A Technical Lemmas

When applying the following self-normalized bound in the analysis of federated bandit algorithm with event-trigger, a subtle difference from the analysis of standard bandit algorithm is that the sequence of data points used to update each client is controlled by the data-dependent event-trigger, e.g. Eq (4), which introduces dependencies on the future data, and thus breaks the standard argument. This problem also exists in prior works of distributed linear bandit, but was not addressed rigorously (see Lemma H.1. of [28]). Specifically, each client i observes the sequence of data points in a different order, i.e., it first observes each newly collected local data points from the environment, and then observes (in the form of their gradients) the batch of new data points that other clients have collected at the end of the epoch. Then, if we consider a data point that is contained in the batch of new data collected by other clients, the index of this data point (as observed by client i) has dependency all the way to the end of this batch, i.e., its index is only determined after some client triggers the global update.

Therefore, in order to avoid this dependency on future data points, when constructing the filtration, we should avoid including the σ-algebra that 'cuts a batch in half', but instead only include the σ-algebra generated by the sequence of data points up to the end of each batch, where we consider each locally observed data point as a batch as well. Denote the sequence of time indices corresponding to these data points as $\{t_p\}_{p\in[P]}$ for some $P > 0$. Then the constructed filtration $\{\mathcal{F}_{t_p}\}_{p\in[P]}$ is essentially a batched version of the standard $\{\mathcal{F}_t\}_{t=1}^{\infty}$. The self-normalized bound below still holds, i.e., by changing the stopping time construction from $\cup_{t\geq 1} B_t(\delta)$ to $\cup_{t\in\{t_p\}_{p\in[P]}} B_t(\delta)$ in the proof of Theorem 1 in [1], where $B_t(\delta)$ denotes the bad event that the bound does not hold. Therefore, instead of holding uniformly over all t, the self-normalized bound now only holds for all $t \in \{t_p\}_{p\in[P]}$, i.e., the sequence of time steps when client i gets updated, which is also what we need.

Lemma A.1 (Vector-valued self-normalized bound (Theorem 1 of [1]))**.** *Let $\{\mathcal{F}_t\}_{t=1}^{\infty}$ be a filtration. Let $\{\eta_t\}_{t=1}^{\infty}$ be a real-valued stochastic process such that η_t is $\mathcal{F}_{t+1}$-measurable, and η_t is conditionally zero mean R-sub-Gaussian for some $R \geq 0$. Let $\{X_t\}_{t=1}^{\infty}$ be a $\mathbb{R}^d$-valued stochastic process such that X_t is $\mathcal{F}_t$-measurable. Assume that V is a $d \times d$ positive definite matrix. For any $t > 0$, define*

$$V_t = V + \sum_{\tau=1}^{t} X_\tau X_\tau^\top \quad \mathcal{S}_t = \sum_{\tau=1}^{t} \eta_\tau X_\tau$$

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

$$||\mathcal{S}_t||_{V_t^{-1}} \leq R\sqrt{2\log\frac{\det(V_t)^{1/2}}{\det(V)^{1/2}\delta}}, \quad \forall t \geq 0$$

Lemma A.2 (Corollary 8 of [2])**.** *Under the same assumptions as Lemma A.1, consider a sequence of real-valued variables $\{Z_t\}_{t=1}^{\infty}$ such that Z_t is $\mathcal{F}_t$-measurable. Then for any $\delta > 0$, with probability at least $1-\delta$,*

$$|\sum_{\tau=1}^{t}\eta_\tau Z_\tau| \leq R\sqrt{2(V+\sum_{\tau=1}^{t} Z_\tau^2)\log\left(\frac{\sqrt{V+\sum_{\tau=1}^{t} Z_\tau^2}}{\delta\sqrt{V}}\right)}, \forall t \geq 0$$

Lemma A.3. *Under Assumption 1, $F_{i,t}(\theta)$ for $i = 1, 2, \dots, N$ is smooth with constant $k_\mu + \frac{\lambda}{Nt}$*

Proof. By Assumption 1, $\mu(\cdot)$ is Lipschitz continuous with constant k_μ, i.e., $|\mu(\mathbf{x}^\top\theta_1)-\mu(\mathbf{x}^\top\theta_2)| \leq k_\mu|\mathbf{x}^\top(\theta_1-\theta_2)|$. Then we can show that

$$
\begin{aligned}
&||\nabla F_{i,t}(\theta_1) - \nabla F_{i,t}(\theta_2)|| \\
&= ||\frac{1}{t}\sum_{s=1}^{t}\mathbf{x}_{s,i}[\mu(\mathbf{x}_{s,i}^\top\theta_1) - \mu(\mathbf{x}_{s,i}^\top\theta_2)] + \frac{\lambda}{Nt}(\theta_1-\theta_2)|| \\
&\leq \frac{1}{t}\sum_{s=1}^{t}||\mathbf{x}_{s,i}[\mu(\mathbf{x}_{s,i}^\top\theta_1) - \mu(\mathbf{x}_{s,i}^\top\theta_2)]|| + \frac{\lambda}{Nt}||\theta_1-\theta_2|| \\
&\leq \frac{1}{t}\sum_{s=1}^{t}|\mu(\mathbf{x}_{s,i}^\top\theta_1) - \mu(\mathbf{x}_{s,i}^\top\theta_2)| + \frac{\lambda}{Nt}||\theta_1-\theta_2|| \\
&\leq \frac{k_\mu}{t}\sum_{s=1}^{t}|\mathbf{x}_{s,i}^\top(\theta_1-\theta_2)| + \frac{\lambda}{Nt}||\theta_1-\theta_2|| \leq (k_\mu + \frac{\lambda}{Nt})||\theta_1-\theta_2||
\end{aligned}
$$

Therefore, $\nabla F_{i,t}(\theta)$ is Lipschitz continuous with constant $k_\mu + \frac{\lambda}{Nt}$, and $\nabla F_t(\theta) = \frac{1}{N}\sum_{i=1}^N \nabla F_{i,t}(\theta)$ is Lipschitz continuous with constant $k_\mu + \frac{\lambda}{Nt}$ as well. □

Lemma A.4 (Matrix Weighted Cauchy-Schwarz). *If $A \in \mathbb{R}^{d\times d}$ is a PSD matrix, then $x^T A y \leq \sqrt{x^T A x \cdot y^T A y}$ holds for any vectors $x, y \in \mathbb{R}^d$.*

Proof. Consider a quadratic function $(x+ty)^T A(x+ty) = x^T Ax + 2(x^T Ay)t + (y^T Ay)t^2$ for some variable $t \in \mathbb{R}$, where $x, y \in \mathbb{R}^d$ are arbitrary vectors. Since A is PSD, the value of this quadratic function $(x+ty)^T A(x+ty) = x^T Ax + 2(x^T Ay)t + (y^T Ay)t^2 \geq 0, \forall t$, which means there can be at most one root. This is equivalent to saying the discriminant of this quadratic function $4(x^T Ay)^2 - 4x^T Ax \cdot y^T Ay \leq 0$, which finishes the proof. □

Lemma A.5 (Confidence Ellipsoid Centered at Global Model). *Consider time step $t \in [T]$ when a global update happens, such that the distributed optimization over N clients is performed to get the globally updated model θ_t. Denote the sub-optimality of the final iteration as ϵ_t, such that $F_t(\theta_t) - F_t(\hat{\theta}_t^{\text{MLE}}) \leq \epsilon_t$; then with probability at least $1-\delta$, for all $t \in [T]$,*

$$||\theta_t - \theta_\star||_{A_t} \leq \alpha_t$$

where $\alpha_t = Nt\sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Ntc_\mu}}\sqrt{\epsilon_t} + \frac{R_{max}}{c_\mu}\sqrt{d\log(1+Ntc_\mu/(d\lambda)) + 2\log(1/\delta)} + \sqrt{\frac{\lambda}{c_\mu}}S$, and $A_t = \frac{\lambda}{c_\mu}I + \sum_{i=1}^N\sum_{s=1}^t \mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top$.

Proof. Recall that the unique minimizer of Eq.(2) is denoted as $\hat{\theta}_t^{\text{MLE}}$, so by taking gradient w.r.t. θ we have, $g_t(\hat{\theta}_t^{\text{MLE}}) = \sum_{i=1}^N\sum_{s=1}^t \mathbf{x}_{s,i}y_{s,i}$, where we define $g_t(\theta) = \lambda\theta + \sum_{i=1}^N\sum_{s=1}^t \mu(\mathbf{x}_{s,i}^\top\theta)\mathbf{x}_{s,i}$. First, we start with standard arguments [8, 20] to show that $\|\theta_t - \theta_\star\|_{A_t} \leq \frac{1}{c_\mu}\|g_t(\theta_t) - g_t(\theta_\star)\|_{A_t^{-1}}$. Specifically, by Assumption 1 and the Fundamental Theorem of Calculus, we have

$$g_t(\theta_t) - g_t(\theta_\star) = G_t(\theta_t - \theta_\star)$$

where $G_t = \int_0^1 \nabla g_t(a\theta_t + (1-a)\theta_\star)da$. Again by Assumption 1, $\nabla g_t(\theta) = \lambda I + \sum_{s=1}^t\sum_{i=1}^N \mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top\dot{\mu}(\mathbf{x}_{s,i}^\top\theta)$ is continuous, and $\nabla g_t(\theta) \succcurlyeq \lambda I + c_\mu\sum_{s=1}^t\sum_{i=1}^N \mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top \succ 0$ for $\theta \in \mathcal{B}_d(S)$, so $G_t \succ 0$, i.e., G_t is invertible. Therefore, we have

$$\theta_t - \theta_\star = G_t^{-1}[g_t(\theta_t) - g_t(\theta_\star)]$$

Note that $G_t \succcurlyeq \lambda I + c_\mu\sum_{s=1}^t\sum_{i=1}^N \mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top = c_\mu A_t$, so $G_t^{-1} \preccurlyeq \frac{1}{c_\mu}A_t^{-1}$. Hence,

$$
\begin{aligned}
\|\theta_t - \theta_\star\|_{A_t} &= \|G_t^{-1}[g_t(\theta_t) - g_t(\theta_\star)]\|_{A_t} \leq \|\frac{1}{c_\mu}A_t^{-1}[g_t(\theta_t) - g_t(\theta_\star)]\|_{A_t} = \frac{1}{c_\mu}\|g_t(\theta_t) - g_t(\theta_\star)\|_{A_t^{-1}} \\
&\leq \frac{1}{c_\mu}\|g_t(\theta_t) - g_t(\hat{\theta}_t^{\text{MLE}})\|_{A_t^{-1}} + \frac{1}{c_\mu}\|g_t(\hat{\theta}_t^{\text{MLE}}) - g_t(\theta_\star)\|_{A_t^{-1}} \qquad (8)
\end{aligned}
$$

where the first term depends on the sub-optimality of the offline regression estimator θ_t to the unique minimizer $\hat{\theta}_t^{(\text{MLE})}$, and the second term is standard for GLB [20].

