol{y}, \lambda) \geq \text{LCB}(\beta, \boldsymbol{x}_t^\star, \lambda)$ since $\boldsymbol{x}_t^\star$ never gets eliminated (Lemma 7.6). Thus,

$$\boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}_\beta) + \epsilon_1(\beta, \boldsymbol{x}_t, \lambda) + \epsilon_2(\beta, \boldsymbol{x}_t, \lambda) \geq \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}_\beta) - \epsilon_1(\beta, \boldsymbol{x}_t^\star, \lambda) - \epsilon_2(\beta, \boldsymbol{x}_t^\star, \lambda)$$

Thus, we get

$$\begin{aligned}
\boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) &\leq 2\epsilon_1(\beta, \boldsymbol{x}_t, \lambda) + 2\epsilon_2(\beta, \boldsymbol{x}_t, \lambda) + 2\epsilon_1(\beta, \boldsymbol{x}_t^\star, \lambda) + 2\epsilon_2(\beta, \boldsymbol{x}_t^\star, \lambda) \\
&\leq 4\max_{\boldsymbol{x}\in\mathcal{X}} \epsilon_1(\beta, \boldsymbol{x}, \lambda) + 4\max_{\boldsymbol{x}\in\mathcal{X}} \epsilon_2(\beta, \boldsymbol{x}, \lambda)
\end{aligned}$$

Taking an expectation on both sides, we get

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right] \leq 4\left(\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} \epsilon_1(\beta, \boldsymbol{x}, \lambda) + \max_{\boldsymbol{x}\in\mathcal{X}} \epsilon_2(\beta, \boldsymbol{x}, \lambda)\right]\right) \\
&\leq \frac{32RK\kappa^{1/2}d\gamma(\lambda)^2}{\tau_\beta}\left\{e^{3S}K^{3/2}S^{-1}\sqrt{\log(Kd)\log d} + 12\kappa^{1/2}d\right\} + \frac{16RK^2d^{1/2}\gamma(\lambda)\sqrt{\log(Kd)}}{\sqrt{\tau_\beta}}
\end{aligned}$$

which follows from Lemma 7.8 and Lemma 7.9.

□

7.3 Supporting Results on Optimal Designs for 7

Recall from Section 7.1,

$$\tilde{\boldsymbol{X}}_\beta = \frac{\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}}}{\sqrt{B_\beta(\boldsymbol{x})}} \otimes \boldsymbol{x}$$

Also, recall that at each round $t \in [T]$, the feasible set of context vectors $\mathcal{X}_t$ is being sampled from some distribution $\mathcal{D}$. For a given batch β, we denote $\mathcal{D}_\beta$ to be the distribution of the pruned arm-sets post the successive elimination procedure (Section 3.1). Thus, we have that $\mathcal{D}_{\beta+1} \subset \mathcal{D}_\beta$.

We now define K different partitions of $\tilde{\boldsymbol{X}}_\beta$ as follows:

$$\tilde{\boldsymbol{x}}_\beta^{(i)} = \frac{\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}}}{\sqrt{B_\beta(\boldsymbol{x})}} \boldsymbol{e}_i \otimes \boldsymbol{x}$$

where $i \in [K]$ and $\boldsymbol{e}_i$ is the $K-$dimensional standard basis vector. We first show a few relations between $\tilde{\boldsymbol{X}}_\beta$ and $\tilde{\boldsymbol{x}}_\beta^{(i)}$:

Lemma 7.11. *Let $\tilde{\boldsymbol{X}}_\beta$ and $\tilde{\boldsymbol{x}}_\beta^{(i)}$ be defined as above. Then, we have*

$$\tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top = \sum_{i=1}^{K} \tilde{\boldsymbol{x}}_\beta^{(i)} \tilde{\boldsymbol{x}}_\beta^{(i)\top}$$

Proof.

$$\begin{aligned}
\sum_{i=1}^{K} \tilde{\boldsymbol{x}}_\beta^{(i)} \tilde{\boldsymbol{x}}_\beta^{(i)\top} &= \sum_{i=1}^{K} \left(\frac{\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}}}{\sqrt{B_\beta(\boldsymbol{x})}} \boldsymbol{e}_i \otimes \boldsymbol{x} \right) \left(\boldsymbol{e}_i^\top \frac{\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}}}{\sqrt{B_\beta(\boldsymbol{x})}} \otimes \boldsymbol{x}^\top \right) \\
&= \frac{1}{B_\beta(\boldsymbol{x})} \sum_{i=1}^{K} \boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}} \boldsymbol{e}_i \boldsymbol{e}_i^\top \boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}} \otimes \boldsymbol{x}\boldsymbol{x}^\top \\
&= \frac{1}{B_\beta(\boldsymbol{x})} \boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}} \left(\sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_i^\top \right) \boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)^{\frac{1}{2}} \otimes \boldsymbol{x}\boldsymbol{x}^\top \\
&= \frac{\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\beta)}{B_\beta(\boldsymbol{x})} \otimes \boldsymbol{x}\boldsymbol{x}^\top = \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top
\end{aligned}$$

where we use the fact that $\sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_i^\top = \boldsymbol{I}_{K \times K}$. □

Lemma 7.12. *Let $\boldsymbol{M} \in \mathbb{R}^{Kd}$ be any positive-semidefinite matrix. Then,*

$$\lambda_{max} \left(\tilde{\boldsymbol{X}}_\beta^\top \boldsymbol{M} \tilde{\boldsymbol{X}}_\beta \right) \leq \sum_{i=1}^{K} \left|\left| \tilde{\boldsymbol{x}}_\beta^{(i)} \right|\right|_{\boldsymbol{M}}^2$$

Proof.

$$\begin{aligned}
\lambda_{max} \left(\tilde{\boldsymbol{X}}_\beta^\top \boldsymbol{M} \tilde{\boldsymbol{X}}_\beta \right) &\overset{(i)}{=} \lambda_{max} \left(\tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top \boldsymbol{M} \right) \overset{(ii)}{=} \lambda_{max} \left(\sum_{i=1}^{K} \tilde{\boldsymbol{x}}_\beta^{(i)} \tilde{\boldsymbol{x}}_\beta^{(i)\top} \boldsymbol{M} \right) \\
&\overset{(iii)}{\leq} \sum_{i=1}^{K} \lambda_{max} \left(\tilde{\boldsymbol{x}}_\beta^{(i)} \tilde{\boldsymbol{x}}_\beta^{(i)\top} \boldsymbol{M} \right) \overset{(iv)}{=} \sum_{i=1}^{K} \lambda_{max} \left(\tilde{\boldsymbol{x}}_\beta^{(i)\top} \boldsymbol{M} \tilde{\boldsymbol{x}}_\beta^{(i)} \right) = \sum_{i=1}^{K} \left|\left| \tilde{\boldsymbol{x}}_\beta^{(i)} \right|\right|_{\boldsymbol{M}}^2
\end{aligned}$$

where (i) follows from the cyclic property of eigenvalues, (ii) follows from Lemma 7.11, (iii) follows from the fact that $\lambda_{max}\left(\boldsymbol{A}+\boldsymbol{B}\right) \leq \lambda_{max}\left(\boldsymbol{A}\right) + \lambda_{max}\left(\boldsymbol{B}\right)$, and (iv) again follows from the cyclic property of eigenvalues. $\square$

We first redefine the Distributional Optimal Design (Definition **??**) for a set $\mathcal{X}$.

$$\pi(\mathcal{X}) = \begin{cases} \pi_G(\mathcal{X}) & \text{w.p.} \frac{1}{2} \\ \pi^S_{\boldsymbol{M}_i}(\mathcal{X}) & \text{w.p.} \frac{p_i}{2} \end{cases}$$

where π_G is the G-optimal design and $\pi^S_{\boldsymbol{M_i}}$ represents the Softmax Policy with respect to $\boldsymbol{M}_i$. We refer the reader to Definition **??** for more details.

We now define a few notations regarding some of the information and design matrices used throughout this section.

1. $\mathbb{I}^\lambda_\mathcal{D}(\pi) = \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi(\mathcal{X})}{\mathbb{E}} \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top\right]$
2. $\mathbb{W}^{(i)}_\mathcal{D}(\pi) = \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi(\mathcal{X})}{\mathbb{E}} \tilde{\boldsymbol{x}}^{(i)}_\beta \tilde{\boldsymbol{x}}^{(i)\,\top}_\beta\right]$
3. $\mathbb{W}^{(0)}_\mathcal{D}(\pi) = \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi(\mathcal{X})}{\mathbb{E}} \boldsymbol{x}\boldsymbol{x}^\top\right]$

Suppose Algorithm 2 is called with the inputs β and $\mathcal{S}_\beta$, where β is the current batch index. Then, the policy returned by the algorithm is denoted by π_β, where

$$\pi_\beta = \frac{1}{K+1}\left(\sum_{i=0}^{K} \pi_{\beta,i}\right)$$

where $\pi_{\beta,0}$ and $\pi_{\beta,i}$ $i \in [K]$ represents the Distributional Optimal Design learned over $\mathcal{S}_\beta$ and $F_i(\mathcal{S}_\beta, \beta), i \in [K]$ respectively. Here F_i is as defined in Equation 8.

We now state a few results that relate the design matrices $\boldsymbol{H}$ and $\boldsymbol{V}$ as well as the matrices $|\mathbb{I}$ and $\mathbb{W}$.

