A Proof of Lemma 3

Proof of Lemma 3. We first prove the upper bound of A_t. The essential proof is actually due to Cheung et al. [2019a] in analyzing sliding window based approach. For self-containedness, we restate here in the notations of our proposed restarted strategy.

$$\begin{aligned}
&\left\| V_{t-1}^{-1}\left(\sum_{s=t_0}^{t-1} X_s X_s^{\mathrm{T}}(\theta_s-\theta_t)\right)\right\|_2\\
=&\left\| V_{t-1}^{-1}\left(\sum_{s=t_0}^{t-1} X_s X_s^{\mathrm{T}}\Big(\sum_{p=s}^{t-1}(\theta_p-\theta_{p+1})\Big)\right)\right\|_2\\
=&\left\| V_{t-1}^{-1}\left(\sum_{p=t_0}^{t-1} \Big(\sum_{s=t_0}^{p} X_s X_s^{\mathrm{T}}(\theta_p-\theta_{p+1})\Big)\right)\right\|_2 \qquad (21)\\
\leq&\sum_{p=t_0}^{t-1}\left\| V_{t-1}^{-1}\Big(\sum_{s=t_0}^{p} X_s X_s^{\mathrm{T}}\Big)(\theta_p-\theta_{p+1})\right\|_2 \qquad (22)\\
\leq&\sum_{p=t_0}^{t-1}\lambda_{\max}\left(V_{t-1}^{-1}\Big(\sum_{s=t_0}^{p} X_s X_s^{\mathrm{T}}\Big)\right)\|\theta_p-\theta_{p+1}\|_2 \qquad (23)\\
\leq&\|\theta_p-\theta_{p+1}\|_2, \qquad (24)
\end{aligned}$$

where (21) holds by rearranging over the index pair of (s,p), (22) holds due to the triangle inequality, (23) and (24) can be obtained by the same argument in Appendix B of Cheung et al. [2019b]. We thus prove the upper bound of A_t.

We proceed to prove the upper bound of B_t. From the self-normalized concentration inequality [Abbasi-Yadkori et al., 2011, Theorem 1], restated in Theorem 5 of Appendix C, we know that

$$\begin{aligned}
&\left\|\sum_{s=t_0}^{t-1}\eta_s X_s\right\|_{V_{t-1}^{-1}}\\
\overset{(32)}{\leq}&\sqrt{2R^2\log\left(\frac{\det(V_{t-1})^{1/2}\det(\lambda I)^{-1/2}}{\delta}\right)}\\
\leq& R\sqrt{2\log\frac{1}{\delta}+d\log\left(1+\frac{(t-t_0)L^2}{d}\right)},
\end{aligned}$$

where the last inequality is obtained from the analysis of the determinant, as shown in the proof of Lemma 4.

Meanwhile, since $V_{t-1}\succeq\lambda I$, we know that

$$\|\lambda\theta_t\|^2_{V_{t-1}^{-1}}\leq 1/\lambda_{\min}(V_{t-1})\|\lambda\theta_t\|_2^2\leq\frac{1}{\lambda}\|\lambda\theta_t\|_2^2\leq\lambda S^2.$$

Therefore, the upper bound of B_t can be immediately obtained by combining the above inequalities. □

B Bandit-over-Bandits Mechanism and Proof of Theorem 4

The RestartUCB algorithm requires prior information of the path-length P_T, which is generally unknown. Such a limitation can be avoided by utilizing the Bandits-over-bandits (BOB) mechanism, proposed by Cheung et al. [2019a] in designing parameter-free algorithm for non-stationary linear bandits based on sliding window least square estimator.

In the following, we first describe how to apply the BOB mechanism to eliminate the requirement of the unknown path-length in RestartUCB. Then, we present the proof of Theorem 4.

B.1 RestartUCB with BOB Mechanism

We name the RestartUCB algorithm with Bandit-over-Bandits mechanism as "RestartUCB-BOB", whose main idea is illustrated in Figure 4.