Recall from Algorithm 3 that $\theta_t = \arg\min_{\theta \in \mathcal{B}_d(S)} \|g_t(\tilde{\theta}_t) - g_t(\theta)\|_{A_t^{-1}}$, where $\tilde{\theta}_t$ denotes the AGD estimator before projection. Therefore, for the first term, using triangle inequality and the definition of $g_t(\cdot)$, we have

$$
\begin{aligned}
&||g_t(\theta_t) - g_t(\hat{\theta}_t^{(\text{MLE})})||_{A_t^{-1}} \leq ||g_t(\theta_t) - g_t(\tilde{\theta}_t)||_{A_t^{-1}} + ||g_t(\tilde{\theta}_t) - g_t(\hat{\theta}_t^{(\text{MLE})})||_{A_t^{-1}} \\
&\leq 2||g_t(\tilde{\theta}_t) - g_t(\hat{\theta}_t^{(\text{MLE})})||_{A_t^{-1}} = 2||\lambda\theta_t + \sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}\mu(\mathbf{x}_{s,i}^\top\theta_t) - \sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}y_{s,i}||_{A_t^{-1}} \\
&= 2||\sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}[-y_{i,s} + \mu(\mathbf{x}_{s,i}^\top\theta_t)] + \lambda\theta_t||_{A_t^{-1}} = 2||Nt\nabla F_t(\theta_t)||_{A_t^{-1}}
\end{aligned}
$$

where the last equality is due to the definition of $F_t(\theta)$ in Eq.(2). We can further bound it using the property of Rayleigh quotient and the fact that $A_t \succcurlyeq \frac{\lambda}{c_\mu}I$, which gives us

$$
||g_t(\theta_t) - g_t(\hat{\theta}_t^{(\text{MLE})})||_{A_t^{-1}} \leq \frac{2Nt||\nabla F_t(\theta_t)||_2}{\sqrt{\lambda_{\min}(A_t)}} \leq \frac{2Nt||\nabla F_t(\theta_t)||_2}{\sqrt{\lambda/c_\mu}}
$$

Based on Lemma A.3, $F_t(\theta)$ is $(k_\mu + \frac{\lambda}{Nt})$-smooth, which means

$$
\frac{1}{2k_\mu + 2\lambda/(Nt)}\|\nabla F_t(\theta_t)\|^2 \leq F_t(\theta_t) - F_t(\hat{\theta}_t^{\text{MLE}}) \leq \epsilon_t
$$

where the second inequality is by definition of ϵ_t. Putting everything together, we have the following bound for the first term

$$
\frac{1}{c_\mu}\|g_t(\theta_t) - g_t(\hat{\theta}_t^{\text{MLE}})\|_{A_t^{-1}} \leq 2Nt\sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Ntc_\mu}}\sqrt{\epsilon_t}
$$

For the second term, similarly, based on the definition of $g_t(\cdot)$, we have

$$
\begin{aligned}
&\frac{1}{c_\mu}||g_t(\hat{\theta}_t^{\text{MLE}}) - g_t(\theta_\star)||_{A_t^{-1}} \\
&= \frac{1}{c_\mu}||\sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}y_{s,i} - \sum_{s=1}^{t}\sum_{i=1}^{N} \mu(\mathbf{x}_{s,i}^\top\theta_\star)\mathbf{x}_{s,i} - \lambda\theta_\star||_{A_t^{-1}} \\
&\leq \frac{1}{c_\mu}||\sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}\eta_{s,i}||_{A_t^{-1}} + \sqrt{\frac{\lambda}{c_\mu}}S
\end{aligned}
$$

Then based on the self-normalized bound in Lemma A.1 (Theorem 1 of [1]), we have $||\sum_{s=1}^{t}\sum_{i=1}^{N} \mathbf{x}_{s,i}\eta_{s,i}||_{A_t^{-1}} \leq R_{max}\sqrt{d\log(1 + Ntc_\mu/d\lambda) + 2\log(1/\delta)}, \forall t$, with probability at least $1 - \delta$.

Substituting the upper bounds for these two terms back into Eq.(8), we have, with probability at least $1 - \delta$,

$$
\begin{aligned}
\|\theta_t - \theta_\star\|_{A_t} &\leq \frac{1}{c_\mu}\|g_t(\theta_t) - g_t(\hat{\theta}_t^{(\text{MLE})})\|_{A_t^{-1}} + \frac{1}{c_\mu}\|g_t(\hat{\theta}_t^{(\text{MLE})}) - g_t(\theta_\star)\|_{A_t^{-1}} \\
&\leq 2Nt\sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Ntc_\mu}}\sqrt{\epsilon_t} + \frac{R_{max}}{c_\mu}\sqrt{d\log(1 + Ntc_\mu/(d\lambda)) + 2\log(1/\delta)} + \sqrt{\frac{\lambda}{c_\mu}}S
\end{aligned}
$$

which finishes the proof. □

B Proof of Lemma 4.1

Proof. Denote the two terms for loss difference as $A_1 = \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[l(\mathbf{x}_{s,i}^\top\theta_{t_{\text{last}}}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i})\big]$, and $A_2 = \sum_{s=t_{\text{last}}+1}^{t}\big[l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i})\big]$. We can upper bound the term A_1 by

$$
\begin{aligned}
A_1 &= \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[l(\mathbf{x}_{s,i}^\top\theta_{t_{\text{last}}}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i})\big] - \frac{\lambda}{2}\|\theta_\star\|_2^2 + \frac{\lambda}{2}\|\theta_\star\|_2^2 \\
&\leq \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\theta_{t_{\text{last}}}, y_{s,i}) - \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}, y_{s,i}) - \frac{\lambda}{2}\|\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}\|_2^2 + \frac{\lambda}{2}\|\theta_\star\|_2^2 \\
&\leq \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\theta_{t_{\text{last}}}, y_{s,i}) + \frac{\lambda}{2}\|\theta_{t_{\text{last}}}\|_2^2 - \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}, y_{s,i}) - \frac{\lambda}{2}\|\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}\|_2^2 + \frac{\lambda}{2}S^2 \\
&\leq N t_{\text{last}}\epsilon_{t_{\text{last}}} + \frac{\lambda}{2}S^2 := B_1
\end{aligned}
$$

where the first inequality is because $\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}$ minimizes Eq.(2), such that $\sum_{s=1}^{t_{\text{last}}} l(\mathbf{x}_{s,i}^\top\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}, y_{s,i}) + \frac{\lambda}{2}\|\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}\|_2^2 \leq \sum_{s=1}^{t_{\text{last}}} l(\mathbf{x}_{s,i}^\top\theta, y_{s,i}) + \frac{\lambda}{2}\|\theta\|_2^2$ for any $\theta \in \mathcal{B}_d(S)$, and the last inequality is because $F_{t_{\text{last}}}(\theta_{t_{\text{last}}}) - F_{t_{\text{last}}}(\hat{\theta}_{t_{\text{last}}}^{\text{MLE}}) \leq \epsilon_{t_{\text{last}}}$ by definition.

Now we start with standard arguments [12, 29] in order to bound the term A_2, which is essentially the online regret of ONS, except that its initial model is the globally updated model $\theta_{t_{\text{last}}}$. First, since $l(z, y)$ is c_μ-strongly-convex w.r.t. z, we have

$$
l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i}) \leq [\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}]\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star) - \frac{c_\mu}{2}||\theta_{s-1,i} - \theta_\star||^2_{\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top} \tag{9}
$$

To further bound the RHS of Eq.(9), recall from the ONS local update rule in Algorithm 2 that, for each client $i \in [N]$ at the end of each time step $s \in [t_{\text{last}} + 1, t]$,

$$
\begin{aligned}
\theta'_{s,i} &= \theta_{s-1,i} - \frac{1}{c_\mu}A_{s,i}^{-1}\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) \\
\theta_{s,i} &= \arg\min_{\theta \in \mathcal{B}_d(S)} ||\theta'_{s,i} - \theta||^2_{A_{s,i}}
\end{aligned}
$$

Then due to the property of generalized projection (Lemma 8 of [10]), we have

$$
\begin{aligned}
&\|\theta_{s,i} - \theta_\star\|^2_{A_{s,i}} \\
&\leq \|\theta_{s-1,i} - \theta_\star - \frac{1}{c_\mu}A_{s,i}^{-1}\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}} \\
&\leq \|\theta_{s-1,i} - \theta_\star\|^2_{A_{s,i}} - \frac{2}{c_\mu}\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})^\top(\theta_{s-1,i} - \theta_\star) + \frac{1}{c_\mu^2}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}}
\end{aligned}
$$

By rearranging terms, we have

$$
\begin{aligned}
&\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})^\top(\theta_{s-1,i} - \theta_\star) \\
&\leq \frac{1}{2c_\mu}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,j}^{-1}} + \frac{c_\mu}{2}\left(\|\theta_{s-1,i} - \theta_\star\|^2_{A_{s,i}} - \|\theta_{s,i} - \theta_\star\|^2_{A_{s,i}}\right) \\
&= \frac{1}{2c_\mu}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}} + \frac{c_\mu}{2}\|\theta_{s-1,i} - \theta_\star\|^2_{A_{s-1,i}} \\
&\quad + \frac{c_\mu}{2}\left(\|\theta_{s-1,i} - \theta_\star\|^2_{A_{s,i}} - \|\theta_{s-1,i} - \theta_\star\|^2_{A_{s-1,i}}\right) - \frac{c_\mu}{2}\|\theta_{s,i} - \theta_\star\|^2_{A_{s,i}} \\
&= \frac{1}{2c_\mu}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}} + \frac{c_\mu}{2}\|\theta_{s-1,i} - \theta_\star\|^2_{A_{s-1,i}} + \frac{c_\mu}{2}\|\theta_{s-1,i} - \theta_\star\|^2_{\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top} - \frac{c_\mu}{2}\|\theta_{s,i} - \theta_\star\|^2_{A_{s,i}}
\end{aligned}
$$

Note that $\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) = \mathbf{x}_{s,i}[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}]$, so with the inequality above, we can further bound the RHS of Eq.(9):

$$
\begin{aligned}
&l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i}) \leq [\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}]\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star) - \frac{c_\mu}{2}||\theta_{s-1,i} - \theta_\star||^2_{\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top} \\
&\leq \frac{1}{2c_\mu}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}} + \frac{c_\mu}{2}\|\theta_{s-1,i} - \theta_\star\|^2_{A_{s-1,i}} - \frac{c_\mu}{2}\|\theta_{s,i} - \theta_\star\|^2_{A_{s,i}}
\end{aligned}
$$

Then summing over $s \in [t_{\text{last}}+1, t]$, we have

$$A_2 \leq \frac{1}{2c_\mu} \sum_{s=t_{\text{last}}+1}^{t} \|\nabla l(\mathbf{x}_{s,i}^\top \theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}} + \frac{c_\mu}{2}\|\theta_{t_{\text{last}},i} - \theta_\star\|^2_{A_{t_{\text{last}},i}} - \frac{c_\mu}{2}\|\theta_{t,i} - \theta_\star\|^2_{A_{t,i}}$$

where $A_{t_{\text{last}},i} = A_{t_{\text{last}}}, \theta_{t_{\text{last}},i} = \theta_{t_{\text{last}}}, \forall i \in [N]$ due to the global update (line 15 in Algorithm 1).

We should note that the second term above itself essentially corresponds to a confidence ellipsoid centered at the globally updated model $\theta_{t_{\text{last}}}$, and its appearance in the upper bound for the loss difference (online regret) of local updates is because the local update is initialized by $\theta_{t_{\text{last}}}$. And based on Lemma A.5, with probability at least $1-\delta$,

$$\begin{aligned}\|\theta_{t_{\text{last}},i} - \theta_\star\|_{A_{t_{\text{last}},i}} \leq 2Nt_{\text{last}}\sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Nt_{\text{last}}c_\mu}}\sqrt{\epsilon_{t_{\text{last}}}} \\ + \frac{1}{c_\mu}R_{\max}\sqrt{d\log\left(1 + Nt_{\text{last}}c_\mu/d\lambda\right) + 2\log\left(1/\delta\right)} + \sqrt{\frac{\lambda}{c_\mu}}S\end{aligned}$$

Therefore, with probability at least $1-\delta$,

$$\begin{aligned}A_2 \leq \frac{1}{2c_\mu} \sum_{s=t_{\text{last}}+1}^{t} \|\nabla l(\mathbf{x}_{s,i}^\top \theta_{s-1,i}, y_{s,i})\|^2_{A_{s,i}^{-1}} + \frac{c_\mu}{2}\big[2Nt_{\text{last}}\sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Nt_{\text{last}}c_\mu}}\sqrt{\epsilon_{t_{\text{last}}}} \\ + \frac{1}{c_\mu}R_{\max}\sqrt{d\log\left(1 + Nt_{\text{last}}c_\mu/d\lambda\right) + 2\log\left(1/\delta\right)} + \sqrt{\frac{\lambda}{c_\mu}}S\big]^2 := B_2\end{aligned}$$

which finishes the proof for Lemma 4.1.