Lemma 7.13. *(Lemma A.16, Sawarni et al. (2024)) Let* $\boldsymbol{V}_\beta$ *and* $\boldsymbol{H}_\beta$ *as defined in Section 7.1 and* $\boldsymbol{W}^{(0)}_\mathcal{D}(\pi_\beta)$ *and* $\boldsymbol{I}^\lambda_\mathcal{D}(\pi_\beta)$ *be as defined in Section 7.3.*

Then, with probability at least $1-\frac{1}{T^2}$*, we have that*

$$\boldsymbol{V}_\beta \succcurlyeq \frac{\tau_\beta}{8}\mathbb{W}^{(0)}_\mathcal{D}(\pi_\beta)$$

$$\boldsymbol{H}_\beta \succcurlyeq \frac{\tau_\beta}{8}\mathbb{I}^\lambda_\mathcal{D}(\pi_\beta)$$

Lemma 7.14. *For all* $i \in [0, K]$*, we have that*

$$(K+1)\mathbb{I}^\lambda_\mathcal{D}(\pi_\beta) \succcurlyeq \mathbb{I}^\lambda_\mathcal{D}(\pi_{\beta,i})$$

Proof.

$$\begin{aligned}\mathbb{I}^\lambda_\mathcal{D}(\pi_\beta) = \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi_\beta(\mathcal{X})}{\mathbb{E}} \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top\right] &\overset{(i)}{\succcurlyeq} (K+1)^{-1}\underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\sum_{i=0}^{K}\underset{\boldsymbol{x}\sim\pi_{\beta,i}(\mathcal{X})}{\mathbb{E}} \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top\right] \\ &\succcurlyeq (K+1)^{-1}\underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi_{\beta,i}(\mathcal{X})}{\mathbb{E}} \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top\right] = (K+1)^{-1}\mathbb{I}^\lambda_\mathcal{D}(\pi_{\beta,i})\end{aligned}$$

where (i) follows from the definition of π_β. □

Lemma 7.15. *For all $i \in [K]$, we have that*

$$\mathbb{I}^{\lambda}_{\mathcal{D}}(\pi) \succcurlyeq \mathbb{W}^{(i)}_{\mathcal{D}}(\pi)$$

Proof.

$$\mathbb{I}^{\lambda}_{\mathcal{D}}(\pi) = \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}}\left[\mathop{\mathbb{E}}_{\boldsymbol{x}\sim\pi(\mathcal{X})} \tilde{\boldsymbol{X}}_\beta \tilde{\boldsymbol{X}}_\beta^\top\right] \overset{(i)}{=} \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}}\left[\mathop{\mathbb{E}}_{\boldsymbol{x}\sim\pi(\mathcal{X})} \sum_{i=1}^{K} \tilde{\boldsymbol{x}}^{(i)}_\beta \tilde{\boldsymbol{x}}^{(i)\top}_\beta\right] \succcurlyeq \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}}\left[\mathop{\mathbb{E}}_{\boldsymbol{x}\sim\pi(\mathcal{X})} \tilde{\boldsymbol{x}}^{(i)}_\beta \tilde{\boldsymbol{x}}^{(i)\top}_\beta\right] = \mathbb{W}^{(i)}_{\mathcal{D}}(\pi)$$

□

Using the lemmas stated above, we now derive a few results.

Lemma 7.16. *Let $\boldsymbol{V}_\beta$ be as defined in Section 7.1 and τ_β be the length of the β batch, i.e $|\mathcal{T}_\beta = \tau_\beta$. Then, we have*

$$\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\boldsymbol{V}_\beta^{-1}}\right] \leq \frac{16}{\tau_\beta}(K+1)d^2$$

Proof.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\boldsymbol{V}_\beta^{-1}}\right] &\overset{(i)}{\leq} \frac{8}{\tau_\beta} \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\mathbb{W}^{(0)\,-1}_{\mathcal{D}_\beta}(\pi_\beta)}\right] \\
&\overset{(ii)}{\leq} \frac{8}{\tau_\beta}(K+1) \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\mathbb{W}^{(0)\,-1}_{\mathcal{D}_\beta}(\pi_{\beta,0})}\right] \\
&\overset{(iii)}{\leq} \frac{16}{\tau_\beta}(K+1) \mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\mathbb{W}^{(0)\,-1}_{\mathcal{D}_\beta}(\pi_G)}\right] \\
&\overset{(iv)}{\leq} \frac{16}{\tau_\beta}(K+1)d^2
\end{aligned}$$

where (i) follows from Lemma 7.13, (ii) follows from Lemma 7.14 and the fact that $\mathcal{D}_{\beta+1} \subset \mathcal{D}_\beta$ and hence, $\mathbb{E}_{\mathcal{D}_{\beta+1}} \leq \mathbb{E}_{\mathcal{D}_\beta}$, (iii) follows from the definition of $\pi_{\beta,0}$ and uses the fact that $\pi_{\beta,0} \succcurlyeq \frac{\pi_G}{2}$, and (iv) follows from Lemma 9.8.

□

Lemma 7.17. *Let $\boldsymbol{V}_\beta$ be as defined in Section 7.1 and τ_β be the length of the β batch, i.e $|\mathcal{T}_\beta| = \tau_\beta$. Then, we have*

$$\mathop{\mathbb{E}}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||_{\boldsymbol{V}_\beta^{-1}}\right] \leq 4\sqrt{\frac{Kd\log d}{\tau_\beta}}$$

Proof.

$$
\begin{aligned}
\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}||\boldsymbol{x}||_{\boldsymbol{V}_\beta^{-1}}\right] &\overset{(i)}{\leq} \sqrt{\frac{8}{\tau_\beta}}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}||\boldsymbol{x}||_{\mathbb{W}_{\mathcal{D}_\beta}^{(0)}{}^{-1}(\pi_\beta)}\right] \\
&\overset{(ii)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}||\boldsymbol{x}||_{\mathbb{W}_{\mathcal{D}_\beta}^{(0)}{}^{-1}(\pi_{\beta,0})}\right] \\
&\overset{(iii)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)d\log d} \\
&\leq 4\sqrt{\frac{Kd\log d}{\tau_\beta}}
\end{aligned}
$$

where (i) follows from Lemma 7.13 and the fact that $\mathcal{D}_{\beta+1}\subset\mathcal{D}_\beta$, (ii) follows in a similar manner as Lemma 7.14, and (iii) follows from Lemma 9.7.

□

Lemma 7.18. *Let $\tilde{\boldsymbol{X}}_\beta$ and $\boldsymbol{H}_\beta$ be as defined in Section 7.1. Denote $\tau_\beta = |\mathcal{T}_\beta|$. Then, we have that*

$$
\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\tilde{\boldsymbol{X}}_\beta^\top\boldsymbol{H}_\beta^{-1/2}\right|\right|_2\right] \leq 4K^2\sqrt{\frac{d\log(Kd)}{\tau_\beta}}
$$

Proof.

$$
\begin{aligned}
\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_{\beta+1}}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\tilde{\boldsymbol{X}}_\beta^\top\boldsymbol{H}_\beta^{-1/2}\right|\right|_2\right] &\overset{(i)}{\leq} \mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\lambda_{max}\left(\boldsymbol{H}_\beta^{-1/2}\tilde{\boldsymbol{X}}_\beta\tilde{\boldsymbol{X}}_\beta^\top\boldsymbol{H}_\beta^{-1/2}\right)}\right] \\
&\overset{(ii)}{=} \mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\lambda_{max}\left(\tilde{\boldsymbol{X}}_\beta^\top\boldsymbol{H}_\beta^{-1}\tilde{\boldsymbol{X}}_\beta\right)}\right] \\
&\overset{(iii)}{\leq} \mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\sum_{i=1}^K\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|^2_{\boldsymbol{H}_\beta^{-1}}}\right] \\
&\overset{(iv)}{\leq} \sqrt{\frac{8}{\tau_\beta}}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\sum_{i=1}^K\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|^2_{\mathbb{I}_{\mathcal{D}_\beta}^{\lambda}{}^{-1}(\pi_\beta)}}\right] \\
&\overset{(v)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\sum_{i=1}^K\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|^2_{\mathbb{I}_{\mathcal{D}_\beta}^{\lambda}{}^{-1}(\pi_{\beta,i})}}\right] \\
&\overset{(vi)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sqrt{\sum_{i=1}^K\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|^2_{\mathbb{W}_{\mathcal{D}_\beta}^{(i)}{}^{-1}(\pi_{\beta,i})}}\right] \\
&\overset{(vii)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)}\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\sum_{i=1}^K\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|_{\mathbb{W}_{\mathcal{D}_\beta}^{(i)}{}^{-1}(\pi_{\beta,i})}\right] \\
&\overset{(viii)}{\leq} \sqrt{\frac{8}{\tau_\beta}(K+1)}\sum_{i=1}^K\mathbb{E}_{\mathcal{X}\sim\mathcal{D}_\beta}\left[\max_{\boldsymbol{x}\in\mathcal{X}}\left|\left|\tilde{\boldsymbol{x}}_\beta^{(i)}\right|\right|_{\mathbb{W}_{\mathcal{D}_\beta}^{(i)}{}^{-1}(\pi_{\beta,i})}\right] \\
&\overset{(ix)}{\leq} K\sqrt{\frac{8}{\tau_\beta}(K+1)Kd\log(Kd)} \leq 4K^2\sqrt{\frac{d\log(Kd)}{\tau_\beta}}
\end{aligned}
$$

where (i) follows from the definition of the norm $||\boldsymbol{A}||_2 = \sqrt{\lambda_{max}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)}$ and the fact that $\mathcal{D}_{\beta+1} \subset \mathcal{D}_\beta$, (ii) follows from the cyclic property of eigenvalues, (iii) follows from Lemma 7.12, (iv) follows from Lemma 7.13, (v) follows from Lemma 7.14, (vi) follows from Lemma 7.15, (vii) uses the fact that for $\{a_i\}_{i=1}^N$, $\sqrt{\sum_{i=1}^N a_i^2} \leq \sqrt{\left(\sum_{i=1}^N a_i\right)^2} = \sum_{i=1}^N a_i$, $(viii)$ uses the linearity of expectations and the fact that $\max_x [f(x) + g(x)] \leq \max_x f(x) + \max_x g(x)$, and (ix) follows from Lemma 9.7. □

8 Rarely Switching Multinomial Contextual Bandit Algorithm: RS-MNL

8.1 Notations

We first define a few matrices, vectors, and scalars that are used throughout this section (here, $\boldsymbol{e}_i$ denotes the $i^{th}-$standard basis vector):