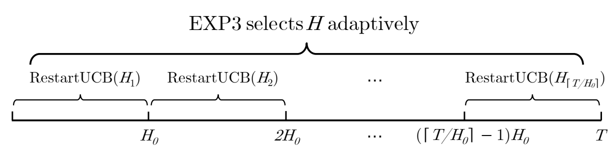

Figure 4: Illustration of Bandit-over-Bandits mechanism with application to RestartUCB algorithm.

From a high-level view, although the exact value of the optimal epoch size (or equivalently, the path-length P_T) is not clear, we can make some random guesses of its possible value, since the P_T is always bounded. Then, we can use a certain meta-algorithm to adaptively track the best epoch size, based on the returned reward returned. Specifically, The RestartUCB-BOB algorithm first sets an update period H_0, and then runs the RestartUCB with a particular epoch size in each period, and the epoch size will be adaptively adjusted by employing EXP3 [Auer et al., 2002] as the meta-algorithm. We refer the reader to Section 7.3 of Cheung et al. [2019b] for more descriptions of design motivations and algorithmic details.

In the configuration of RestartUCB-BOB, we set $H_0=\lceil d\sqrt{T}\rceil$ and the pool of epoch sizes J as

$$J=\{H_i=\lfloor (d/(2S))^{2/3}\cdot 2^{i-1}\rfloor \mid i=1,2,\cdots,N\},$$

where $N=\lceil\ln(d^{1/3}T^{1/2}(2S)^{2/3})\rceil+1$.

Denoted by $H_{\min}$ ($H_{\max}$) the minimal (maximal) epoch size in the pool J, we know that

$$H_{\min}=\lfloor (d/(2S))^{2/3}\rfloor,\ H_{\max}=\lfloor d\sqrt{T}\rfloor\leq H_0. \qquad (25)$$

B.2 Proof of Theorem 4

Proof of Theorem 4. We begin with the following decomposition of the dynamic regret.

$$\begin{aligned}
&\sum_{t=1}^{T}\langle X_t^*,\theta_t\rangle - \langle X_t,\theta_t\rangle \\
=& \underbrace{\sum_{t=1}^{T}\langle X_t^*,\theta_t\rangle - \sum_{i=1}^{\lceil T/H_0\rceil}\sum_{t=(i-1)H_0+1}^{iH_0}\langle X_t(H^\dagger),\theta_t\rangle}_{\texttt{term (i)}} \\
&+ \underbrace{\sum_{i=1}^{\lceil T/H_0\rceil}\sum_{t=(i-1)H_0+1}^{iH_0}\langle X_t(H^\dagger),\theta_t\rangle - X_t(H_i),\theta_t\rangle,}_{\texttt{term (ii)}}
\end{aligned}$$

where $H^\dagger$ is the best epoch size to approximate the optimal epoch size H^* in the pool J, and $H^* = \lfloor (dT/(1+P_T))^{2/3}\rfloor$. Hence, it suffices to bound terms (i) and (ii). In the following, we consider two cases, either $(1+P_T)\geq d^{-1/2}T^{1/4}$ or $(1+P_T) < d^{-1/2}T^{1/4}$.

Case 1. when $(1+P_T)\geq d^{-1/2}T^{1/4}$.

In this case, it is easy to verify that $H^*\leq H_{\max}$ and we thus conclude that H^* lies in the the range of $[H_{\min}, H_{\max}]$. Furthermore, from the configuration of the pool J, we confirm that there exists an epoch size $H^\dagger\in J$ such that $H^\dagger \leq H^* \leq 2H^\dagger$. So term (ii) can be upper bounded by

$$\begin{aligned}
\texttt{term (ii)} &\leq \sum_{i=1}^{\lceil T/H_0\rceil}\widetilde{O}\Big(H^\dagger P_i + \frac{dH_0}{\sqrt{H^\dagger}}\Big) &(26)\\
&= \widetilde{O}\Big(H^\dagger P_T + \frac{dT}{\sqrt{H^\dagger}}\Big) &(27)\\
&\leq \widetilde{O}\Big(H^* P_T + \frac{dT}{\sqrt{2H^*}}\Big)\\
&= \widetilde{O}(d^{2/3}P_T^{1/3}T^{2/3}),
\end{aligned}$$

where (26) is due to Theorem 2 and P_i denotes the path-length in the i-th update period. (27) follows by summing over all update periods, and the last inequality holds since the optimal epoch size H^* is provably in the range of $[H_{\min}, H_{\max}]$ and satisfies $H^\dagger\leq H^*\leq 2H^\dagger$.