□

C Proof of Lemma 4.2 and Corollary 4.2.1

Proof of Lemma 4.2. Due to c_μ-strongly convexity of $l(z, y)$ w.r.t. z, we have $l(\mathbf{x}_{s,i}^\top\theta, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i}) \geq \big[\mu(\mathbf{x}_{s,i}^\top\theta_\star) - y_{s,i}\big]\mathbf{x}_{s,i}^\top(\theta - \theta_\star) + \frac{c_\mu}{2}\big[\mathbf{x}_{s,i}^\top(\theta - \theta_\star)\big]^2$. Substituting this to the LHS of Eq.(6) and Eq.(7), we have

$$\begin{aligned}B_1 &\geq \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[l(\mathbf{x}_{s,i}^\top\theta_{t_{\text{last}}}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i})\big] \\ &\geq \sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mu(\mathbf{x}_{s,i}^\top\theta_\star) - y_s\big]\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}} - \theta_\star) + \frac{c_\mu}{2}\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}} - \theta_\star)\big]^2 \\ B_2 &\geq \sum_{s=t_{\text{last}}+1}^{t}\big[l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) - l(\mathbf{x}_{s,i}^\top\theta_\star, y_{s,i})\big] \\ &\geq \sum_{s=t_{\text{last}}+1}^{t}\big[\mu(\mathbf{x}_{s,i}^\top\theta_\star) - y_s\big]\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star) + \frac{c_\mu}{2}\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star)\big]^2\end{aligned}$$

By rearranging the terms, we have

$$\begin{aligned}\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}} - \theta_\star)\big]^2 &\leq \frac{2}{c_\mu}B_1 + \frac{2}{c_\mu}\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}} - \theta_\star) \\ \sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star)\big]^2 &\leq \frac{2}{c_\mu}B_2 + \frac{2}{c_\mu}\sum_{s=t_{\text{last}}+1}^{t}\eta_{s,i}\mathbf{x}_{s,i}^\top(\theta_{s-1,i} - \theta_\star)\end{aligned}$$

where the LHS is quadratic in $\theta_\star$. For the RHS, we will further upper bound the second term as shown below.

• **Upper Bound for $\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2$** Note that $\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)$ is $\mathcal{F}_{s,i}$-measurable, and $\eta_{s,i}$ is $\mathcal{F}_{s+1,i}$-measurable and conditionally R_{max}-sub-Gaussian. By applying Lemma A.2 (Corollary 8 of [2]) w.r.t. client i's filtration $\{\mathcal{F}_{s,i}\}_{s=t_{\text{last}}+1}^{\infty}$, where $\mathcal{F}_{s,i}=\sigma\big([\mathbf{x}_{k,j},\eta_{k,j}]_{k,j:k\leq t_{\text{last}}\cap j\leq N},[\mathbf{x}_{k,j},\eta_{k,j}]_{k,j:t_{\text{last}}+1\leq k\leq s-1\cap j=i},\mathbf{x}_{s,i}\big)$, and taking union bound over all $i\in[N]$, with probability at least $1-\delta$, for all $t\in[T], i\in[N]$,

$$\sum_{s=t_{\text{last}}+1}^{t}\eta_{s,i}\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\leq$$

$$R_{\max}\sqrt{2\big(1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2\big)\cdot\log\big(\frac{N}{\delta}\sqrt{1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2}\big)}$$

Therefore,

$$\begin{aligned}1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2&\leq 1+\frac{2}{c_\mu}B_2\\&+\frac{2R_{\max}}{c_\mu}\sqrt{2\big(1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2\big)\cdot\log\big(\frac{N}{\delta}\sqrt{1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2}\big)}\end{aligned}\tag{10}$$

Then by applying Lemma 2 of [12], i.e., if $q^2\leq a+fq\sqrt{\log(\frac{q}{\delta/N})}$ then $q^2\leq 2a+f^2\log(\frac{\sqrt{4a+f^4/(4\delta^2)}}{\delta/N})$ (for $a,f\geq 0, q\geq 1$). And by setting $q=\sqrt{1+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2}$, $a=1+\frac{2}{c_\mu}B_2$, $f=\frac{2\sqrt{2}R_{\max}}{c_\mu}$, we have

$$\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2\leq 1+\frac{4B_2}{c_\mu}+\frac{8R_{\max}^2}{c_\mu^2}\log\left(\frac{N}{\delta}\sqrt{4+\frac{8}{c_\mu}B_2+\frac{64R_{\max}^4}{c_\mu^4\cdot 4\delta^2}}\right),\forall t,i\tag{11}$$

with probability at least $1-\delta$.

• **Upper Bound for $\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2$** Note that $\theta_{t_{\text{last}}}$ depends on all data samples in $\{(\mathbf{x}_{s,i},y_{s,i})\}_{s\in[t_{\text{last}}]}$ as a result of the offline regression method, and therefore $\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)$ is no longer $\mathcal{F}_{s,i}$-measurable for $s\in[1,t_{\text{last}})$. Hence, we cannot use Lemma A.2 as before. Instead, we have

$$\begin{aligned}&\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)=\big(\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}\big)^\top(\theta_{t_{\text{last}}}-\theta_\star)\\&=\big(\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}\big)^\top\big(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top\big)^{-1}\big(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top\big)(\theta_{t_{\text{last}}}-\theta_\star)\\&\leq\sqrt{\big(\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}\big)^\top\big(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top\big)^{-1}\big(\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}\big)\cdot(\theta_{t_{\text{last}}}-\theta_\star)^\top\big(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top\big)(\theta_{t_{\text{last}}}-\theta_\star)}\\&=\sqrt{\|\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\eta_{s,i}\mathbf{x}_{s,i}\|^2_{(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)^{-1}}\cdot\|\theta_{t_{\text{last}}}-\theta_\star\|^2_{(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}}\\&\leq R_{\max}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)\cdot\|\theta_{t_{\text{last}}}-\theta_\star\|^2_{(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}},\end{aligned}$$

with probability at least $1-\delta$, where the first inequality is due to the matrix-weighted Cauchy-Schwarz inequality in Lemma A.4, such that $x^\top A^{-1}Ay\leq\sqrt{x^\top A^{-1}x\cdot y^\top A^\top A^{-1}Ay}=\sqrt{x^\top A^{-1}x\cdot y^\top Ay}$ for symmetric PD matrix A, and the second inequality is obtained by applying the self-normalized bound in Lemma A.1 w.r.t. the filtration $\{\mathcal{F}_s\}_{s\in\{t_p\}_{p=1}^{B}}$, where $\mathcal{F}_s=\sigma\big([\mathbf{x}_{k,j},\eta_{k,j}]_{k,j:k\leq s-1\cap j\leq N},[\mathbf{x}_{k,j},\eta_{k,j}]_{k,j:k=s\cap j\leq N-1},\mathbf{x}_{s,N}\big)$ and $\{t_p\}_{p=1}^{B}$ denotes the sequence of time steps when global update happens, and B denotes the total number of global updates.

By substituting it back, we have

$$
\begin{aligned}
&\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2 \\
&\le \frac{2}{c_\mu}B_1 + \frac{2R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)\cdot\|\theta_{t_{\text{last}}}-\theta_\star\|^2_{I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top}} \\
&\le \frac{2}{c_\mu}B_1 + \frac{2R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)\cdot\big(\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2+\|\theta_{t_{\text{last}}}-\theta_\star\|_2^2\big)}
\end{aligned}
\tag{12}
$$

Then by applying the Proposition 9 of [2], i.e. if $z^2 \le a + bz$ then $z \le b + \sqrt{a}$ (for $a, b \ge 0$), and setting $z = \sqrt{\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2+\|\theta_{t_{\text{last}}}-\theta_\star\|_2^2}$, $a = \|\theta_{t_{\text{last}}}-\theta_\star\|_2^2 + \frac{2}{c_\mu}B_1$, $b = \frac{2R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)}$,we have

$$
\begin{aligned}
&\sqrt{\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2+\|\theta_{t_{\text{last}}}-\theta_\star\|_2^2} \\
\le&\frac{2R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)} + \sqrt{\|\theta_{t_{\text{last}}}-\theta_\star\|_2^2+B_1}
\end{aligned}
\tag{13}
$$

Taking square on both sides, and rearranging terms, we have

$$
\begin{aligned}
&\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2 \\
\le&\frac{8R_{\max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)+B_1 \\
&+\frac{4R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)}\sqrt{\|\theta_{t_{\text{last}}}-\theta_\star\|_2^2+B_1} \\
\le&\frac{8R_{\max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)+B_1 \\
&+\frac{4R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)}(\|\theta_{t_{\text{last}}}-\theta_\star\|_2+\sqrt{B_1}) \\
\le&\frac{8R_{\max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)+B_1 \\
&+\frac{4R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)}(\|\theta_{t_{\text{last}}}\|_2+\|\theta_\star\|_2+\sqrt{B_1})
\end{aligned}
\tag{14}
$$

Now putting everything together, we have the following confidence region for $\theta_\star$,

$$
P\big(\forall t,i,\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\big[\mathbf{x}_{s,i}^\top(\theta_{t_{\text{last}}}-\theta_\star)\big]^2+\sum_{s=t_{\text{last}}+1}^{t}\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2\le\beta_{t,i}\big)\ge 1-2\delta \tag{15}
$$

where $\beta_{t,i} = \frac{8R_{\max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big) + B_1 + \frac{4R_{\max}}{c_\mu}\sqrt{2\log\big(\frac{1}{\delta}\sqrt{\det(I+\sum_{s=1}^{t_{\text{last}}}\sum_{i=1}^{N}\mathbf{x}_{s,i}\mathbf{x}_{s,i}^\top)}\big)}(\|\theta_{t_{\text{last}}}\|_2 + \|\theta_\star\|_2 + \sqrt{B_1}) + 1 + \frac{4B_2}{c_\mu} + \frac{8R_{\max}^2}{c_\mu^2}\log\big(\frac{N}{\delta}\sqrt{4+\frac{8}{c_\mu}B_2+\frac{64R_{\max}^4}{c_\mu^4\cdot 4\delta^2}}\big)$.

Denote $\mathbf{X}_{t,i} = \begin{bmatrix}\mathbf{x}_{1,1}^\top\\ \cdots\\ \mathbf{x}_{t_{\text{last}},N}^\top\\ \mathbf{x}_{i,t_{\text{last}}+1}^\top\\ \cdots\\ \mathbf{x}_{i,t}^\top\end{bmatrix} \in \mathbb{R}^{(Nt_{\text{last}}+t-t_{\text{last}})\times d}$, and $\mathbf{z}_{t,i} = \begin{bmatrix}\mathbf{x}_{1,1}^\top\theta_{t_{\text{last}}}\\ \cdots\\ \mathbf{x}_{t_{\text{last}},N}^\top\theta_{t_{\text{last}}}\\ \mathbf{x}_{i,t_{\text{last}}+1}^\top\theta_{t_{\text{last}},i}\\ \cdots\\ \mathbf{x}_{i,t}^\top\theta_{t-1,i}\end{bmatrix} \in \mathbb{R}^{Nt_{\text{last}}+t-t_{\text{last}}}$. We can rewrite the inequality above as

$$
\begin{aligned}
&\|\mathbf{z}_{t,i}-\mathbf{X}_{t,i}\theta_\star\|_2^2 + \frac{\lambda}{c_\mu}\|\theta_\star\|_2^2 \le \beta_{t,i} + \frac{\lambda}{c_\mu}\|\theta_\star\|_2^2 \le \beta_{t,i} + \frac{\lambda}{c_\mu}S^2\\
\Leftrightarrow &\|\mathbf{z}_{t,i}-\mathbf{X}_{t,i}\theta_\star\|_2^2 + \frac{\lambda}{c_\mu}\|\theta_\star\|_2^2 - \|\mathbf{z}_{t,i}-\mathbf{X}_{t,i}\hat\theta_{t,i}\|_2^2 - \frac{\lambda}{c_\mu}\|\hat\theta_{t,i}\|_2^2 + \|\mathbf{z}_{t,i}-\mathbf{X}_{t,i}\hat\theta_{t,i}\|_2^2 + \frac{\lambda}{c_\mu}\|\hat\theta_{t,i}\|_2^2\\
&\le \beta_{t,i} + \frac{\lambda}{c_\mu}S^2
\end{aligned}
$$

where $\hat\theta_{t,i} = A_{t,i}^{-1}\mathbf{X}_{t,i}^\top\mathbf{z}_{t,i}$ denotes the Ridge regression estimator based on the predicted rewards given by the past sequence of model updates, and the regularization parameter is $\frac{\lambda}{c_\mu}$. Note that by expanding $\hat\theta_{t,i}$, we can show $\hat\theta_{t,i}^\top A_{i,t}\hat\theta_{t,i} = \mathbf{z}_{i,t}^\top\mathbf{X}_{i,t}\hat\theta_{t,i}$, and $\hat\theta_{t,i}^\top A_{i,t}\theta_\star = \mathbf{z}_{i,t}^\top\mathbf{X}_{i,t}\theta_\star$. Therefore, we have

$$
\|\hat\theta_{t,i}-\theta_\star\|_{A_{t,i}}^2 \le \beta_{t,i} + \frac{\lambda}{c_\mu}S^2 - (\|\mathbf{z}_{t,i}\|_2^2 - \hat\theta_{t,i}^\top\mathbf{X}_{t,i}^\top\mathbf{z}_{t,i})
$$

which finishes the proof of Lemma 4.2. □

Proof of Corollary 4.2.1. Under the condition that $\epsilon_{t_{\text{last}}} \le \frac{1}{N^2 t_{\text{last}}^2}$,

$$
\begin{aligned}
B_1 &\le \frac{1}{Nt_{\text{last}}} + \frac{\lambda}{2}S^2 = O(1)\\
B_2 &\le \frac{1}{2c_\mu}\sum_{s=t_{\text{last}}+1}^{t}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i})\|_{A_{s,i}^{-1}}^2\\
&\quad + \frac{c_\mu}{2}\big[2\sqrt{\frac{2k_\mu}{\lambda c_\mu}+\frac{2}{Nt_{\text{last}}c_\mu}} + \frac{1}{c_\mu}R_{\max}\sqrt{d\log(1+Nt_{\text{last}}c_\mu/d\lambda)+2\log(1/\delta)} + \sqrt{\frac{\lambda}{c_\mu}}S\big]^2
\end{aligned}
$$