1. $\boldsymbol{V}_t = \lambda \boldsymbol{I}_{d\times d} + \sum_{s\in[t]} \boldsymbol{x}_s \boldsymbol{x}_s^\top$
2. $\tilde{\boldsymbol{V}}_t = \boldsymbol{I}_{K\times K} \otimes \boldsymbol{V}_t$
3. $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}) = \text{diag}(\boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})) - \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})\boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})^\top$
4. $\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \int_0^1 \boldsymbol{A}(\boldsymbol{x}, v\boldsymbol{\theta}_1 + (1-v)\boldsymbol{\theta}_2)\, dv$
5. $\boldsymbol{H}_t^\star = \lambda \boldsymbol{I}_{Kd\times Kd} + \sum_{s\in[t]} \boldsymbol{A}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_s \boldsymbol{x}_t^\top$
6. $\gamma(\delta) = CS^{5/4}\sqrt{Kd\log(T/\delta)}$
7. $B_t(\boldsymbol{x}) = \exp\left(\sqrt{6}\min\left\{2\kappa^{1/2}\gamma(\delta)\|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}}, 2S\right\}\right)$
8. $\boldsymbol{H}_t(\boldsymbol{\theta}) = \lambda \boldsymbol{I}_{Kd\times Kd} + \sum_{s\in[t]} \frac{\boldsymbol{A}(\boldsymbol{x}_s, \boldsymbol{\theta})}{B_s(\boldsymbol{x}_s)} \otimes \boldsymbol{x}_s \boldsymbol{x}_s^\top$
9. $\boldsymbol{H}_t(\boldsymbol{\theta}) = \lambda \boldsymbol{I}_{Kd\times Kd} + \sum_{s\in[t]} \frac{\boldsymbol{A}(\boldsymbol{x}_s, \boldsymbol{\theta})}{B_s(\boldsymbol{x}_s)} \otimes \boldsymbol{x}_s \boldsymbol{x}_s^\top$
10. $\tilde{\boldsymbol{X}}_t(\boldsymbol{\theta}) = \frac{\boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta})^{\frac{1}{2}}}{\sqrt{B_t(\boldsymbol{x}_t)}} \otimes \boldsymbol{x}_t$
11. $\tilde{\boldsymbol{x}}_t^{(i)}(\boldsymbol{\theta}) = \frac{\boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta})^{\frac{1}{2}}}{\sqrt{B_t(\boldsymbol{x}_t)}} \boldsymbol{e}_i \otimes \boldsymbol{x}_t$
12. $\boldsymbol{H}_t^i(\boldsymbol{\theta}) = \sum_{s\in[t]} \tilde{\boldsymbol{x}}_s^{(i)}(\boldsymbol{\theta}) \tilde{\boldsymbol{x}}_s^{(i)}(\boldsymbol{\theta})^\top + \lambda \boldsymbol{I}$

We now present the regret upper bound for RS-MNL by restating Theorem 4.1.

Theorem 8.1. *With high probability, the regret incurred by Algorithm 3 is bounded above by R_T where:*

$$R_T \leq CRK^{3/2}S^{5/4}(\log T\log(T/\delta))^{1/2}d\sqrt{T} + CRK^2d^2S^{5/2}\log T\log(T/\delta)\kappa^{1/2}e^{2S}(e^S + K\kappa^{1/2})$$

Proof. For any round $t \in [T]$, let $\tau_t \leq t$ denote the last round at which a switch was made. Then, using the value of $\gamma(\delta)$ alongside Lemma 8.8, Lemma 8.12, and Lemma 8.13, we get:

$$\begin{aligned}
R(T) &\leq \sum_{t\in[T]} |\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t^\star, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)| \leq \sum_{t\in[T]} 2\epsilon_1(t, \tau_t, \boldsymbol{x}_t) + 2\epsilon_2(t, \tau_t, \boldsymbol{x}_t) \\
&\leq 4RKd^{1/2}(\log T)^{1/2}\gamma(\delta)\sqrt{T} + 8RKd\log T\kappa^{1/2}e^{3S}\gamma(\delta)^2 + 24dRK^2e^{2S}\kappa\gamma(\delta)^2\log T \\
&\leq CRK^{3/2}S^{5/4}(\log T\log(T/\delta))^{1/2}d\sqrt{T} + CRK^2d^2S^{5/2}\log T\log(T/\delta)\kappa^{1/2}e^{2S}(e^S + K\kappa^{1/2})
\end{aligned}$$

□

8.2 Supporting Lemmas for 8

Lemma 8.1. *Let $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_\tau\}$ be a set of arms and $\{r_1, \dots, r_\tau\}$ be the set of corresponding rewards associated with the arms. Define $\hat{\boldsymbol{\theta}}_\tau$ be the MLE estimate calculated using this set of arms and rewards, i.e*

$$\hat{\boldsymbol{\theta}}_\tau = \arg\min_{\boldsymbol{\theta}} \sum_{s\in[\tau]} \sum_{i=1}^{K} \mathbb{1}\{y_s = i\} \log z_i(\boldsymbol{x}_s, \boldsymbol{\theta}) + \frac{\lambda}{2}\|\boldsymbol{\theta}\|_2^2$$

Let $\boldsymbol{H}_\tau^\star$ be as defined in Section 8.1. Then, with high probability, and the choice of $\lambda = KdS^{-1/2}\log(T/\delta)$, we have that

$$\|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\|_{\boldsymbol{H}_\tau^\star} \leq CS^{5/4}\sqrt{Kd\log(T/\delta)}$$

Proof. We define $G_\tau(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ as:

$$G_\tau(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \sum_{t\in[\tau]} M(\boldsymbol{x}_t, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \lambda \boldsymbol{I}$$

where $M(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ is as defined in Section 8.1. Thus, from Lemma 9.2, we have that

$$(1+2S)^{-1}\boldsymbol{G}_\tau \succcurlyeq \boldsymbol{H}_\tau^\star$$

Thus, we have

$$
\begin{aligned}
\left\|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right\|_{\boldsymbol{H}_\tau^\star} &\leq \sqrt{1+2S}\left\|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right\|_{\boldsymbol{G}_\tau(\hat{\boldsymbol{\theta}}_\tau, \boldsymbol{\theta}^\star)} \\
&\leq \sqrt{1+2S}\left\|\boldsymbol{G}_\tau(\hat{\boldsymbol{\theta}}_\tau, \boldsymbol{\theta}^\star)\left(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right)\right\|_{\boldsymbol{G}_\tau^{-1}(\hat{\boldsymbol{\theta}}_\tau, \boldsymbol{\theta}^\star)} \\
&\overset{(i)}{\leq} \sqrt{1+2S}\left\|\left[\sum_{t\in[\tau]} \boldsymbol{M}(\hat{\boldsymbol{\theta}}_\tau, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \lambda \boldsymbol{I}_{Kd\times Kd}\right]\left(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right)\right\|_{\boldsymbol{G}_\tau^{-1}(\hat{\boldsymbol{\theta}}_\tau, \boldsymbol{\theta}^\star)} \\
&\overset{(ii)}{\leq} \sqrt{1+2S}\left\|\sum_{t\in[\tau]}\left[\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}^\star, \hat{\boldsymbol{\theta}}_\tau) \otimes \boldsymbol{x}_t^\top\right]\left(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau\right) \otimes \boldsymbol{x}_t + \lambda\left(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right)\right\|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \\
&\overset{(iii)}{\leq} \sqrt{1+2S}\left\|\sum_{t\in[\tau]}\left[\boldsymbol{z}(\boldsymbol{x_t}, \hat{\boldsymbol{\theta}}_\tau) - \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right] \otimes \boldsymbol{x}_t + \lambda\left(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right)\right\|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \\
&\overset{(iv)}{\leq} \sqrt{1+2S}\left\|\sum_{t\in[\tau]}\left[\boldsymbol{m}_t - \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right] \otimes \boldsymbol{x}_t - \lambda\boldsymbol{\theta}^\star\right\|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \\
&\overset{(v)}{\leq} \sqrt{1+2S}\left\|\sum_{t\in[\tau]}\left[\boldsymbol{m}_t - \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right] \otimes \boldsymbol{x}_t\right\|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} + \lambda\sqrt{1+2S}\left|\left|\boldsymbol{\theta}^\star\right|\right|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \\
&\leq (1+2S)\left\|\sum_{t\in[\tau]}\left[\boldsymbol{m}_t - \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right] \otimes \boldsymbol{x}_t\right\|_{\boldsymbol{H}_\tau^{\star\,-1}} + \lambda\sqrt{1+2S}\left|\left|\boldsymbol{\theta}^\star\right|\right|_{\boldsymbol{G}_\tau^{-1}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)} \\
&\overset{(vi)}{\leq} (1+2S)\left\|\sum_{t\in[\tau]} \boldsymbol{\epsilon}_t \otimes \boldsymbol{x}_t\right\|_{\boldsymbol{H}_\tau^{\star\,-1}} + S\sqrt{1+2S}\sqrt{\lambda} \\
&\overset{(vi)}{\leq} 3S\left\|\sum_{t\in[\tau]} \boldsymbol{\epsilon}_t \otimes \boldsymbol{x}_t\right\|_{\boldsymbol{H}_\tau^{\star\,-1}} + \sqrt{3}\lambda^{1/2}S^{3/2}
\end{aligned}
$$

where (i) follows from Lemma 9.2 , (ii) follows from Mixed Product Property , (iii) follows from the Mean value Theorem, (iv) from Lemma 9.3, (v) follows from Cauchy-Schwarz, and (vi) follows from the fact that $\boldsymbol{G}_\tau \geq \lambda \boldsymbol{I}$ and $||\boldsymbol{\theta}||_2 \leq S$.