Next, we bound the term (i),

$$\begin{aligned}
\texttt{term (i)} &\leq \widetilde{O}(\sqrt{H_0 NT})\\
&\leq \widetilde{O}(d^{1/2}T^{3/4}) &(28)\\
&\leq \widetilde{O}(d^{2/3}T^{2/3}(1+P_T)^{1/3}),
\end{aligned}$$

where the first inequality follows by the same argument as in the sliding window based approach [Cheung et al., 2019b, Lemma 13], building upon the of EXP3. In addition, the last inequality holds due to the fact that $(1+P_T)\geq d^{-1/2}T^{1/4}$ implies,

$$d^{1/2}T^{3/4} = d^{2/3}T^{2/3}d^{-1/3}T^{1/6} \leq d^{2/3}T^{2/3}(1+P_T)^{1/3}.$$

Hence, by combining the upper bounds of term (i) and term (ii), we know that the dynamic regret of RestartUCB-BOB is bounded by $\widetilde{O}(d^{2/3}T^{2/3}(1+P_T)^{1/3})$ under the condition of $(1+P_T)\geq d^{-1/2}T^{1/4}$.

Case 2. when $(1+P_T) < d^{-1/2}T^{1/4}$.

In this case, we cannot guarantee that the optimal epoch size H^* lies in the range of $[H_{\min}, H_{\max}]$, so we set $H^\dagger$ as H_0,

$$\begin{aligned}
\texttt{term (ii)} &\leq \widetilde{O}\Big(H^\dagger P_T + \frac{dT}{\sqrt{H^\dagger}}\Big)\\
&\leq \widetilde{O}\Big(H_0 P_T + \frac{dT}{\sqrt{H_0}}\Big)\\
&= \widetilde{O}\Big(d\sqrt{T}P_T + d^{1/2}T^{3/4}\Big)\\
&\leq \widetilde{O}\Big(d^{1/2}T^{3/4}\Big)
\end{aligned}$$

where the last inequality holds by exploiting the condition of $(1+P_T)\leq d^{-1/2}T^{1/4}$. The result in conjunction with the upper bound of term (i) in (28) gives the $\widetilde{O}(d^{1/2}T^{3/4})$ dynamic regret under this condition.

Finally, note that the dynamic regret of above two cases can be rewritten in the following unified form,

$$\texttt{term (i)}+\texttt{term (ii)} \leq \widetilde{O}\Big(d^{\frac{2}{3}}T^{\frac{2}{3}}\big(\max\{P_T, d^{-\frac{1}{2}}T^{\frac{1}{4}}\}\big)^{\frac{1}{3}}\Big).$$

Hence, we complete the proof of Theorem 4. □

C Technical Lemmas

In this section, we provide several technical lemmas that frequently used in the proofs.

Theorem 5 (Self-Normalized Bound for Vector-Valued Martingales [Abbasi-Yadkori et al., 2011, Theorem 1])**.** *Let $\{F_t\}_{t=0}^\infty$ be a filtration. Let $\{\eta_t\}_{t=0}^\infty$ be a real-valued stochastic process such that η_t is F_t-measurable and conditionally R-sub-Gaussian for some $R>0$, namely,*

$$\forall\lambda\in\mathbb{R},\quad \mathbb{E}[\exp(\lambda\eta_t)|F_{t-1}] \leq \exp\left(\frac{\lambda^2R^2}{2}\right). \qquad (29)$$

Let $\{X_t\}_{t=1}^\infty$ be an $\mathbb{R}^d$-valued stochastic process such that X_t is F_{t-1}-measurable. Assume that V is a $d\times d$ positive definite matrix. For any $t\geq 0$, define

$$\bar{V}_t = V + \sum_{\tau=1}^{t}X_\tau X_\tau^{\mathrm{T}},\quad S_t = \sum_{\tau=1}^{t}\eta_\tau X_\tau. \qquad (30)$$