Note that $\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i}, y_{s,i}) = \mathbf{x}_{s,i}[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}]$. We can upper bound the squared prediction error by

$$
\begin{aligned}
&\big[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}\big]^2\\
&= \big[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - \mu(\mathbf{x}_{s,i}^\top\theta_\star) - \eta_{s,i}\big]^2\\
&\le 2\big[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - \mu(\mathbf{x}_{s,i}^\top\theta_\star)\big]^2 + 2\eta_{s,i}^2\\
&\le 2k_\mu^2\big[\mathbf{x}_{s,i}^\top(\theta_{s-1,i}-\theta_\star)\big]^2 + 2\eta_{s,i}^2\\
&\le 8k_\mu^2 S^2 + 2\eta_{s,i}^2
\end{aligned}
$$

where the first inequality is due to AM-QM inequality, and the second inequality is due to the k_μ-Lipschitz continuity of $\mu(\cdot)$ according to Assumption 1. Since $|\eta_{s,i}| \le R_{\max}$, $\big[\mu(\mathbf{x}_{s,i}^\top\theta_{s-1,i}) - y_{s,i}\big]^2 \le k_\mu^2 S^2 + R_{\max}^2$. In addition, due to Lemma 11 of [1], i.e., $\sum_{s=t_{\text{last}}+1}^{t}\|\mathbf{x}_{s,i}\|_{A_{s,i}^{-1}}^2 \le$

$2\log(\frac{\det(A_{t,i})}{\det(\lambda I)})$ Therefore,

$$\frac{1}{2c_\mu}\sum_{s=t_{\text{last}}+1}^{t}\|\nabla l(\mathbf{x}_{s,i}^\top\theta_{s-1,i},y_{s,i})\|^2_{A_{s,i}^{-1}}=O\big(\frac{d\log NT}{c_\mu}[k_\mu^2 S^2+R_{\max}^2]\big)$$

so $B_2=O\big(\frac{d\log NT}{c_\mu}[k_\mu^2S^2+R_{\max}^2]\big)$. Hence,

$$\beta_{t,i}=O(d\frac{R_{\max}^2}{c_\mu^2}\log NT+d\frac{k_\mu^2}{c_\mu^2}\log NT+d\frac{R_{\max}^2}{c_\mu^2}\log NT)=O(\frac{d\log NT}{c_\mu^2}[k_\mu^2+R_{\max}^2])$$

which finishes the proof. □

D Proof of Theorem 4.3

Proof. Since $\mu(\cdot)$ is k_μ-Lipschitz continuous, we have $\mu(\mathbf{x}_{t,\star}^\top\theta_\star)-\mu(\mathbf{x}_{t,i}^\top\theta_\star)\le k_\mu(\mathbf{x}_{t,\star}^\top\theta_\star-\mathbf{x}_{t,i}^\top\theta_\star)$. Then we have the following upper bound on the instantaneous regret,

$$\begin{aligned}\frac{r_{t,i}}{k_\mu}&\le\mathbf{x}_{t,\star}^\top\theta_\star-\mathbf{x}_{t,i}^\top\theta_\star\le\mathbf{x}_{t,i}^\top\tilde\theta_{t-1,i}-\mathbf{x}_{t,i}^\top\theta_\star\\&=\mathbf{x}_{t,i}^\top(\tilde\theta_{t-1,i}-\hat\theta_{t-1,i})+\mathbf{x}_{t,i}^\top(\hat\theta_{t-1,i}-\theta_\star)\\&\le\|\mathbf{x}_{t,i}\|_{A_{t-1,i}^{-1}}\|\tilde\theta_{t-1,i}-\hat\theta_{t-1,i}\|_{A_{t-1,i}}+\|\mathbf{x}_{t,i}\|_{A_{t-1,i}^{-1}}\|\hat\theta_{t-1,i}-\theta_\star\|_{A_{t-1,i}}\\&\le 2\alpha_{t-1,i}\cdot\|\mathbf{x}_{t,i}\|_{A_{t-1,i}^{-1}}\end{aligned}$$

which holds for all $i\in[N],t\in[T]$, with probability at least $1-2\delta$. And $\tilde\theta_{t-1.i}$ denotes the optimistic estimate in the confidence ellipsoid that maximizes the UCB score when client i selects arm at time step t.

Now consider an imaginary centralized agent that has direct access to all clients' data, and we denote its covariance matrix as $\tilde A_{t,i}=\frac{\lambda}{c_\mu}I+\sum_{s=1}^{t-1}\sum_{j=1}^{N}\mathbf{x}_{s,j}\mathbf{x}_{s,j}+\sum_{j=1}^{i}\mathbf{x}_{t,j}\mathbf{x}_{t,j}^\top$, i.e., $\tilde A_{t,i}$ is immediately updated after any client obtains a new data sample from the environment. Then we can obtain the following upper bound for $r_{t,i}$, which is dependent on the determinant ratio between the covariance matrix of the imaginary centralized agent and that of client i, i.e., $\det(\tilde A_{t-1,i})/\det(A_{t-1,i})$.

$$r_{t,i}\le 2k_\mu\alpha_{t-1,i}\sqrt{\mathbf{x}_{t,i}^\top A_{t-1,i}^{-1}\mathbf{x}_{t,i}}\le 2k_\mu\alpha_{t-1,i}\sqrt{\mathbf{x}_{t,i}^\top\tilde A_{t-1,i}^{-1}\mathbf{x}_{t,i}\cdot\frac{\det(\tilde A_{t-1,i})}{\det(A_{t-1,i})}}$$

We refer to the time period in-between two consecutive global updates as an epoch, and denote the total number of epochs as $B\in\mathbb{R}$, i.e., the p-th epoch refers to the period from $t_{p-1}+1$ to t_p, for $p\in[B]$, where t_p denotes the time step when the p-th global update happens. Then the p-th epoch is called a 'good' epoch if the determinant ratio $\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})}\le 2$, where A_{t_p} is the aggregated sufficient statistics computed at the p-th global update. Otherwise, it is called a 'bad' epoch. In the following, we bound the cumulative regret in 'good' and 'bad' epochs separately.

Suppose the p-th epoch is a good epoch, then for any client $i\in[N]$, and time step $t\in[t_{p-1}+1,t_p]$, we have $\frac{\det(\tilde A_{t-1,i})}{\det(A_{t-1,i})}\le\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})}\le 2$, because $A_{t-1,i}\succcurlyeq A_{t_{p-1}}$ and $\tilde A_{t-1,i}\preccurlyeq A_{t_p}$. Therefore, the instantaneous regret incurred by any client i at any time step t of a good epoch can be bounded by

$$r_{t,i}\le 2\sqrt{2}k_\mu\alpha_{t-1,i}\sqrt{\mathbf{x}_{t,i}^\top\tilde A_{t-1,i}^{-1}\mathbf{x}_{t,i}}$$

with probability at least $1-2\delta$. Therefore, using standard arguments for UCB-type algorithms, e.g., Theorem 2 in [20], the cumulative regret for all the 'good epochs' is

$$\begin{aligned}REG_{good}&\le 2\sqrt{2}k_\mu\alpha_{t-1,i}\sum_{t=1}^{T}\sum_{i=1}^{N}\|\mathbf{x}_{t,i}\|_{\tilde A_{t-1,i}^{-1}}\\&=O\left(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}d\sqrt{NT}\log NT\right)\end{aligned}$$

which matches the regret upper bound of GLOC [12].

Now suppose the p-th epoch is bad. Then the cumulative regret incurred by all N clients during this 'bad epoch' can be upper bounded by:

$$\begin{aligned}
&\sum_{t=t_{p-1}+1}^{t_p}\sum_{i=1}^{N} r_{t,i} \\
&\leq O(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}\sqrt{d\log(NT)})\sum_{t=t_{p-1}+1}^{t_p}\sum_{i=1}^{N}\min(1,||\mathbf{x}_{t,i}||_{A_{t-1,i}^{-1}}) \\
&\leq O(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}\sqrt{d\log(NT)})\sum_{i=1}^{N}\sqrt{(t_p-t_{p-1})\log\frac{\det(A_{t_p-1,i})}{\det(A_{t_p-1,i}-\Delta A_{t_p-1,i})}} \\
&\leq O(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}N\sqrt{d\log(NT)D})
\end{aligned}$$

where the last inequality is due to the event-trigger design in Algorithm 1. Following the same argument as [28], there can be at most $R=\lceil d\log(1+\frac{NTc_\mu}{\lambda d})\rceil=O\big(d\log(NT)\big)$ 'bad epochs', because $\det(A_{t_B})\leq\det(\tilde{A}_{T,N})\leq(\frac{\lambda}{c_\mu}+\frac{NT}{d})^d$. Therefore, the cumulative regret for all the 'bad epochs' is

$$REG_{bad}=O\left(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}d^{1.5}\log^{1.5}(NT)ND^{0.5}\right)$$

Combining the regret upper bound for 'good' and 'bad' epochs, the cumulative regret

$$R_T=O\left(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}(d\sqrt{NT}\log(NT)+d^{1.5}\log^{1.5}(NT)ND^{0.5})\right).$$

To obtain upper bound for the communication cost C_T, we first upper bound the total number of epochs B. Denote the length of an epoch, i.e., the number of time steps between two consecutive global updates, as $\alpha>0$, so that there can be at most $\lceil\frac{T}{\alpha}\rceil$ epochs with length longer than α. For a particular epoch p with less than α time steps, we have $t_p-t_{p-1}<\alpha$. Moreover, due to the event-trigger design in Algorithm 1, we have $(t_p-t_{p-1})\log\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})}>D$, which means $\log\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})}>\frac{D}{\alpha}$. Since $\sum_{p=1}^{B}\log\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})}\leq R$, the number of epochs with less than α time steps is at most $\lceil\frac{R\alpha}{D}\rceil$. Therefore, the total number of epochs.

$$B\leq\lceil\frac{T}{\alpha}\rceil+\lceil\frac{R\alpha}{D}\rceil$$

which is minimized it by choosing $\alpha=\sqrt{\frac{DT}{R}}$, so $B\leq\sqrt{\frac{TR}{D}}=O(d^{0.5}\log^{0.5}(NT)T^{0.5}D^{-0.5})$.