Note that $\boldsymbol{\epsilon}_t = \boldsymbol{m}_t - \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)$ and since $\mathbb{E}\left[\boldsymbol{m}_t\right] = \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)$, we get $\mathbb{E}\left[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^\top\right] = \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)$. Also, note that $\|\epsilon_t\|_1 \leq \|\boldsymbol{m}_t\|_1 + \|\boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\|_1 \leq 2$. Thus, using Lemma 9.10, we get

$$
\left\|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right\|_{\boldsymbol{H}_\tau^\star} \leq 3S\left(\frac{\sqrt{\lambda}}{4} + \frac{4}{\sqrt{\lambda}}\log\left(\frac{\det \boldsymbol{H}_\tau^{1/2}}{\delta\lambda^{\frac{dK}{2}}}\right) + \frac{4}{\sqrt{\lambda}}Kd\log 2\right) + 2S^{3/2}\lambda^{1/2}
$$

where $\boldsymbol{H}_\tau = \lambda\boldsymbol{I} + \sum_{t\in\tau} A(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top$.

We can calculate $\det \boldsymbol{H}_\tau$ as follows:

$$\begin{aligned}\det \boldsymbol{H}_\tau &\stackrel{(i)}{\leq} \left(\frac{\text{trace } \boldsymbol{H}_\tau}{Kd}\right)^{Kd} \\ &\leq \left(\frac{\text{trace } \lambda \boldsymbol{I} + \text{trace } \sum_{t\in\tau} A(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top}{Kd}\right)^{Kd} \\ &\stackrel{(ii)}{\leq} \left(\frac{\lambda Kd + \tau \|\boldsymbol{x}_t\|_2^2}{Kd}\right)^{Kd} \\ &\stackrel{(iii)}{\leq} \lambda^{Kd}\left(1 + \frac{\tau}{\lambda Kd}\right)^{Kd}\end{aligned}$$

where (i) follows from Lemma 9.11, (ii) follows from the fact that tr $(\boldsymbol{A} \otimes \boldsymbol{B}) = \sum \lambda(\boldsymbol{A})\lambda(\boldsymbol{B})$ and the fact that $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}^\star) \preccurlyeq \boldsymbol{I}$ and the only non-zero eigenvalue of $\boldsymbol{x}_t\boldsymbol{x}_t^\top$ is $\|\boldsymbol{x}_t\|_2^2$, and (iii) follows since $\|\boldsymbol{x}\| \leq 1$.

Thus, we have

$$\begin{aligned}\left|\left|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right|\right|_{\boldsymbol{H}_\tau^\star} &\leq 3S\left(\frac{\sqrt{\lambda}}{4} + \frac{4}{\sqrt{\lambda}}\log\left(\frac{\left(1+\frac{\tau}{\lambda Kd}\right)^{\frac{Kd}{2}}}{\delta}\right) + \frac{4}{\sqrt{\lambda}}Kd\log 2\right) + 2S^{3/2}\lambda^{1/2} \\ &= 3S\left(\frac{\sqrt{\lambda}}{4} + \frac{2Kd}{\sqrt{\lambda}}\log\left(1+\frac{\tau}{\lambda d}\right) + \frac{4}{\sqrt{\lambda}}\log\frac{1}{\delta} + \frac{4}{\sqrt{\lambda}}Kd\log 2\right) + 2S^{3/2}\lambda^{1/2}\end{aligned}$$

Finally, by setting $\lambda = KdS^{-1/2}\log(T/\delta)$ and simplifying the constants, we get that for some appropriately tuned constant C

$$\left|\left|\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}^\star\right|\right|_{\boldsymbol{H}_\tau^\star} \leq CS^{5/4}\sqrt{Kd\log(T/\delta)}$$

□

From here on, we shall use the notation $\gamma(\delta) = CS^{5/4}\sqrt{Kd\log(T/\delta)}$.

Lemma 8.2. *Let $\tilde{\boldsymbol{V}}_t$ and $\boldsymbol{H}_t^\star$ be defined as in Section 8.1. Then, for any round $t \in [T]$, we have that*

$$\tilde{\boldsymbol{V}}_t \preccurlyeq \kappa \boldsymbol{H}_t^\star$$

Proof. From the definition of κ, we have $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}) \succcurlyeq \frac{1}{\kappa}\boldsymbol{I}$. Hence, using the fact that $\kappa \geq 1$, we have

$$\begin{aligned}\tilde{\boldsymbol{V}}_t = \boldsymbol{I}_{K\times K} \otimes \boldsymbol{V}_t &= \boldsymbol{I}_{K\times K} \otimes \left(\lambda \boldsymbol{I}_{d\times d} + \sum_{s\in[t]} \boldsymbol{x}_s\boldsymbol{x}_s^\top\right) \\ &= \lambda \boldsymbol{I}_{Kd\times Kd} + \boldsymbol{I}_{K\times K} \otimes \sum_{s\in[t]} \boldsymbol{x}_t \boldsymbol{x}_t^\top \\ &\preccurlyeq \kappa\lambda \boldsymbol{I}_{Kd\times Kd} + \kappa \sum_{s\in[t]} \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top \\ &\preccurlyeq \kappa \boldsymbol{H}_t^\star\end{aligned}$$

□

Lemma 8.3. *Let $\tilde{\boldsymbol{V}}_t$ and $\boldsymbol{H}_t(\boldsymbol{\theta})$ be defined as in Section 8.1. Then, for any round $t \in [T]$, we have that*

$$\tilde{\boldsymbol{V}}_t \preccurlyeq \kappa \boldsymbol{H}_t(\boldsymbol{\theta})$$

Proof. From the definition of κ, we have $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}) \succcurlyeq \frac{1}{\kappa}\boldsymbol{I}$. Hence, using the fact that $\kappa \geq 1$, we have

$$\begin{aligned}
\tilde{\boldsymbol{V}}_t &= \boldsymbol{I}_{K\times K} \otimes \boldsymbol{V}_t = \boldsymbol{I}_{K\times K} \otimes \left(\lambda \boldsymbol{I}_{d\times d} + \sum_{s\in[t]} \boldsymbol{x}_s \boldsymbol{x}_s^\top\right) \\
&= \lambda \boldsymbol{I}_{Kd\times Kd} + \boldsymbol{I}_{K\times K} \otimes \sum_{s\in[t]} \boldsymbol{x}_t \boldsymbol{x}_t^\top \\
&\preccurlyeq \kappa\lambda \boldsymbol{I}_{Kd\times Kd} + \kappa \sum_{s\in[t]} \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}) \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top \\
&\preccurlyeq \kappa\lambda \boldsymbol{I}_{Kd\times Kd} + \kappa \sum_{s\in[t]} \frac{\boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta})}{B_s(\boldsymbol{x}_s)} \otimes \boldsymbol{x}_t \boldsymbol{x}_t^\top \\
&\preccurlyeq \kappa \boldsymbol{H}_t(\boldsymbol{\theta})
\end{aligned}$$

where the second to last inequality follows since $B_t(\boldsymbol{x}) \geq 1$. □

Lemma 8.4. *Let $1, \tau_1, \ldots, \tau_m$ be the rounds at which a switch occurs, i.e det $\boldsymbol{H}_{\tau_{i+1}}(\hat{\boldsymbol{\theta}}_{\tau_i}) \geq 2$ det $\boldsymbol{H}_{\tau_i}(\hat{\boldsymbol{\theta}}_{\tau_i}) \forall i \in [m]$. Let $\boldsymbol{H}_t(\boldsymbol{\theta})$ and $\boldsymbol{H}_t^\star$ be defined as in Section 8.1. Then, for all $i \in [m]$, we have that*

$$\boldsymbol{H}_{\tau_i}(\hat{\boldsymbol{\theta}}_{\tau_i}) \preccurlyeq \boldsymbol{H}_{\tau_i}^\star$$

Proof. From Lemma 9.1, for some $\boldsymbol{x}$ such that $\|\boldsymbol{x}\| \leq 1$ and some $\tau \in \{\tau_1, \ldots, \tau_m\}$, we have that

$$A(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\tau) \preccurlyeq A(\boldsymbol{x}, \boldsymbol{\theta}^\star) \exp\left(\sqrt{6}\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau)\right\|_2\right)$$

Now, we can bound $\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau)\right\|_2$ as follows:

$$\begin{aligned}
\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_{\tau_t})\right\|_2 &\overset{(i)}{\leq} 2S\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\|_2 \overset{(ii)}{=} 2S\sqrt{\lambda_{max}\left((\boldsymbol{I} \otimes \boldsymbol{x})(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\right)} \\
&\overset{(iii)}{=} 2S\sqrt{\lambda_{max}\left(\boldsymbol{I} \otimes \boldsymbol{x}\boldsymbol{x}^\top\right)} \overset{(iv)}{\leq} 2S
\end{aligned}$$

where (i) uses Cauchy-Schwarz inequality and the fact that $\|\boldsymbol{\theta}\|_2 \leq S$, (ii) uses the definition of the norm as $\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{max}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)}$, (iii) follows from the mixed product property of tensor products, and (iv) follows from the fact that $\lambda_{max}\left(\boldsymbol{A} \otimes \boldsymbol{B}\right) = \lambda_{max}\left(\boldsymbol{A}\right)\lambda_{max}\left(\boldsymbol{B}\right)$ and $\lambda_{max}\left(\boldsymbol{x}\boldsymbol{x}^\top\right) = \|\boldsymbol{x}\|_2^2 \leq 1$.