Then, for any $\delta > 0$, with probability at least $1-\delta$, for all $t \geq 0$,

$$\|S_t\|^2_{\bar{V}_t^{-1}} \leq 2R^2 \log\left(\frac{\det(\bar{V}_t)^{1/2}\det(V)^{-1/2}}{\delta}\right). \quad (31)$$

Lemma 4 (Elliptical Potential Lemma). *Suppose $U_0 = \lambda I$, $U_t = U_{t-1} + X_t X_t^{\mathrm{T}}$, and $\|X_t\|_2 \leq L$, then*

$$\sum_{t=1}^{T} \|U_{t-1}^{-\frac{1}{2}} X_t\|_2 \leq \sqrt{2dT \log\left(1 + \frac{L^2 T}{\lambda d}\right)}. \quad (32)$$

Proof. First, we have the following decomposition,

$$U_t = U_{t-1} + X_t X_t^{\mathrm{T}} = U_{t-1}^{\frac{1}{2}}(I + U_{t-1}^{-\frac{1}{2}} X_t X_t^{\mathrm{T}} U_{t-1}^{-\frac{1}{2}}) U_{t-1}^{\frac{1}{2}}.$$

Taking the determinant on both sides, we get

$$\det(U_t) = \det(U_{t-1}) \det(I + U_{t-1}^{-\frac{1}{2}} X_t X_t^{\mathrm{T}} U_{t-1}^{-\frac{1}{2}}),$$

which in conjunction with Lemma 5 yields

$$\begin{aligned}\det(U_t) &= \det(U_{t-1})(1 + \|U_{t-1}^{-\frac{1}{2}} X_t\|_2^2) \\ &\geq \det(U_{t-1}) \exp(\|U_{t-1}^{-\frac{1}{2}} X_t\|_2^2 / 2).\end{aligned}$$

Note that in the first inequality, we utilize the fact that $1 + x \geq \exp(x/2)$ holds for any $x \in [0, 1]$. By taking advantage of the telescope structure, we have

$$\sum_{t=1}^{T} \|U_{t-1}^{-\frac{1}{2}} X_t\|_2^2 \leq 2 \log \frac{\det(U_T)}{\det(U_0)} \leq 2d \log\left(1 + \frac{L^2 T}{\lambda d}\right),$$

where the last inequality follows from the fact that $\mathrm{Tr}(U_T) \leq \mathrm{Tr}(U_0) + L^2 T = \lambda d + L^2 T$, and thus $\det(U_T) \leq (\lambda + L^2 T/d)^d$.

Therefore, Cauchy-Schwartz inequality gives,

$$\begin{aligned}\sum_{t=1}^{T} \|U_{t-1}^{-\frac{1}{2}} X_t\|_2 &\leq \sqrt{T \sum_{t=1}^{T} \|U_{t-1}^{-\frac{1}{2}} X_t\|_2^2} \\ &\leq \sqrt{2dT \log\left(1 + \frac{L^2 T}{\lambda d}\right)}.\end{aligned}$$

□

Lemma 5.

$$\det(I + \mathbf{v}\mathbf{v}^{\mathrm{T}}) = 1 + \|\mathbf{v}\|_2^2. \quad (33)$$

Proof. Notice that

(i) $(I + \mathbf{v}\mathbf{v}^{\mathrm{T}})\mathbf{v} = (1 + \|\mathbf{v}\|_2^2)\mathbf{v}$, therefore, $\mathbf{v}$ is its eigenvector with $(1 + \|\mathbf{v}\|_2^2)$ as the eigenvalue;

(ii) $(I + \mathbf{v}\mathbf{v}^{\mathrm{T}})\mathbf{v}^{\perp} = \mathbf{v}^{\perp}$, therefore, $\mathbf{v}^{\perp} \perp \mathbf{v}$ is its eigenvector with 1 as the eigenvalue.

Consequently, $\det(I + \mathbf{v}\mathbf{v}^{\mathrm{T}}) = 1 + \|\mathbf{v}\|_2^2$. □