At the end of each epoch, FedGLB-UCB has a global update step that executes AGD among all N clients. As mentioned in Section 4.1, the number of iterations required by AGD has upper bound

$$J_t\leq 1+\sqrt{\frac{k_\mu}{\lambda}Nt+1}\log\frac{(k_\mu+\frac{2\lambda}{Nt})\|\theta_t^{(1)}-\hat{\theta}_t^{\text{MLE}}\|_2^2}{2\epsilon_t},$$

and under the condition that $\epsilon_t=\frac{1}{N^2t^2},\forall t\in[T]$, we have $J_t=O\big(\sqrt{NT}\log(NT)\big),\forall t\in[T]$. Moreover, each iteration of AGD involves communication with N clients, so the communication cost

$$C_T=O(d^{0.5}\log^{1.5}(NT)TN^{1.5}D^{-0.5})$$

In order to match the regret under centralized setting, we set the threshold $D=\frac{T}{Nd\log(NT)}$, which gives us $R_T=O(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}d\sqrt{NT}\log(NT))$, and $C_T=O(dN^2\sqrt{T}\log^2(NT))$. □

E Theoretical Analysis for Variants of FedGLB-UCB

In this section, we describe and analyze the variants of FedGLB-UCB listed in Table 1. The first variant, FedGLB-UCB$_1$, completely disables local update, and we can see that it requires a linear communication cost in T to attain the $O(d\sqrt{NT}\log(NT))$ regret. As we mentioned in Section 4.1, this is because in the absence of local update, FedGLB-UCB$_1$ requires more frequent global updates, i.e., $\sqrt{NT}$ in total, to control the sub-optimality of the employed bandit model w.r.t the growing training set. The second variant, denoted as FedGLB-UCB$_2$, is exactly the same as FedGLB-UCB, except for its fixed communication schedule. This leads to additional $d\sqrt{N}$ global updates, as fixed update schedule cannot adapt to the actual quality of collected data. The third variant, denoted as FedGLB-UCB$_3$, uses ONS for both local and global update, such that only one round of gradient aggregation among N clients is performed for each global update, i.e., lazy ONS update over batched data. It incurs the least communication cost among all variants, but its regret grows at a rate of $(NT)^{3/4}$ due to the inferior quality of its lazy ONS update.

E.1 FedGLB-UCB$_1$: scheduled communication + no local update

Though many real-world applications are online problems in nature, i.e., the clients continuously collect new data samples from the users, standard federated/distributed learning methods do not provide a principled solution to adapt to the growing datasets. A common practice is to manually set a fixed global update schedule in advance, i.e., periodically update and deploy the model.

To demonstrate the advantage of FedGLB-UCB over this straightforward solution, we present and analyze the first variant FedGLB-UCB$_1$, which completely disables local update, and performs global update according to a fixed schedule $\mathcal{S} = \{t_1 := \lfloor \frac{T}{B} \rfloor, t_2 := 2\lfloor \frac{T}{B} \rfloor, \dots, t_B := B\lfloor \frac{T}{B} \rfloor\}$, where B is the total number of global updates up to time step T. The description of FedGLB-UCB$_1$ is presented in Algorithm 4.

Algorithm 4 FedGLB-UCB$_1$

Input: communication schedule $\mathcal{S}$, regularization parameter $\lambda > 0$, $\delta \in (0,1)$ and c_μ.
Initialize $\forall i \in [N]$: $\theta_{0,i} = \mathbf{0} \in \mathbb{R}^d, A_{0,i} = \frac{\lambda}{c_\mu}\mathbf{I} \in \mathbb{R}^{d\times d}, \mathbf{X}_{0,i} = \mathbf{0} \in \mathbb{R}^{0\times d}, \mathbf{y}_{0,i} = \mathbf{0} \in \mathbb{R}^0$, $t_{\text{last}} = 0$
for $t = 1, 2, ..., T$ **do**
  for client $i = 1, 2, ..., N$ **do**
    Observe arm set $\mathcal{A}_{t,i}$ for client i
    Select arm $x_{t,i} \in \mathcal{A}_{t,i}$ according to Eq. (16) and observe reward $y_{t,i}$
    Update client i: $\mathbf{X}_{t,i} = \begin{bmatrix} \mathbf{X}_{t-1,i} \\ \mathbf{x}_{t,i}^\top \end{bmatrix}, \mathbf{y}_{t,i} = \begin{bmatrix} \mathbf{y}_{t-1,i} \\ y_{t,i} \end{bmatrix}$
  if $t \notin \mathcal{S}$ **then**
    Clients: set $\theta_{t,i} = \theta_{t-1,i}, A_{t,i} = A_{t-1,i}, \forall i \in [N]$
  else
    Clients: send $\{\mathbf{X}_{t,i}^\top \mathbf{X}_{t,i}\}_{i\in[N]}$ to server
    Server compute $A_t = \frac{\lambda}{c_\mu}\mathbf{I} + \sum_{i=1}^N \mathbf{X}_{t,i}^\top \mathbf{X}_{t,i}$ and send A_t to all clients.
    Clients: set $A_{t,i} = A_t$, for $i \in [N]$
    Server update global model $\theta_t = \text{AGD-Update}(\theta_{t_{\text{last}}}, J_t)$, and set $t_{\text{last}} = t$
    Clients set local models $\theta_{t,i} = \theta_t, \forall i \in [N]$

In FedGLB-UCB$_1$, each client stores a local model $\theta_{t-1,i}$, and the corresponding covariance matrix $A_{t-1,i}$. Note that $\{\theta_{t-1,i}, A_{t-1,i}\}_{i\in[N]}$ are only updated at time steps $t \in \mathcal{S}$, and remain unchanged for $t \notin \mathcal{S}$. At time t, client i selects the arm that maximizes the following UCB score:

$$\mathbf{x}_{t,i} = \arg\max_{\mathbf{x}\in\mathcal{A}_{t,i}} \mathbf{x}^\top \theta_{t-1,i} + \alpha_{t-1,i}||\mathbf{x}||_{A_{t-1,i}^{-1}} \tag{16}$$

where $\alpha_{t-1,i}$ is given in Lemma A.5. The regret and communication cost of FedGLB-UCB$_1$ is given in the following theorem.

Theorem E.1 (Regret and Communication Cost Upper Bound of FedGLB-UCB$_1$). *Under the condition that $\epsilon_t = \frac{1}{N^2t^2}$, and the total number of global synchronizations $B = \sqrt{NT}$, the cumulative* *regret R_T has upper bound*

$$R_T = O\left(\frac{k_\mu R_{\max} d}{c_\mu}\sqrt{NT}\log(NT/\delta)\right)$$

with probability at least $1-\delta$. The cumulative communication cost has upper bound

$$C_T = O(N^2 T \log(NT))$$

Proof. First, based on Lemma A.5 and under the condition that $\epsilon_t = \frac{1}{N^2t^2}$, we have

$$\|\theta_t - \theta_\star\|_{A_t} \leq \alpha_t$$

holds $\forall t$, where $\alpha_t = \sqrt{\frac{2k_\mu}{\lambda c_\mu} + \frac{2}{Ntc_\mu}} + \frac{R_{max}}{c_\mu}\sqrt{d\log\left(1 + Ntc_\mu/(d\lambda)\right) + 2\log\left(1/\delta\right)} + \sqrt{\frac{\lambda}{c_\mu}}S = O(\frac{R_{max}}{c_\mu}\sqrt{d\log(Nt)})$, which matches the order in [20].

Similar to the proof of Theorem 4.3, we decompose all B epochs into 'good' and 'bad' epochs according to the log-determinant ratio: the p-th epoch, for $p \in [B]$, is a 'good' epoch if the determinant ratio $\frac{\det(A_{t_p})}{\det(A_{t_{p-1}})} \leq 2$. Otherwise, it is a 'bad' epoch. In the following, we bound the cumulative regret in 'good' and 'bad' epochs separately.

Suppose epoch p is a good epoch, then for any client $i \in [N]$, and time step $t \in [t_{p-1}+1, t_p]$, we have $\frac{\det(\tilde{A}_{t-1,i})}{\det(A_{t-1,i})} \leq \frac{\det(A_{t_p})}{\det(A_{t_{p-1}})} \leq 2$, because $A_{t-1,i} = A_{t_{p-1}}$ and $\tilde{A}_{t-1,i} \preccurlyeq A_{t_p}$. Therefore, the instantaneous regret incurred by any client i at any time step t of a good epoch p can be bounded by

$$\begin{aligned} r_{t,i} &\leq 2k_\mu \alpha_{t_{p-1}} \sqrt{\mathbf{x}_{t,i}^\top A_{t-1,i} \mathbf{x}_{t,i}} \leq 2k_\mu \alpha_{t_{p-1}} \sqrt{\mathbf{x}_{t,i}^\top A_{t-1}^{-1} \mathbf{x}_{t,i} \cdot \frac{\det(\tilde{A}_{t-1,i})}{\det(A_{t-1,i})}} \\ &\leq 2\sqrt{2} k_\mu \alpha_T \sqrt{\mathbf{x}_{t,i}^\top A_{t-1}^{-1} \mathbf{x}_{t,i}} \end{aligned}$$

By standard arguments [1, 20], the cumulative regret incurred in all good epochs can be upper bounded by $O(\frac{k_\mu R_{\max}}{c_\mu} d\sqrt{NT}\log(NT/\delta))$ with probability at least $1-\delta$.

By Assumption 1, $\mu(\cdot)$ is Lipschitz continuous with constant k_μ, i.e., $|\mu(\mathbf{x}^\top\theta_1) - \mu(\mathbf{x}^\top\theta_2)| \leq k_\mu|\mathbf{x}^\top(\theta_1 - \theta_2)|$, so the instantaneous regret $r_{t,i}$ is uniformly bounded $\forall t \in [T], i \in [N]$ by $2k_\mu S$. Now suppose epoch p is bad, then we can upper bound the cumulative regret in this bad epoch by $2k_\mu S\frac{NT}{B}$, where $\frac{NT}{B}$ is the number of time steps in each epoch. Since there can be at most $O(d\log NT)$ bad epochs, the cumulative regret incurred in all bad epochs can be upper bounded by $O(\frac{NT}{B}k_\mu S d\log(NT))$. Combining both parts together, the cumulative regret upper bound is

$$R_T = O\left(\frac{NT}{B}k_\mu S d\log(NT) + \frac{k_\mu R_{\max} d}{c_\mu}\sqrt{NT}\log(NT)\right)$$

To recover the regret under centralized setting, we set $B = \sqrt{NT}$, so

$$R_T = O\left(\frac{k_\mu R_{max}}{c_\mu} d\sqrt{NT}\log(NT)\right)$$

Note that FedGLB-UCB$_1$ has $B = \sqrt{NT}$ global updates in total, and during each global update, there are J_t rounds of communications, for $t \in \mathcal{S}$. As mentioned earlier, for AGD to attain $\epsilon_t = \frac{1}{N^2t^2}$ sub-optimality, the required number of inner iterations

$$J_t \leq 1 + \sqrt{\frac{k_\mu + \frac{\lambda}{Nt}}{\frac{\lambda}{Nt}}} \log\frac{(k_\mu + \frac{\lambda}{Nt} + \frac{\lambda}{Nt})\|\theta_t^{(0)} - \hat{\theta}_t^{\text{MLE}}\|_2^2}{2\epsilon_t} = O\left(\sqrt{Nt}\log(Nt)\right)$$

Therefore, the communication cost over time horizon T is

$$
\begin{aligned}
C_T &= N \cdot \sum_{t \in \mathcal{S}} J_t \\
&= N \cdot \big[\sqrt{\sqrt{NT}}\log(\sqrt{NT}) + \sqrt{2\sqrt{NT}}\log(2\sqrt{NT}) + \cdots + \sqrt{\sqrt{NT}\cdot\sqrt{NT}}\log(\sqrt{NT}\cdot\sqrt{NT})\big] \\
&\leq N^{5/4}T^{1/4}\log(NT)\big[\sqrt{1} + \sqrt{2} + \cdots + \sqrt{\sqrt{NT}}\big] \\
&\leq N^{5/4}T^{1/4}\log(NT)\cdot\frac{3}{2}(\sqrt{NT} + \frac{1}{2})^{3/2} \\
&= O(N^2 T \log(NT))
\end{aligned}
$$

which finishes the proof. □

E.2 FedGLB-UCB$_2$: scheduled communication

For the second variant FedGLB-UCB$_2$, we enabled local update on top of FedGLB-UCB$_1$. Therefore, compared with the original algorithm FedGLB-UCB, the only difference is that FedGLB-UCB$_2$ uses scheduled communication instead of event-triggered communication. Its description is given in Algorithm 5.