We can also bound $\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau)\right\|_2$ in the following way (note that the $d-$ dimensional unit ball is represented as $\mathcal{B}_2(d)$):

$$
\begin{aligned}
\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau)\|_2 &= \|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_\tau^{\star\,-1/2}\boldsymbol{H}_\tau^{\star\,1/2}(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_\tau)\|_2 \\
&\overset{(i)}{\leq} \|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_\tau^{\star\,-1/2}\|_2\gamma(\delta) \\
&\overset{(ii)}{\leq} \kappa^{1/2}\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\tilde{\boldsymbol{V}}_\tau^{-1/2}\|_2\gamma(\delta) \\
&\overset{(iii)}{\leq} \kappa^{1/2}\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{I} \otimes \boldsymbol{V}_\tau^{-1/2})\|_2\gamma(\delta) \\
&\overset{(iv)}{=} \kappa^{1/2}\sqrt{\lambda_{max}\left((\boldsymbol{I} \otimes \boldsymbol{V}_\tau^{-1/2})(I \otimes \boldsymbol{x})(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{I} \otimes \boldsymbol{V}_\tau^{-1/2})\right)}\gamma(\delta) \\
&\overset{(v)}{=} \kappa^{1/2}\gamma(\delta)|\boldsymbol{x}\|_{\boldsymbol{V}_\tau^{-1}} \\
&\leq 2\kappa^{1/2}\gamma(\delta)|\boldsymbol{x}\|_{\boldsymbol{V}_\tau^{-1}}
\end{aligned}
$$

where (i) is obtained from the fact that $\|\boldsymbol{A}\boldsymbol{x}\|_2 \leq \|\boldsymbol{A}\|_2\|\boldsymbol{x}\|_2$ and from Lemma 8.1, (ii) follows from Lemma 8.2, (iii) is obtained from the definition of $\tilde{\boldsymbol{V}}$ and the fact that $(\boldsymbol{A} \otimes \boldsymbol{B})^n = \boldsymbol{A}^n \otimes \boldsymbol{B}^n$, (iv) follows from the definition of the norm, i.e, $\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{max}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)}$, and (v) follows from the cyclic property of eigenvalues and the fact that $\lambda_{max}\left(\boldsymbol{A} \otimes \boldsymbol{B}\right) = \lambda_{max}\left(\boldsymbol{A}\right)\lambda_{max}\left(\boldsymbol{B}\right)$.

Thus, by combining both bounds, we obtain

$$
A(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\tau) \preccurlyeq A(\boldsymbol{x}, \boldsymbol{\theta}^\star)\exp\left(\sqrt{6}\min\left\{\sqrt{2}\kappa^{1/2}\gamma(\delta)|\boldsymbol{x}\|_{\boldsymbol{V}_\tau^{-1}}, 2S\right\}\right)
$$

Let $B_\tau(\boldsymbol{x})$ denote the value $\exp\left(\sqrt{6}\min\left\{\sqrt{2}\kappa^{1/2}\gamma(\delta)|\boldsymbol{x}\|_{\boldsymbol{V}_\tau^{-1}}, 2S\right\}\right)$. Then, we have that

$$
\boldsymbol{H}_\tau^\star = \lambda\boldsymbol{I} + \sum_{s\in[\tau]}\boldsymbol{A}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_s\boldsymbol{x}_s \succcurlyeq \lambda\boldsymbol{I} + \sum_{s\in[\tau]}\frac{\boldsymbol{A}(\boldsymbol{x}_s, \hat{\boldsymbol{\theta}}_\tau)}{B_\tau(\boldsymbol{x}_s)} \otimes \boldsymbol{x}_s\boldsymbol{x}_s = \boldsymbol{H}_\tau(\hat{\boldsymbol{\theta}}_\tau)
$$

□

Lemma 8.5. *For time round t, let $\tau_t \leq t$ be the last time round at which a switch occurred, i.e det $\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t}) \leq 2$ det $\boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})$. Let $\boldsymbol{H}_t(\boldsymbol{\theta})$ and $\boldsymbol{H}_t^\star$ be defined as in Section 8.1.*

$$
\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t}) \preccurlyeq \boldsymbol{H}_t^\star
$$

Proof. Similar to Lemma 8.4 for some $\boldsymbol{x}$ such that $\|\boldsymbol{x}\| \leq 1$, we have that

$$
A(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_{\tau_t}) \preccurlyeq A(\boldsymbol{x}, \boldsymbol{\theta}^\star)\exp\left(\sqrt{6}\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_{\tau_t})\right\|_2\right)
$$

Now, we can bound $\left\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_{\tau_t})\right\|_2$ in two different ways: the first way results in $2S$, following the same method as Lemma 8.4. We can also bound it in the following way:

$$\begin{aligned}
\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_{\tau_t})\|_2 &= \|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}^{\star}_{\tau_t}{}^{-1/2}\boldsymbol{H}^{\star}_{\tau_t}{}^{1/2}(\boldsymbol{\theta}^\star - \hat{\boldsymbol{\theta}}_{\tau_t})\|_2 \\
&\overset{(i)}{\leq} \|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}^{\star}_{\tau_t}{}^{-1/2}\|_2\gamma(\delta) \\
&\overset{(ii)}{\leq} \|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\|_2\gamma(\delta) \\
&\overset{(iii)}{\leq} \sqrt{2}\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_{t}(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\|_2\gamma(\delta) \\
&\overset{(iv)}{\leq} \sqrt{2}\kappa^{1/2}\|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\tilde{\boldsymbol{V}}^{-1/2}\|_2\gamma(\delta) \\
&\leq 2\kappa^{1/2}\gamma(\delta)|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}}
\end{aligned}$$

where (i) is obtained from the fact that $\|\boldsymbol{A}\boldsymbol{x}\|_2 \leq \|\boldsymbol{A}\|_2\|\boldsymbol{x}\|_2$ and from Lemma 8.1, (ii) follows from Lemma 8.4, (iii) follows from the combination of Lemma 9.13 and the fact that det $\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t}) \leq$ 2 det $\boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})$, (iv) follows from Lemma 8.3, and (v) follows from the same steps used in Lemma 8.6.

Combining the bounds in the same manner as Lemma 8.4 finishes the proof.

□

Lemma 8.6. *For time round t, let $\tau_t \leq t$ be the last time round at which a switch occurred. Let $\boldsymbol{H}_t^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})$ and $\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})$ be defined as in Section 8.1. Then, we have*

$$\boldsymbol{H}_t^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t}) \preccurlyeq \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})$$

Proof. We have:

$$\begin{aligned}
\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t}) &= \lambda \boldsymbol{I} + \sum_{s \in [t]} \tilde{\boldsymbol{X}}_s(\hat{\boldsymbol{\theta}}_{\tau_t})\tilde{\boldsymbol{X}}_s(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \\
&\overset{(i)}{=} \lambda \boldsymbol{I} + \sum_{s \in [t]} \sum_{i=1}^{K} \tilde{\boldsymbol{x}}_s^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})\tilde{\boldsymbol{x}}_s^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \\
&\succcurlyeq \lambda \boldsymbol{I} + \sum_{s \in [t]} \tilde{\boldsymbol{x}}_s^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})\tilde{\boldsymbol{x}}_s^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \\
&= \boldsymbol{H}_\tau^{i}(\hat{\boldsymbol{\theta}}_{\tau_t})
\end{aligned}$$

where (i) follows from Lemma 7.11. □

Lemma 8.7. *Let $\tau_t \leq t$ be the last time round at which a switch was made. In other words, det $\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t}) \leq 2$ det $\boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})$. Then, for any arm $\boldsymbol{x}$, we have that,*

$$\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_\tau)\right| \leq \epsilon_1(t, \tau_t, \boldsymbol{x}) + \epsilon_2(t, \tau_t, \boldsymbol{x})$$

where

$$\epsilon_1(t, \tau_t, \boldsymbol{x}) = \sqrt{2}\gamma(\delta)\left|\left|\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}(\boldsymbol{I} \otimes \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{\rho}\right|\right|_2$$

$$\epsilon_2(t, \tau_t, \boldsymbol{x}) = 6R\gamma(\delta)^2\left|\left|(\boldsymbol{I} \otimes \boldsymbol{x}^\top)\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2$$

Proof. The proof follows on the same lines as that of Lemma 7.5 and uses the fact that $\frac{\det \boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}}{\det \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}} \leq 2$ combined with Lemma 9.13 to convert $\boldsymbol{H}_{\tau_t}(\hat{\boldsymbol{\theta}}_{\tau_t})$ to $\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})$. □

Lemma 8.8. *Let $\tau_t \leq t$ be the last time step at which a switch was made. Let $\epsilon_1(t, \tau_t, \boldsymbol{x})$ and $\epsilon_2(t, \tau_t, \boldsymbol{x})$ be as defined in Lemma 8.7. Then, the regret at time step t can be bounded as*

$$\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}^*, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right| \leq 2\epsilon_1(t, \tau_t, \boldsymbol{x}_t) + 2\epsilon_2(t, \tau_t, \boldsymbol{x}_t)$$

Proof.

$$\begin{aligned}\left|\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}^\star, \boldsymbol{\theta}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star)\right| &\overset{(i)}{\leq} \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}^\star, \hat{\boldsymbol{\theta}}_{\tau_t}) + \epsilon_1(t, \tau_t, \boldsymbol{x}^\star) + \epsilon_2(t, \tau_t, \boldsymbol{x}^\star) - \boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}_t, \hat{\boldsymbol{\theta}}_{\tau_t}) + \epsilon_1(t, \tau_t, \boldsymbol{x}_t) + \epsilon_2(t, \tau_t, \boldsymbol{x}_t) \\ &\overset{(ii)}{\leq} 2\epsilon_1(t, \tau_t, \boldsymbol{x}_t) + 2\epsilon_2(t, \tau_t, \boldsymbol{x}_t)\end{aligned}$$

where (i) follows from Lemma 8.7 and (ii) follows from the fact that $\boldsymbol{x}_t = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \mathrm{UCB}(t, \tau_t, \boldsymbol{x}) = \arg\max_{\boldsymbol{x} \in \mathcal{X}} \left[\boldsymbol{\rho}^\top \boldsymbol{z}(\boldsymbol{x}, \hat{\boldsymbol{\theta}}_{\tau_t}) + \epsilon_1(t, \tau_t, \boldsymbol{x}) + \epsilon_2(t, \tau_t, \boldsymbol{x})\right]$ and hence, gets selected at round t. □

Lemma 8.9. *Let $B_t(\boldsymbol{x})$ be as defined in Section 8.1. Then, we have that*

$$\sqrt{B_t(\boldsymbol{x})} \leq e^{3S} \kappa^{1/2} \gamma(\delta) \|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}} + 1$$

Proof.

$$\begin{aligned}\sqrt{B_t(\boldsymbol{x})} &= \exp\left(\sqrt{6} \min\left\{\kappa^{1/2} \gamma(\delta) \|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}}, S\right\}\right) \\ &\overset{(i)}{\leq} e^{3S} \kappa^{1/2} \gamma(\delta) \|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}} + 1\end{aligned}$$

where (i) follows from Lemma 9.6 by choosing $\min\left\{\kappa^{1/2}\gamma(\delta)\|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}}, S\right\} = \kappa^{1/2}\gamma(\delta)\|\boldsymbol{x}\|_{\boldsymbol{V}_t^{-1}}$ and $M = \sqrt{6}S$.