Algorithm 5 FedGLB-UCB$_2$

Input: communication schedule $\mathcal{S}$, regularization parameter $\lambda > 0$, $\delta \in (0,1)$ and c_μ.
Initialize $\forall i \in [N]$: $A_{0,i} = \frac{\lambda}{c_\mu}\mathbf{I} \in \mathbb{R}^{d\times d}, b_{0,i} = \mathbf{0} \in \mathbb{R}^d, \theta_{0,i} = \mathbf{0} \in \mathbb{R}^d, \Delta A_{0,i} = \mathbf{0} \in \mathbb{R}^{d\times d}$; $A_0 = \frac{\lambda}{c_\mu}\mathbf{I} \in \mathbb{R}^{d\times d}, b_0 = \mathbf{0} \in \mathbb{R}^d, \theta_0 = \mathbf{0} \in \mathbb{R}^d$, $t_{\text{last}} = 0$
for $t = 1, 2, ..., T$ **do**
  for client $i = 1, 2, ..., N$ **do**
    Observe arm set $\mathcal{A}_{t,i}$ for client i
    Select arm $\mathbf{x}_{t,i} \in \mathcal{A}_{t,i}$ by Eq.(5), and observe reward $y_{t,i}$
    Update client i: $A_{t,i} = A_{t-1,i} + \mathbf{x}_{t,i}\mathbf{x}_{t,i}^\top$, $\Delta A_{t,i} = \Delta A_{t-1,i} + \mathbf{x}_{t,i}\mathbf{x}_{t,i}^\top$
  if $t \notin \mathcal{S}$ **then**
    Clients $\forall i \in [N]$: $\theta_{t,i} =$ ONS-Update$(\theta_{t-1,i}, A_{t,i}, \nabla l(\mathbf{x}_{t,i}^\top\theta_{t-1,i}, y_{t,i}))$, $b_{t,i} = b_{t-1,i} + \mathbf{x}_{t,i}\mathbf{x}_{t,i}^\top\theta_{t-1,i}$
  else
    Clients $\forall i \in [N]$: send $\Delta A_{t,i}$ to server, and reset $\Delta A_{t,i} = \mathbf{0}$
    Server compute $A_t = A_{t_{\text{last}}} + \sum_{i=1}^N \Delta A_{t,i}$
    Server perform global model update $\theta_t =$ AGD-Update$(\theta_{t_{\text{last}}}, J_t)$ (see Eq.(3) for choice of J_t), $b_t = b_{t_{\text{last}}} + \sum_{i=1}^N \Delta A_{t,i}\theta_t$, and set $t_{\text{last}} = t$
    Clients $\forall i \in [N]$: set $\theta_{t,i} = \theta_t, A_{t,i} = A_t, b_{t,i} = b_t$

The regret and communication cost of FedGLB-UCB$_2$ is given in the following theorem.

Theorem E.2 (Regret and Communication Cost Upper Bound of FedGLB-UCB$_2$)**.** *Under the condition that $\epsilon_t = \frac{1}{N^2t^2}$, and the total number of global synchronizations $B = d^2 N\log(NT)$, the cumulative regret R_T has upper bound*

$$R_T = O\left(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu} d\sqrt{NT}\log(NT/\delta)\sqrt{\log\frac{T}{d^2N\log NT}}\right)$$

with probability at least $1-\delta$. The cumulative communication cost has upper bound

$$C_T = O(d^2N^{2.5}\sqrt{T}\log^2(NT))$$

Proof. Compared with the analysis for FedGLB-UCB, the main difference in the analysis for FedGLB-UCB$_2$ is how we bound the regret incurred in 'bad epochs'. Using the same argument, the cumulative regret for the 'good epochs' is $REG_{good} = O(\frac{k_\mu(k_\mu+R_{\max})}{c_\mu}d\sqrt{NT}\log NT/\delta)$.

Now consider a particular bad epoch $p \in [B]$. Then the cumulative regret incurred by all N clients during this 'bad epoch' can be upper bounded by:

$$\sum_{t=t_{p-1}+1}^{t_p} \sum_{i=1}^{N} r_{t,i}$$
$$\leq O(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu}\sqrt{d\log(NT/\delta)}) \sum_{t=t_{p-1}+1}^{t_p} \sum_{i=1}^{N} \min(1, ||\mathbf{x}_{t,i}||_{A_{t-1,i}^{-1}})$$
$$\leq O(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu}\sqrt{d\log(NT/\delta)}) \sum_{i=1}^{N} \sqrt{(t_p - t_{p-1})\log\frac{\det(A_{t_p-1,i})}{\det(A_{t_p-1,i} - \Delta A_{t_p-1,i})}}$$
$$\leq O(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu} dN\sqrt{\log(NT/\delta)}\sqrt{\frac{T}{B}\log(\frac{T}{B})})$$

where the last inequality is because all epochs has length $\frac{T}{B}$ as defined by $\mathcal{S}$. Again, since there can be at most $O(d\log NT)$ 'bad epochs', the cumulative regret for the 'bad epochs' is upper bounded by

$$REG_{bad} = O(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu} d^2 \log^{1.5}(NT/\delta) N\sqrt{\frac{T}{B}\log(\frac{T}{B})}).$$

Combining the cumulative regret for both 'good' and 'bad' epochs, and setting $B = d^2 N\log(NT)$, we have

$$R_T = O\left(\frac{k_\mu(k_\mu + R_{\max})}{c_\mu} d\sqrt{NT}\log(NT/\delta)\sqrt{\log(\frac{T}{d^2 N\log NT})}\right)$$

Now that FedGLB-UCB$_2$ has $B = d^2 N\log(NT)$ global updates in total, and during each global update, there are $J_t = O(\sqrt{NT}\log(NT))$ rounds of communications, for $t \in \mathcal{S}$. Therefore, the communication cost over time horizon T is

$$C_T = N \cdot \sum_{t\in\mathcal{S}} J_t = O(N \cdot d^2 N\log(NT) \cdot \sqrt{NT}\log(NT))$$
$$= O(d^2 N^{2.5}\sqrt{T}\log^2(NT))$$

which finishes the proof. □

E.3 FedGLB-UCB$_3$: scheduled communication + ONS for global update

The previous two variants both adopt iterative optimization method, i.e., AGD, for the global update, which introduces a $\sqrt{NT}\log(NT)$ factor in the communication cost. In this section, we try to avoid this by studying the third variant FedGLB-UCB$_3$ that adopts ONS for both local and global update, such that only one step of ONS is performed (based on all new data samples N clients collected in this epoch). It can be viewed as the ONS-GLM algorithm [12] with lazy batch update.

Recall that the update schedule is denoted as $\mathcal{S} = \{t_1 := \lfloor\frac{T}{B}\rfloor, t_2 := 2\lfloor\frac{T}{B}\rfloor, \dots, t_q := q\lfloor\frac{T}{B}\rfloor, \dots, t_B := B\lfloor\frac{T}{B}\rfloor\}$, where B denotes the total number of global updates up to T. Compared with [12], the main difference in our construction is that the loss function in the online regression problem may contain multiple data samples, i.e., for global update, or one single data sample, i.e., for local update. Then for a client $i \in [N]$ at time step t (suppose t is in the $(q+1)$-th epoch, so $t \in [t_q + 1, t_{q+1}]$), the sequence of loss functions observed by the online regression estimator till time t is:

$$\underbrace{\sum_{s=1}^{t_1}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\theta_0, y_{s,i}), \sum_{s=t_1+1}^{t_2}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\theta_{t_1}, y_{s,i}), \dots, \sum_{s=t_{q-1}+1}^{t_q}\sum_{i=1}^{N} l(\mathbf{x}_{s,i}^\top\theta_{t_{q-1}}, y_{s,i})}_{\text{global updates at } t_1, t_2, \dots, t_q}, \underbrace{l(\mathbf{x}_{t_q+1,i}^\top\theta_{t_q}, y_{t_q+1,i}), \dots, l(\mathbf{x}_{t,i}^\top\theta_{t-1,i}, y_{t,i})}_{\text{local updates at } t_q+1, \dots, t}$$

We can see that the first q terms correspond to the global ONS updates that are computed using the whole batch of data collected by N clients in each epoch, and the remaining $t - t_q$ terms are local

Algorithm 6 FedGLB-UCB$_3$

Input: communication schedule $\mathcal{S}$, regularization parameter $\lambda > 0$, $\delta \in (0,1)$ and c_μ
Initialize $\forall i \in [N]$: $\theta_{0,i} = \mathbf{0} \in \mathbb{R}^d, A_{0,i} = \lambda\mathbf{I} \in \mathbb{R}^{d\times d}, V_{0,i} = \lambda\mathbf{I} \in \mathbb{R}^{d\times d}, b_{0,i} = \mathbf{0} \in \mathbb{R}^d$; $\theta_0 = \mathbf{0} \in \mathbb{R}^d, A_0 = \lambda\mathbf{I} \in \mathbb{R}^{d\times d}, V_0 = \lambda\mathbf{I} \in \mathbb{R}^{d\times d}, b_0 = \mathbf{0} \in \mathbb{R}^d$, $t_{\text{last}} = 0$
for $t = 1, 2, ..., T$ **do**
  for client $i = 1, 2, ..., N$ **do**
    Observe arm set $\mathcal{A}_{t,i}$ for client $i \in [N]$
    Select arm $\mathbf{x}_{t,i} = \arg\max_{\mathbf{x}\in\mathcal{A}_{t,i}} \mathbf{x}^\top \hat{\theta}_{t-1,i} + \alpha_{t-1,i}\|\mathbf{x}\|_{V_{t-1,i}^{-1}}$, where $\hat{\theta}_{t-1,i} = V_{t-1,i}^{-1} b_{t-1,i}$ and $\alpha_{t-1,i}$ is given in Lemma E.4; and then observe reward $y_{t,i}$
    Compute loss $l(z_{t,i}, y_{t,i})$, where $z_{t,i} = \mathbf{x}_{t,i}^\top \theta_{t-1,i}$
    Update client i: $A_{t,i} = A_{t-1,i} + \nabla l(z_{t,i}, y_{t,i})\nabla l(z_{t,i}, y_{t,i})^\top$, $V_{t,i} = V_{t-1,i} + \mathbf{x}_{t,i}\mathbf{x}_{t,i}^\top$
  if $t \notin \mathcal{S}$ **then**
    Clients $\forall i \in [N]$: $\theta_{t,i} =$ ONS-Update$(\theta_{t-1,i}, A_{t,i}, \nabla l(z_{t,i}, y_{t,i}))$, $b_{t,i} = b_{t-1,i} + \mathbf{x}_{t,i} z_{t,i}$
  else
    Clients $\forall i \in [N]$: send gradient $\nabla F_{t,i}(\theta_{t_{\text{last}}}) = \sum_{s=t_{\text{last}}+1}^{t} \nabla l(\mathbf{x}_{s,i}^\top \theta_{t_{\text{last}}}, y_{s,i})$ and $\Delta V_{t,i} = V_{t,i} - V_{t_{\text{last}},i}$ to server
    Server $A_t = A_{t_{\text{last}}} + (\sum_{i=1}^N \nabla F_{t,i}(\theta_{t_{\text{last}}}))(\sum_{i=1}^N \nabla F_{t,i}(\theta_{t_{\text{last}}}))^\top$, $V_t = V_{t_{\text{last}}} + \sum_{i=1}^N \Delta V_{t,i}$, $b_t = b_{t_{\text{last}}} + \sum_{i=1}^N \Delta V_{t,i}\theta_{t_{\text{last}}}$, $\theta_t =$ ONS-Update$(\theta_{t_{\text{last}}}, A_t, \sum_{i=1}^N \nabla F_{t,i}(\theta_{t_{\text{last}}}))$
    Clients $\forall i \in [N]$: $\theta_{t,i} = \theta_t, A_{t,i} = A_t, V_t = V_t, b_{t,i} = b_t$
    Set $t_{\text{last}} = t$

ONS updates that are computed using each new data sample collected by client i in the $(q+1)$-th epoch.

To facilitate further analysis, we introduce a new set of indices for the data samples, so that we can unify the notations for the loss functions above. Imagine all the arm pulls are performed by an imaginary centralized agent, such that, in each time step $t \in [T]$, it pulls an arm for clients $1, 2, \ldots, N$ one by one. Therefore, the sequence of data sample obtained by this imaginary agent can be denoted as $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_s, y_s), \ldots, (\mathbf{x}_{NT}, y_{NT})$. Moreover, we denote n_p as the total number of data samples collected by all N clients till the p-th ONS update (including both global and local ONS update), and denote the updated model as θ_p, for $p \in [P]$. Note that P denotes the total number of updates up to time t (total number of terms in the sequence above), such that $P = q + t - t_q$. Then this sequence of loss functions can be rewritten as:

$$\underbrace{F_1(\theta_0), F_2(\theta_1), \ldots, F_q(\theta_{q-1})}_{\text{global updates}}, \underbrace{F_{q+1}(\theta_q), \ldots, F_P(\theta_{P-1})}_{\text{local updates}}$$

where $F_p(\theta_{p-1}) = \sum_{s=n_{p-1}+1}^{n_p} l(\mathbf{x}_s^\top \theta_{p-1}, y_s)$, for $p \in [P]$.

• **Online regret upper bound for lazily-updated ONS** To construct the confidence ellipsoid based on this sequence of global and local ONS updates, we first need to upper bound the online regret that ONS incurs on this sequence of loss functions, which is given in Lemma E.3.

Lemma E.3 (Online regret upper bound)**.** *Under the condition that the learning rate of ONS is set to* $\gamma = \frac{1}{2}\min(\frac{1}{4S\sqrt{k_\mu^2 S^2 + R_{\max}^2}}, \frac{c_\mu}{(k_\mu^2 S^2 + R_{\max}^2)\max_{p\in[P]}(n_p - n_{p-1})})$*, then the cumulative online regret over* P *steps*

$$\sum_{p=1}^{P} F_p(\theta_{p-1}) - F_p(\theta_\star) \leq B_P$$

where $B_P = \frac{1}{2\gamma}\sum_{p=1}^{P} ||\nabla F_p(\theta_{p-1})||_{A_p^{-1}}^2 + 2\gamma\lambda S^2$.

Proof of Lemma E.3. Recall from the proof of Corollary 4.2.1 that $|\mu(\mathbf{x}_s^\top \theta_{p-1}) - y_s| \leq \sqrt{k_\mu^2 S^2 + R_{\max}^2} := G, \forall s$. First, we need to show that $F_p(\theta_{p-1}) = \sum_{s=n_{p-1}+1}^{n_p} l(\mathbf{x}_s^\top \theta_{p-1}, y_s)$ is $\frac{c_\mu}{(n_p - n_{p-1})G^2}$-exp-concave, or equivalently, $\nabla^2 F_p(\theta_{p-1}) \succcurlyeq \frac{c_\mu}{(n_p - n_{p-1})G^2}\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top$

(Lemma 4.2 in [9]). Taking first and second order derivative of $F_p(\theta_{p-1})$ w.r.t. θ_{p-1}, we have

$$\nabla F_p(\theta_{p-1}) = \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s[-y_s + \mu(\mathbf{x}_s^\top \theta_{p-1})] = \mathbf{X}_p^\top[\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p],$$
$$\nabla^2 F_p(\theta_{p-1}) = \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s\mathbf{x}_s^\top \dot{\mu}(\mathbf{x}_s^\top\theta_{p-1})$$

where $\mathbf{X}_p = [\mathbf{x}_{n_{p-1}+1}, \mathbf{x}_{n_{p-1}+2}, \dots, \mathbf{x}_{n_p}]^\top \in \mathbb{R}^{(n_p-n_{p-1})\times d}$, and $\mathbf{y}_p = [y_{n_{p-1}+1}, y_{n_{p-1}+2}, \dots, y_{n_p}]^\top \in \mathbb{R}^{n_p-n_{p-1}}$. Then due to Assumption 1, we have $\nabla^2 F_p(\theta_{p-1}) \succcurlyeq c_\mu \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s\mathbf{x}_s^\top = c_\mu \mathbf{X}_p^\top\mathbf{X}_p$. For any vector $u \in \mathbb{R}^d$, we can show that,

$$\begin{aligned}
&u^\top \nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top u \\
&= u^\top \mathbf{X}_p^\top[\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p][\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p]^\top \mathbf{X}_p u \\
&= \big[(\mathbf{X}_p u)^\top[\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p]\big]^2 \\
&\le \|\mathbf{X}_p u\|_2^2 \cdot \|\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p\|_2^2 \\
&\le u^\top \mathbf{X}_p^\top\mathbf{X}_p u \cdot (n_p - n_{p-1})G^2
\end{aligned}$$

where the first inequality is due to Cauchy-Schwarz inequality, and the second inequality is because $\|\mu(\mathbf{X}_p\theta_{p-1}) - \mathbf{y}_p\|_2^2 = \sum_{s=n_{p-1}+1}^{n_p}[-y_s + \mu(\mathbf{x}_s^\top\theta_{p-1})]^2 \le (n_p - n_{p-1})G^2$. Therefore, $\mathbf{X}_p^\top\mathbf{X}_p \succcurlyeq \frac{1}{(n_p-n_{p-1})G^2}\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top$, which gives us

$$\nabla^2 F_p(\theta_{p-1}) \succcurlyeq \frac{c_\mu}{(n_p - n_{p-1})G^2}\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top$$

Then due to Lemma 4.3 of [9], under the condition that $\gamma_p \le \frac{1}{2}\min(\frac{1}{4GS}, \frac{c_\mu}{(n_p-n_{p-1})G^2})$, we have

$$\begin{aligned}
&F_p(\theta_{p-1}) - F_p(\theta_\star) \\
&\le \nabla F_p(\theta_{p-1})^\top(\theta_{p-1} - \theta_\star) - \frac{\gamma_p}{2}(\theta_{p-1} - \theta_\star)^\top \nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top(\theta_{p-1} - \theta_\star)
\end{aligned} \tag{17}$$

Then we start to upper bound the RHS of the inequality above. Recall that the ONS update rule is:

$$\theta_p' = \theta_{p-1} - \frac{1}{\gamma}A_p^{-1}\nabla F_p(\theta_{p-1})$$
$$\theta_p = \arg\min_{\theta\in\Theta} ||\theta_p' - \theta||_{A_p}^2$$

where $A_p = \sum_{\rho=1}^p \nabla F_\rho(\theta_{\rho-1})\nabla F_\rho(\theta_{\rho-1})^\top$, and γ is set to $\min_{p\in[P]}\gamma_p = \frac{1}{2}\min(\frac{1}{4GS}, \frac{c_\mu}{G^2\max_{p\in[P]}(n_p-n_{p-1})})$. So we have

$$\theta_p' - \theta_\star = \theta_{p-1} - \theta_\star - \frac{1}{\gamma}A_p^{-1}\nabla F_p(\theta_{p-1})$$

Then due to the property of the generalized projection, and by substituting into the update rule, we have

$$||\theta_p - \theta_\star||_{A_p}^2 \le ||\theta_p' - \theta_\star||_{A_p}^2 \le ||\theta_{p-1} - \theta_\star||_{A_p}^2 - \frac{2}{\gamma}(\theta_{p-1} - \theta_\star)^\top\nabla F_p(\theta_{p-1}) + \frac{1}{\gamma^2}||\nabla F_p(\theta_{p-1})||_{A_p^{-1}}^2$$

By rearranging terms,

$$\nabla F_p(\theta_{p-1})^\top(\theta_{p-1} - \theta_\star) \le \frac{1}{2\gamma}||\nabla F_p(\theta_{p-1})||_{A_p^{-1}}^2 + \frac{\gamma}{2}\big(||\theta_{p-1} - \theta_\star||_{A_p}^2 - ||\theta_p - \theta_\star||_{A_p}^2\big)$$

After summing over P steps, we have

$$\sum_{p=1}^P \nabla F_p(\theta_{p-1})^\top(\theta_{p-1} - \theta_\star) \le \frac{1}{2\gamma}\sum_{p=1}^P ||\nabla F_p(\theta_{p-1})||_{A_p^{-1}}^2 + \frac{\gamma}{2}\sum_{p=1}^P\big(||\theta_{p-1} - \theta_\star||_{A_p}^2 - ||\theta_p - \theta_\star||_{A_p}^2\big)$$

The second term can be simplified,

$$
\begin{aligned}
&\sum_{p=1}^{P}\big(||\theta_{p-1}-\theta_\star||^2_{A_p}-||\theta_p-\theta_\star||^2_{A_p}\big)\\
&=||\theta_0-\theta_\star||^2_{A_1}+\sum_{p=2}^{P}\big(||\theta_{p-1}-\theta_\star||^2_{A_p}-||\theta_{p-1}-\theta_\star||^2_{A_{p-1}}\big)-||\theta_P-\theta_\star||^2_{A_P}\\
&\le ||\theta_0-\theta_\star||^2_{A_1}+\sum_{p=2}^{P}\big(||\theta_{p-1}-\theta_\star||^2_{A_p}-||\theta_{p-1}-\theta_\star||^2_{A_{p-1}}\big)\\
&=||\theta_0-\theta_\star||^2_{A_1}+\sum_{p=2}^{P}||\theta_{p-1}-\theta_\star||^2_{\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top}\\
&=||\theta_0-\theta_\star||^2_{A_1}+\sum_{p=1}^{P}||\theta_{p-1}-\theta_\star||^2_{\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top}-||\theta_0-\theta_\star||^2_{\nabla F_1(\theta_0)\nabla F_1(\theta_0)^\top}\\
&=4\lambda S^2+\sum_{p=1}^{P}||\theta_{p-1}-\theta_\star||^2_{\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top}
\end{aligned}
$$

which leads to

$$
\begin{aligned}
\sum_{p=1}^{P}\nabla F_p(\theta_{p-1})^\top(\theta_{p-1}-\theta_\star)\le \frac{1}{2\gamma}\sum_{p=1}^{P}||\nabla F_p(\theta_{p-1})||^2_{A_p^{-1}}+2\gamma\lambda S^2\\
+\frac{\gamma}{2}\sum_{p=1}^{P}||\theta_{p-1}-\theta_\star||^2_{\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top}
\end{aligned}
$$

By rearranging terms, we have

$$
\begin{aligned}
&\sum_{p=1}^{P}\big[\nabla F_p(\theta_{p-1})^\top(\theta_{p-1}-\theta_\star)-\frac{\gamma}{2}||\theta_{p-1}-\theta_\star||^2_{\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top}\big]\\
&\le \frac{1}{2\gamma}\sum_{p=1}^{P}||\nabla F_p(\theta_{p-1})||^2_{A_p^{-1}}+2\gamma\lambda S^2
\end{aligned}
$$

Combining with Eq.(17), we obtain the following upper bound for the P-step online regret

$$
\sum_{p=1}^{P}F_p(\theta_{p-1})-F_p(\theta_\star)\le \frac{1}{2\gamma}\sum_{p=1}^{P}||\nabla F_p(\theta_{p-1})||^2_{A_p^{-1}}+2\gamma\lambda S^2
$$

where $A_p=\sum_{\rho=1}^{p}\nabla F_\rho(\theta_{\rho-1})\nabla F_\rho(\theta_{\rho-1})^\top$. ☐

Corollary E.3.1 (Order of B_P)**.** *Under the condition that* $\gamma=\frac{1}{2}\min(\frac{1}{4S\sqrt{k_\mu^2S^2+R_{\max}^2}},\frac{c_\mu}{(k_\mu^2S^2+R_{\max}^2)\max_{p\in[P]}(n_p-n_{p-1})})$, *the online regret upper bound* $B_P=O(\frac{k_\mu^2+R_{max}^2}{c_\mu}d\log(n_P)\max_{p\in[P]}(n_p-n_{p-1}))$.