□

Lemma 8.10. *Let $\tilde{\boldsymbol{X}}_\tau(\boldsymbol{\theta})$ and $\tilde{\boldsymbol{x}}_\tau^{(i)}(\boldsymbol{\theta})$ be defined as in Section 8.1. Then, we have*

$$\tilde{\boldsymbol{X}}_\tau(\boldsymbol{\theta}) \tilde{\boldsymbol{X}}_\tau(\boldsymbol{\theta})^\top = \sum_{i=1}^{K} \tilde{\boldsymbol{x}}_\tau^{(i)}(\boldsymbol{\theta}) \tilde{\boldsymbol{x}}_\tau^{(i)}(\boldsymbol{\theta})^\top$$

Proof. The proof follows on the same lines as Lemma 7.11. □

Lemma 8.11. *Let $\boldsymbol{M} \in \mathbb{R}^{Kd}$ be any positive-semidefinite matrix. Then,*

$$\lambda_{max}\left(\tilde{\boldsymbol{X}}_\tau(\boldsymbol{\theta})^\top \boldsymbol{M} \tilde{\boldsymbol{X}}_\tau(\boldsymbol{\theta})\right) \leq \sum_{i=1}^{K} \left|\left|\tilde{\boldsymbol{x}}_\tau^{(i)}(\boldsymbol{\theta})\right|\right|_{\boldsymbol{M}}^2$$

Proof. The proof follows on the same lines as Lemma 7.12. □

Lemma 8.12. *Let $\epsilon_1(t,\tau,\boldsymbol{x})$ be as defined in Lemma 8.7, and τ_t be the last switching round before round t. Then, we have that*

$$\sum_{t\in[T]} \epsilon_1(t,\tau_t,\boldsymbol{x}_t) \le 8RKd\log T\kappa^{1/2}e^{3S}\gamma(\delta)^2 + 4RKd^{1/2}(\log T)^{1/2}\gamma(\delta)\sqrt{T}$$

Proof.

$$\begin{aligned}
\sum_{t\in[T]} \epsilon_1(t,\tau_t,\boldsymbol{x}_t) &= \sqrt{2}\gamma(\delta)\sum_{t\in[T]} \left|\left|\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}(\boldsymbol{I}\otimes\boldsymbol{x}_t)\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{\rho}\right|\right|_2 \\
&\overset{(i)}{\le} \sqrt{2}\gamma(\delta)\sum_{t\in[T]} \left|\left|\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}(\boldsymbol{I}\otimes\boldsymbol{x}_t)\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_{\tau_t})^{1/2}\right|\right|_2 ||\boldsymbol{\rho}||_{\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_{\tau_t})} \\
&\le \sqrt{2}R\gamma(\delta)\sum_{t\in[T]} \left|\left|\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_{\tau_t})^{1/2}(\boldsymbol{I}\otimes\boldsymbol{x}_t^\top)\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 \\
&\overset{(ii)}{\le} \sqrt{2}R\gamma(\delta)\sum_{t\in[T]} \left|\left|\sqrt{B_t(\boldsymbol{x}_t)}\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 \\
&\overset{(iii)}{\le} \sqrt{2}R\gamma(\delta)\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 \left\{e^{3S}\kappa^{1/2}\gamma(\delta)\|\boldsymbol{x}_t\|_{\boldsymbol{V}_t^{-1}}+1\right\}
\end{aligned}$$

where (i) follows from $||\boldsymbol{A}\boldsymbol{x}||_2 \le ||\boldsymbol{A}||_2\,||\boldsymbol{x}||_2$, (ii) follows from the definition of $\tilde{\boldsymbol{X}}(\boldsymbol{\theta})$, and (iii) follows from Lemma 8.9.

We now bound the term $\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2$:

$$\begin{aligned}
\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 &= \sum_{t\in[T]} \sqrt{\lambda_{max}\left(\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right)} \\
&= \sum_{t\in[T]} \sqrt{\lambda_{max}\left(\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\right)} \\
&\overset{(i)}{=} \sum_{t\in[T]} \sqrt{\sum_{i=1}^{K}\left|\left|\tilde{\boldsymbol{x}}_t^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})\right|\right|^2_{\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}}} \\
&\overset{(ii)}{\le} \sum_{t\in[T]} \sqrt{\sum_{i=1}^{K}\left|\left|\tilde{\boldsymbol{x}}_t^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})\right|\right|^2_{\boldsymbol{H}^i_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}}} \\
&\overset{(iii)}{\le} \sqrt{T}\sqrt{\sum_{t\in[T]}\sum_{i=1}^{K}\left|\left|\tilde{\boldsymbol{x}}_t^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_t})\right|\right|^2_{\boldsymbol{H}^i_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1}}} \\
&\overset{(iv)}{\le} 2K\sqrt{dT\log T}
\end{aligned}$$

where (i) follows from Lemma 8.11, (ii) follows from Lemma 8.6, (iii) follows from Cauchy-Schwarz, and (iv) follows from Lemma 9.12 and the fact that $\left|\left|\tilde{\boldsymbol{x}}^{(i)}(\boldsymbol{\theta})\right|\right|_2 \le ||\boldsymbol{A}(\boldsymbol{x},\boldsymbol{\theta})||_2\,||\boldsymbol{x}||_2 \le 1$.

We also bound the term $\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 \|\boldsymbol{x}_t\|_{\boldsymbol{V}_t^{-1}}$ as follows:

$$\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2 \|\boldsymbol{x}_t\|_{\boldsymbol{V}_t^{-1}} \overset{(i)}{\leq} \sqrt{\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2} \sqrt{\sum_{t\in[T]} \|\boldsymbol{x}_t\|^2_{\boldsymbol{V}_t^{-1}}}$$
$$\overset{(ii)}{\leq} 2K\sqrt{d\log T}\sqrt{\sum_{t\in[T]} \|\boldsymbol{x}_t\|^2_{\boldsymbol{V}_t^{-1}}}$$
$$\overset{(ii)}{\leq} 4Kd\log T$$

where (i) follows from Cauchy-Schwarz, (ii) follows from the same steps used to bound $\sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^\top \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2$ above, and (iii) follows from Lemma 9.12.

Substituting back, we get

$$\sum_{t\in[T]} \epsilon_1(t,\tau_t,\boldsymbol{x}_t) \leq 4\sqrt{2}RKd\log T\kappa^{1/2}e^{3S}\gamma(\delta)^2 + 2\sqrt{2}RKd^{1/2}(\log T)^{1/2}\gamma(\delta)\sqrt{T}$$
$$\leq 8RKd\log T\kappa^{1/2}e^{3S}\gamma(\delta)^2 + 4RKd^{1/2}(\log T)^{1/2}\gamma(\delta)\sqrt{T}$$

□

Lemma 8.13. *Let $\epsilon_2(t,\tau,\boldsymbol{x})$ be as defined in Lemma 8.7, and τ_t be the last switching round before round t. Then, we have that*

$$\sum_{t\in[T]} \epsilon_2(t,\tau_t,\boldsymbol{x}_t) \leq 24dRK^2e^{2S}\kappa\gamma(\delta)^2\log T$$

Proof.

$$\sum_{t\in[T]} \epsilon_2(t,\tau,\boldsymbol{x}_t) = 6R\gamma(\delta)^2 \sum_{t\in[T]} \left|\left|(\boldsymbol{I}\otimes\boldsymbol{x}^\top)\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2$$
$$\overset{(i)}{=} 6R\gamma(\delta)^2 \sum_{t\in[T]} \left|\left|\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_\tau)^{-1/2}\right|\right|_2 \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2 B_t(\boldsymbol{x}_t)$$
$$\overset{(ii)}{\leq} 6R\gamma(\delta)^2e^{2S} \sum_{t\in[T]} \left|\left|\boldsymbol{A}(\boldsymbol{x}_t,\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2 \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2$$
$$\overset{(iii)}{\leq} 6R\gamma(\delta)^2e^{2S}\kappa \sum_{t\in[T]} \left|\left|\tilde{\boldsymbol{X}}_t(\hat{\boldsymbol{\theta}}_{\tau_t})\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_t})^{-1/2}\right|\right|_2^2$$
$$\overset{(iv)}{\leq} 24dRK^2e^{2S}\kappa\gamma(\delta)^2\log T$$

where (i) follows from the definition of $\tilde{\boldsymbol{X}}$ and Lemma 8.6, (ii) follows from the definition of $B_t(\boldsymbol{x})$, (iii) follows from the fact that $\boldsymbol{A}(\boldsymbol{x},\boldsymbol{\theta}) \leq \frac{1}{\kappa}\boldsymbol{I}$, and (iv) follows from the methods used in Lemma 8.12. □

Lemma 8.14. *Let Algorithm 3 be run for t rounds. Then, the switching criterion is triggered a maximum of $dK\log(1+\frac{t}{d\lambda})$ times.*

Proof. Let $\tau_0,\tau_1,\ldots,\tau_m \in [1,t]$ be the time steps at which the switching criterion in Algorithm 3 is triggered, i.e, 2 det $\boldsymbol{H}_{\tau_i}(\hat{\boldsymbol{\theta}}_{\tau_i}) \leq$ det $\boldsymbol{H}_{\tau_{i+1}}(\hat{\boldsymbol{\theta}}_{\tau_i})$ for $i\in[m-1]$, and $\tau_m = t$. Note that $\boldsymbol{H}_{\tau_0}(\boldsymbol{\theta}) = \lambda\boldsymbol{I}_{Kd\times Kd}$.

$$
\begin{aligned}
\frac{\det \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})}{\det \boldsymbol{H}_{\tau_0}(\boldsymbol{\theta})} &= \frac{\det \boldsymbol{H}_{\tau_m}(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})}{\det \boldsymbol{H}_{\tau_{m-1}}(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})} \times \frac{\det \boldsymbol{H}_{\tau_{m-1}}(\hat{\boldsymbol{\theta}}_{\tau_{m-2}})}{\det \boldsymbol{H}_{\tau_{m-2}}(\hat{\boldsymbol{\theta}}_{\tau_{m-2}})} \times \ldots \times \frac{\det \boldsymbol{H}_{\tau_1}(\hat{\boldsymbol{\theta}}_{\tau_0})}{\det \boldsymbol{H}_{\tau_0}(\boldsymbol{\theta})} \\
&\geq 2^m
\end{aligned}
$$

and hence, $\det \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_{m-1}}) \geq 2^m \lambda^{Kd}$ since $\det \boldsymbol{H}_1 = \lambda^{Kd}$. Also, we can say that:

$$
\begin{aligned}
\det \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_{m-1}}) &\overset{(i)}{\leq} \left(\frac{\text{trace } \boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})}{Kd}\right)^{Kd} \\
&\overset{(ii)}{\leq} \left(\frac{\sum_{i\in[K]} \text{trace } \boldsymbol{H}_t^i(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})}{Kd}\right)^{Kd} \\
&\overset{(iii)}{\leq} \left(\frac{\lambda Kd + \sum_{i\in[K]} \sum_{s\in[t]} \|\tilde{\boldsymbol{x}}_s^{(i)}(\hat{\boldsymbol{\theta}}_{\tau_{m-1}})\|_2^2}{Kd}\right)^{Kd} \\
&\overset{(iv)}{\leq} \left(\lambda + \frac{t}{d}\right)^{Kd}
\end{aligned}
$$

Here (i) follows from Lemma 9.11, (ii) follows from Lemma 8.6 alongside the linearity of the trace operator, (iii) follows from the definition of $\boldsymbol{H}_t^i(\boldsymbol{\theta})$ and the fact that the only non-zero eigenvalue of $\boldsymbol{x}\boldsymbol{x}^\top$ is $\|\boldsymbol{x}\|_2^2$, and (iv) is due to the fact that $\|\tilde{\boldsymbol{x}}_t^{(i)}(\boldsymbol{\theta})\|_2 \leq \|\boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta})\| \leq 1$. Thus, we have

$$
2^m \lambda^{Kd} \leq \det(\boldsymbol{H}_t(\hat{\boldsymbol{\theta}}_{\tau_{m-1}}) \leq \left(\lambda + \frac{t}{d}\right)^{Kd}
$$

and hence, $2^m \leq \left(1 + \frac{t}{\lambda d}\right)^{Kd}$. This finishes the proof.

$\square$

9 General Lemmas and Results

Lemma 9.1. *(Self-Concordance) Let $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta}) = \nabla \boldsymbol{z}(\boldsymbol{x}, \boldsymbol{\theta})$. Then, $\boldsymbol{A}(\boldsymbol{x}, \boldsymbol{\theta})$ is $(M, v)-$generalized self-concordant with $v = 1$ and $M = \sqrt{6}$. In other words, for any given $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2$, denote $\boldsymbol{A}_1 = \boldsymbol{A}(\boldsymbol{x}_1, \boldsymbol{\theta}_1)$ and $\boldsymbol{A}_2 = \boldsymbol{A}(\boldsymbol{x}_2, \boldsymbol{\theta}_2)$. Then, we have*

$$
\boldsymbol{A}_2 \exp\left(-\sqrt{6}\left|\left|\left(\boldsymbol{I} \otimes \boldsymbol{x}_1^\top\right)\boldsymbol{\theta}_1 - \left(\boldsymbol{I} \otimes \boldsymbol{x}_2^\top\right)\boldsymbol{\theta}_2\right|\right|_2\right) \preccurlyeq \boldsymbol{A}_1 \preccurlyeq \boldsymbol{A}_2 \exp\left(\sqrt{6}\left|\left|\left(\boldsymbol{I} \otimes \boldsymbol{x}_1^\top\right)\boldsymbol{\theta}_1 - \left(\boldsymbol{I} \otimes \boldsymbol{x}_2^\top\right)\boldsymbol{\theta}_2\right|\right|_2\right)
$$

Lemma 9.2. *(Lemma 13, Amani & Thrampoulidis (2021)) Let $\beta = \{t_1, \ldots, t_N\}$ be a set of time indices and define*

$$
\boldsymbol{G}_\beta(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \sum_{t\in\beta} \boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \lambda \boldsymbol{I}_{Kd\times Kd}
$$

and

$$
\boldsymbol{H}_\beta^\star = \sum_{t\in\beta} \boldsymbol{A}(\boldsymbol{x}_t, \boldsymbol{\theta}^\star) \otimes \boldsymbol{x}_t\boldsymbol{x}_t^\top + \lambda \boldsymbol{I}_{Kd\times Kd}
$$

where

$$
\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \int_0^1 \boldsymbol{A}(\boldsymbol{x}, v\boldsymbol{\theta}_1 + (1-v)\boldsymbol{\theta}_2)\ \mathrm{d}v
$$

Then,

$$\boldsymbol{G}_\beta(\boldsymbol{\theta}, \boldsymbol{\theta}^\star) \succcurlyeq \frac{1}{1+2S}\boldsymbol{H}_\beta^\star$$

Lemma 9.3. *Define the log-likelihood function as follows:*

$$\mathcal{L}_t(\boldsymbol{\theta}) = \sum_{s=1}^{t-1}\sum_{i=1}^{K} \mathbb{1}\left\{y_s = i\right\} \log \frac{1}{\boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})} + \frac{\lambda}{2}\|\boldsymbol{\theta}\|_2^2$$

Let $\hat{\boldsymbol{\theta}}$ be the MLE of $\boldsymbol{\theta}^\star$, i.e., $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{L}_t(\boldsymbol{\theta})$, then

$$\sum_{s=1}^{t-1} \boldsymbol{z}(\boldsymbol{x}_s, \hat{\boldsymbol{\theta}}) \otimes \boldsymbol{x}_s + \lambda\hat{\boldsymbol{\theta}} = \sum_{s=1}^{t-1} \boldsymbol{m}_s \otimes \boldsymbol{x}_s$$

where $\boldsymbol{m}_s = \left(\mathbb{1}\left\{y_s = 1\right\}, \ldots, \mathbb{1}\left\{y_s = K\right\}\right)^\top$ is the user-response vector.

Proof. For the sake of convenience, define the loss incurred at round t (without the regularization term) as

$$l_t(\boldsymbol{\theta}) = \sum_{i=1}^{K} \mathbb{1}\left\{y_s = i\right\} \log \frac{1}{\boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})}$$

Then, it is easy to see that

$$\begin{aligned}
\frac{\partial l_t(\boldsymbol{\theta})}{\partial \theta_m} &= -\sum_{i=1}^{K} \mathbb{1}\left\{y_s = i\right\} \frac{1}{\boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})} \frac{\partial \boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})}{\partial \theta_m} \\
&= -\sum_{i=1}^{K} \mathbb{1}\left\{y_s = i\right\} \frac{1}{\boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})} \left[\mathbb{1}\left\{i = m\right\} \boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta}) - \boldsymbol{z}_i(\boldsymbol{x}_s, \boldsymbol{\theta})\boldsymbol{z}_m(\boldsymbol{x}_s, \boldsymbol{\theta})\right] \otimes \boldsymbol{x}_s \\
&= \left[\mathbb{1}\left\{y_s = m\right\} - \boldsymbol{z}_m(\boldsymbol{x}_s, \boldsymbol{\theta})\right] \otimes \boldsymbol{x}_s
\end{aligned}$$

and hence,

$$\nabla l_t(\boldsymbol{\theta}) = \left[\boldsymbol{m}_s - \boldsymbol{z}(\boldsymbol{x}_s, \boldsymbol{\theta})\right] \otimes \boldsymbol{x}_s$$

Since $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathcal{L}_t(\boldsymbol{\theta})$, we have that $\nabla\mathcal{L}_t(\hat{\boldsymbol{\theta}}) = \arg\min_{\boldsymbol{\theta}} \sum_{s=1}^{t-1} l_s(\hat{\boldsymbol{\theta}}) + \lambda\hat{\boldsymbol{\theta}} = 0$, which results in the claim.

□

Lemma 9.4. *(Bernstein's Inequality) Let $X_1, \ldots, X_n$ be a sequence of independent random variables with $|X_t - \mathbb{E}[X_t]| \leq b$. Let $S = \sum_{t=1}^{n} (X_t - \mathbb{E}[X_t])$ and $v = \sum_{t=1}^{n} \mathbb{V}[X_t]$. Then, for any $\delta \in [0, 1]$, we have*

$$\mathbb{P}\left\{S \geq \sqrt{2v\log\frac{1}{\delta}} + \frac{2b}{3}\log\frac{1}{\delta}\right\} \leq \delta$$

Lemma 9.5. *Let $\boldsymbol{m}_s = \left(\mathbb{1}\left\{y_s = 1\right\}, \ldots, \mathbb{1}\left\{y_s = K\right\}\right)$ be the user-response vector as defined in Section 7.1. Then,*

$$\mathbb{E}\left[\boldsymbol{m}_s\right] = \boldsymbol{z}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star) \text{ and } \mathbb{E}\left[\boldsymbol{m}_s\boldsymbol{m}_s^\top\right] = \mathit{diag}(\boldsymbol{z}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star))$$

Proof. Since $\boldsymbol{m}_s = (\mathbb{1}\{y_s = 1\}, \ldots, \mathbb{1}\{y_s = K\})$, we have

$$\mathbb{E}[\boldsymbol{m}_s] = (\mathbb{E}[\mathbb{1}\{y_s = 1\}], \ldots, \mathbb{E}[\mathbb{1}\{y_s = K\}]) = (z_1(\boldsymbol{x}_s, \boldsymbol{\theta}^\star), \ldots, z_K(\boldsymbol{x}_s, \boldsymbol{\theta}^\star)) = \boldsymbol{z}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star)$$

For the second part, note that

$$\left[\boldsymbol{m}_s \boldsymbol{m}_s^\top\right]_{i,j} = \mathbb{1}\{y_s = i\}\,\mathbb{1}\{y_s = j\} = \begin{cases} \mathbb{1}\{y_s = i\} & i = j \\ 0 & i \neq j \end{cases}$$

Thus, we have

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{m}_s \boldsymbol{m}_s^\top\right] &= \mathbb{E}\left[\operatorname{diag}\left(\mathbb{1}\{y_s = 1\}, \ldots, \mathbb{1}\{y_s = K\}\right)\right] = \operatorname{diag}\left(\mathbb{E}[\mathbb{1}\{y_s = 1\}], \ldots, \mathbb{E}[\mathbb{1}\{y_s = K\}]\right) \\ &= \operatorname{diag}\left(z_1(\boldsymbol{x}_s, \boldsymbol{\theta}^\star), \ldots, z_K(\boldsymbol{x}_s, \boldsymbol{\theta}^\star)\right) = \operatorname{diag}(\boldsymbol{z}(\boldsymbol{x}_s, \boldsymbol{\theta}^\star))\end{aligned}$$

$\square$

Lemma 9.6. *(Claim A.8, Sawarni et al. (2024)) For any $x \in [0, M]$,*

$$e^x \leq e^M \left(\frac{x}{M}\right) + 1$$

Lemma 9.7. *(Theorem 5, Ruan et al. (2021)) Let π represent the G-Optimal Distributional Design learnt from $\mathcal{X}_1 \ldots \mathcal{X}_s \stackrel{i.i.d}{\sim} \mathcal{D}$ and let $\boldsymbol{W}$ be the expected data matrix, i.e. $\boldsymbol{W} = \lambda \boldsymbol{I} + \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\sim\pi(\mathcal{X})}{\mathbb{E}} \boldsymbol{x}\boldsymbol{x}^\top | \mathcal{X}\right]$, then, we have*

$$P\left\{\underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||_{\boldsymbol{W}^{-1}}\right] \leq O(\sqrt{d\log d})\right\} \geq 1 - \exp\left(O(d^4 \log^2 d) - sd^{-12}2^{-16}\right)$$

Lemma 9.8. *(Lemma 4, Ruan et al. (2021)) Let π_G represent the G-Optimal design and define the design matrix $\boldsymbol{W}_G = \lambda \boldsymbol{I} + \underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\underset{\boldsymbol{x}\in\pi_G(\mathcal{X})}{\mathbb{E}} \boldsymbol{x}\boldsymbol{x}^\top \mid \mathcal{X}\right]$, then we have*

$$\underset{\mathcal{X}\sim\mathcal{D}}{\mathbb{E}}\left[\max_{\boldsymbol{x}\in\mathcal{X}} ||\boldsymbol{x}||^2_{\boldsymbol{W}_G^{-1}}\right] \leq O(d^2)$$

Lemma 9.9. *(Lemma A.15, Sawarni et al. (2024) , Ruan et al. (2021)) Let $\boldsymbol{x}_1 \ldots \boldsymbol{x}_n \sim \mathcal{D}$ be vectors with $||x||_2 \leq 1$, then*

$$\mathbb{P}\left\{3\epsilon N \boldsymbol{I} + \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^\top \not\succcurlyeq \frac{n}{8} \underset{\boldsymbol{x}\sim\mathcal{D}}{\mathbb{E}}\left[\boldsymbol{x}\boldsymbol{x}^\top\right]\right\} \geq 1 - 2d\exp\left(-\frac{\epsilon n}{8}\right)$$

Lemma 9.10. *(Lemma 6, Zhang & Sugiyama (2023)) Let $\{\mathcal{F}_t\}_{t=1}^\infty$ be a filteration and $\{\boldsymbol{x}_t\}_{t=1}^\infty$ be a stochastic process in $\mathcal{B}_2(d) = \left\{\boldsymbol{x} \in \mathbb{R}^d \mid ||\boldsymbol{x}_t||_2 \leq 1\right\}$ such that $\boldsymbol{x}_t$ is $\mathcal{F}_t-$ measurable. Let $\{\boldsymbol{\epsilon_t}\}_{t=1}^\infty$ be a martingale difference sequence such that $\boldsymbol{\epsilon}_t$ is $\mathcal{F}_{t+1}-$measurable. Assume that conditioned on $\mathcal{F}_t$, we have $||\boldsymbol{\epsilon}_t||_1 \leq 2$ almost surely, and is denoted by $\boldsymbol{\eta_t} = \mathbb{E}\left[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^\top \mid \mathcal{F}_t\right]$. Let $\lambda > 0$ and for any $t \geq 1$, define*

$$\boldsymbol{S}_t = \sum_{s=1}^{t-1} \boldsymbol{\epsilon}_s \otimes \boldsymbol{x}_s \text{ and } \boldsymbol{H}_t = \lambda \boldsymbol{I}_{dK\times dK} + \sum_{s=1}^{t-1} \boldsymbol{\eta}_s \otimes \boldsymbol{x}_s \boldsymbol{x}_s^\top$$

Then, for any $\delta \in (0, 1)$, we have

$$\mathbb{P}\left\{\exists t > 1, ||\boldsymbol{S}_t||_{\boldsymbol{H}_t^{-1}} \geq \frac{\sqrt{\lambda}}{4} + \frac{4}{\sqrt{\lambda}} \log\left(\frac{det\, \boldsymbol{H}_t^{1/2}}{\delta \lambda^{\frac{dK}{2}}}\right) + \frac{4}{\sqrt{\lambda}} K d \log 2\right\} \leq \delta$$

Lemma 9.11. *(Determinant-Trace Inequality) Let the determinant and trace of a p.s.d matrix $\boldsymbol{A} \in \mathbb{R}^{d\times d}$ be denoted by det $\boldsymbol{A}$ and trace $\boldsymbol{A}$. Then, we have*

$$\text{det } \boldsymbol{A} \leq \left(\frac{\text{trace } \boldsymbol{A}}{d}\right)^d$$

Proof. Let the eigenvalues of $\boldsymbol{A}$ be denoted by $\lambda(\boldsymbol{A}) \geq 0$ since $\boldsymbol{A} \succcurlyeq 0$. Then, we know, det $\boldsymbol{A} = \prod \lambda(\boldsymbol{A})$ and trace $\boldsymbol{A} = \sum \lambda(\boldsymbol{A})$. Thus, applying the inequality for arithmetic means and geometric means, we get that

$$\left(\prod \lambda(\boldsymbol{A})\right)^{1/d} \leq \frac{\sum \lambda(\boldsymbol{A})}{d} \implies \text{det } \boldsymbol{A} \leq \left(\frac{\text{trace } \boldsymbol{A}}{d}\right)^d$$

□

Lemma 9.12. *(Elliptical Potential Lemma, Lemma 11, Abbasi-Yadkori et al. (2011)) Let $\{\boldsymbol{x}_s\}_{s=1}^t$ represent a set of vectors in $\mathbb{R}^d$ and let $||\boldsymbol{x}_s||_2 \leq L$. Let $\boldsymbol{V}_s = \lambda \boldsymbol{I}_{d\times d} + \sum\limits_{m=1}^{s-1} x_m x_m^\top$. Then, for $\lambda \geq 1$*

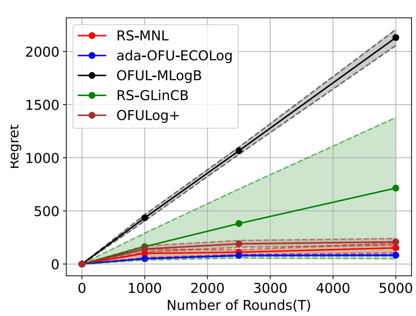
$$\sum_{s=1}^{t} ||\boldsymbol{x}_s||^2_{\boldsymbol{V}_s^{-1}} \leq 2d\log\left(1 + \frac{tL^2}{\lambda d}\right) \leq 4d\log(tL^2)$$

Lemma 9.13. *(Lemma 12, Abbasi-Yadkori et al. (2011)) If $\boldsymbol{A} \succcurlyeq \boldsymbol{B} \succcurlyeq 0$, then*

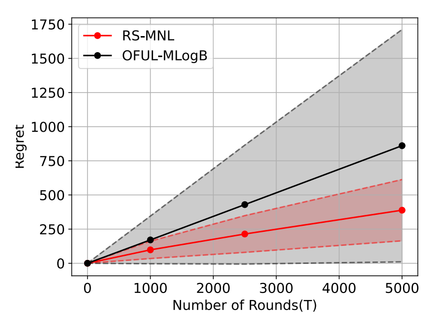
$$\sup_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}} \leq \frac{\text{det }(\boldsymbol{A})}{\text{det }(\boldsymbol{B})}$$

10 Additional Experiments

In this section, we supplement the experiments from Section 5 (in particular, **Experiment 1** and **Experiment 2**).



(a) Regret vs. T: Logistic ($K = 1$) Setting



(b) Regret vs. T: $K = 3$

Experiment 1 ($R(T)$ vs. T for the Logistic ($K = 1$) Setting): In this experiment, we use the same instance as in **Experiment 1** (Section 5) and average the regret over 10 different seeds for sampling rewards. The averaged results with two standard deviations can be found in Figure 2a.

Experiment 2 ($R(T)$ vs. T for $K = 3$): In this experiment, we use the same instance as in **Experiment 2** (Section 5) and average the regret over 10 different seeds for sampling rewards. The averaged results with two standard deviations are reported in Figure 2b.