Proof of Corollary E.3.1. Recall that $A_p=\sum_{\rho=1}^{p}\nabla F_\rho(\theta_{\rho-1})\nabla F_\rho(\theta_{\rho-1})^\top$. Therefore, we have

$$
\begin{aligned}
\sum_{p=1}^{P}||\nabla F_p(\theta_{p-1})||^2_{A_p^{-1}}&\le\log\frac{\det(A_P)}{\det(\lambda I)}=\log\frac{\det(\lambda I+\sum_{p=1}^{P}\nabla F_p(\theta_{p-1})\nabla F_p(\theta_{p-1})^\top)}{\det(\lambda I)}\\
&\le d\log\big(1+\frac{1}{d\lambda}\sum_{p=1}^{P}\|\nabla F_p(\theta_{p-1})\|_2^2\big)
\end{aligned}
$$

where the first inequality is due to Lemma 11 of [1], and the second due to the determinant-trace inequality (Lemma 10 of [1]), i.e., $\det(\lambda I + \sum_{p=1}^{P} \nabla F_p(\theta_{p-1}) \nabla F_p(\theta_{p-1})^\top) \leq \big(\frac{tr(\lambda I + \sum_{p=1}^{P} \nabla F_p(\theta_{p-1}) \nabla F_p(\theta_{p-1})^\top)}{d}\big)^d = \big(\frac{d\lambda + \sum_{p=1}^{P} \|\nabla F_p(\theta_{p-1})\|_2^2}{d}\big)^d$. Also note that $\nabla F_p(\theta_{p-1}) = \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s \big[\mu(\mathbf{x}_s^\top \theta_{p-1}) - y_s\big]$, so we have

$$\sum_{p=1}^{P} ||\nabla F_p(\theta_{p-1})||_2^2 = \sum_{p=1}^{P} || \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s \big[\mu(\mathbf{x}_s^\top \theta_{p-1}) - y_s\big] ||_2^2$$

$$\leq G^2 \sum_{p=1}^{P} || \sum_{s=n_{p-1}+1}^{n_p} \mathbf{x}_s ||_2^2 \leq G^2 \sum_{p=1}^{P} (n_p - n_{p-1})^2 \leq G^2 n_P^2$$

where the second inequality is due to Jensen's inequality and the assumption that $\|\mathbf{x}_s\| \leq 1, \forall s$. Substituting this back gives us

$$\sum_{p=1}^{P} F_p(\theta_{p-1}) - F_p(\theta_\star) \leq \frac{1}{2\gamma} d \log\big(1 + \frac{1}{d\lambda} G^2 n_P^2\big) + 2\gamma\lambda S^2$$

$$= \frac{(k_\mu^2 S^2 + R_{\max}^2) \max_{p\in[P]}(n_p - n_{p-1})}{c_\mu} d \log\big(1 + \frac{1}{d\lambda}(k_\mu^2 S^2 + R_{\max}^2) n_P^2\big)$$

$$+ \frac{c_\mu}{(k_\mu^2 S^2 + R_{\max}^2) \max_{p\in[P]}(n_p - n_{p-1})} \lambda S^2$$

where the equality is because $\max_{p\in[P]}(n_p - n_{p-1})$ dominates $\gamma = \frac{1}{2}\min(\frac{1}{4GS}, \frac{c_\mu}{G^2 \max_{p\in[P]}(n_p - n_{p-1})})$. □

• **Construct Confidence Ellipsoid for FedGLB-UCB$_3$** With the online regret bound B_P in Lemma E.3, the steps to construct the confidence ellipsoid largely follows that of Theorem 1 in [12], with the main difference in our batch update. We include the full proof here for the sake of completeness.

Lemma E.4 (Confidence Ellipsoid for FedGLB-UCB$_3$)**.** *Under the condition that the learning rate of ONS* $\gamma = \frac{1}{2}\min(\frac{1}{4S\sqrt{k_\mu^2 S^2 + R_{\max}^2}}, \frac{c_\mu}{(k_\mu^2 S^2 + R_{\max}^2)\max_{p\in[P]}(n_p - n_{p-1})})$*, we have* $\forall t \in [T], i \in [N]$

$$\|\theta_\star - \hat{\theta}_{t,i}\|_{V_{t,i}}^2 \leq \lambda S^2 + 1 + \frac{4}{c_\mu} B_P + \frac{8R_{max}^2}{c_\mu^2} \log\big(\frac{N}{\delta}\sqrt{4 + \frac{8}{c_\mu} B_P + \frac{64R_{max}^2}{c_\mu^4 \cdot 4\delta^2}}\big)\}$$

$$- \hat{\theta}_{t,i}^\top b_{t,i} - \sum_{s=1}^{n_P} z_s^2 := \alpha_{t,i}^2$$

with probability at least $1 - \delta$.

Proof of Lemma E.4. Due to c_μ-strongly convexity of $l(z, y)$ w.r.t. z, we have $l(\mathbf{x}_s^\top \theta_{p-1}, y_s) - l(\mathbf{x}_s^\top \theta_\star, y_s) \geq \big[\mu(\mathbf{x}_s^\top \theta_\star) - y_s\big]\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star) + \frac{c_\mu}{2}\big[\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star)\big]^2$. Therefore,

$$F_p(\theta_{p-1}) - F_p(\theta_\star) = \sum_{s=n_{p-1}+1}^{n_p} l(\mathbf{x}_s^\top \theta_{p-1}, y_s) - l(\mathbf{x}_s^\top \theta_\star, y_s)$$

$$\geq \sum_{s=n_{p-1}+1}^{n_p} \big[\mu(\mathbf{x}_s^\top \theta_\star) - y_s\big]\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star) + \frac{c_\mu}{2} \sum_{s=n_{p-1}+1}^{n_p} \big[\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star)\big]^2$$

$$= - \sum_{s=n_{p-1}+1}^{n_p} \eta_s \mathbf{x}_s^\top(\theta_{p-1} - \theta_\star) + \frac{c_\mu}{2} \sum_{s=n_{p-1}+1}^{n_p} \big[\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star)\big]^2$$

where η_s is the R-sub-Gaussian noise in the reward y_s. Summing over P steps we have

$$B_P \geq \sum_{p=1}^{P} F_p(\theta_{p-1}) - F_p(\theta_\star) \geq \sum_{p=1}^{P} \sum_{s=n_{p-1}+1}^{n_p} \eta_s \mathbf{x}_s^\top(\theta_{p-1} - \theta_\star) + \frac{c_\mu}{2} \sum_{p=1}^{P} \sum_{s=n_{p-1}+1}^{n_p} \big[\mathbf{x}_s^\top(\theta_{p-1} - \theta_\star)\big]^2$$

By rearranging terms, we have

$$\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\big[\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)\big]^2 \le \frac{2}{c_\mu}\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\eta_s\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)+\frac{2}{c_\mu}B_P$$

Then as $\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)$ for $s\in[n_{p-1}+1,n_p]$ is $\mathcal{F}_s$-measurable for lazily updated online estimator θ_{p-1}, we can use Corollary 8 from [2], which leads to

$$\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\eta_s\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)\le$$
$$R_{max}\sqrt{\big(2+2\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}(\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star))^2\big)\cdot\log\big(\frac{1}{\delta}\sqrt{1+\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}(\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star))^2}\big)}$$

Then we have

$$\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\big[\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)\big]^2\le\frac{2}{c_\mu}B_P$$
$$+\frac{2R_{max}}{c_\mu}\sqrt{\big(2+2\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}(\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star))^2\big)\cdot\log\big(\frac{1}{\delta}\sqrt{1+\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}(\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star))^2}\big)}$$

Then by applying Lemma 2 from [12], we have

$$\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\big[\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)\big]^2\le 1+\frac{4}{c_\mu}B_P+\frac{8R_{max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{4+\frac{8}{c_\mu}B_P+\frac{64R_{max}^2}{c_\mu^4\cdot 4\delta^2}}\big)$$

Therefore, we have the following confidence ellipsoid (regularized with parameter λ):

$$\{\theta:\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\big[\mathbf{x}_s^\top(\theta_{p-1}-\theta_\star)\big]^2+\lambda\|\theta\|_2^2\le\lambda S^2+1+\frac{4}{c_\mu}B_P+\frac{8R_{max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{4+\frac{8}{c_\mu}B_P+\frac{64R_{max}^2}{c_\mu^4\cdot 4\delta^2}}\big)\}$$

And this can be rewritten as a ellipsoid centered at ridge regression estimator $\hat{\theta}_{t,i}=V_{t,i}^{-1}b_{t,i}$, where $V_{t,i}=\lambda I+\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\mathbf{x}_s\mathbf{x}_s^\top$ and $b_{t,i}=\sum_{p=1}^{P}\sum_{s=n_{p-1}+1}^{n_p}\mathbf{x}_s z_s$ (recall that ONS's prediction at time s is denoted as $z_s=\mathbf{x}_s^\top\theta_{p-1}$), i.e., $\forall t\in[T]$

$$\|\theta_\star-\hat{\theta}_{t,i}\|_{V_{t,i}}^2\le\lambda S^2+1+\frac{4}{c_\mu}B_P+\frac{8R_{max}^2}{c_\mu^2}\log\big(\frac{1}{\delta}\sqrt{4+\frac{8}{c_\mu}B_P+\frac{64R_{max}^2}{c_\mu^4\cdot 4\delta^2}}\big)\}+\hat{\theta}_{t,i}^\top b_{t,i}-\sum_{s=1}^{n_P}z_s^2$$

with probability at least $1-\delta$. Then taking union bound over all N clients, we have, $\forall t\in[T],i\in[N]$

$$\|\theta_\star-\hat{\theta}_{t,i}\|_{V_{t,i}}^2\le\lambda S^2+1+\frac{4}{c_\mu}B_P+\frac{8R_{max}^2}{c_\mu^2}\log\big(\frac{N}{\delta}\sqrt{4+\frac{8}{c_\mu}B_P+\frac{64R_{max}^2}{c_\mu^4\cdot 4\delta^2}}\big)\}+\hat{\theta}_{t,i}^\top b_{t,i}-\sum_{s=1}^{n_P}z_s^2$$

with probability at least $1-\delta$. □

• Regret and Communication Upper Bounds for FedGLB-UCB$_3$

The regret and communication cost of FedGLB-UCB$_3$ is given in the following theorem.

Theorem E.5 (Regret and Communication Cost Upper Bound of FedGLB-UCB$_3$)**.** *Under the condition that the learning rate of ONS* $\gamma=\frac{1}{2}\min(\frac{1}{4S\sqrt{k_\mu^2S^2+R_{\max}^2}},\frac{c_\mu}{(k_\mu^2S^2+R_{\max}^2)\sqrt{NT}})$*, and the total number of global synchronizations* $B=\sqrt{NT}$*, the cumulative regret* R_T *has upper bound*

$$R_T=O\left(\frac{k_\mu(k_\mu+R_{max})}{c_\mu}dN^{3/4}T^{3/4}\log(NT/\delta)\right)$$

with probability at least $1-\delta$*. The cumulative communication cost has upper bound*

$$C_T=O(N^{1.5}\sqrt{T})$$

Proof. Similar to the proof for the previous two variants of FedGLB-UCB, we divide the epochs into 'good' and 'bad' ones according to the determinant ratio, and then bound their cumulative regret separately.

Recall that the instantaneous regret $r_{t,i}$ incurred by client $i \in [N]$ at time step $t \in [T]$ has upper bound

$$\begin{aligned}\frac{r_{t,i}}{k_\mu} &\leq \mathbf{x}_{t,\star}^\top \theta_\star - \mathbf{x}_{t,i}^\top \theta_\star \leq \mathbf{x}_{t,i}^\top \tilde{\theta}_{i,t} - \mathbf{x}_{t,i}^\top \theta_\star \\ &= \mathbf{x}_{t,i}^\top (\tilde{\theta}_{i,t} - \hat{\theta}_{t,i}) + \mathbf{x}_{t,i}^\top (\hat{\theta}_{t,i} - \theta_\star) \\ &\leq \|\mathbf{x}_{t,i}\|_{V_{t,i}^{-1}} \|\tilde{\theta}_{i,t} - \hat{\theta}_{t,i}\|_{V_{t,i}} + \|\mathbf{x}_{t,i}\|_{V_{t,i}^{-1}} \|\hat{\theta}_{t,i} - \theta_\star\|_{V_{t,i}} \\ &\leq 2\alpha_{t,i} \|\mathbf{x}_{t,i}\|_{V_{t,i}^{-1}}\end{aligned}$$

Note that due to the update schedule $\mathcal{S}$, we have $\max_{p \in [P]} (n_p - n_{p-1}) = \frac{NT}{B}$. Then based on Corollary E.3.1, $\alpha_{t,i} = O(\frac{k_\mu + R_{max}}{c_\mu} \sqrt{d \log(NT)} \sqrt{\frac{NT}{B}})$, so we have, $\forall t \in [T], i \in [N]$,

$$r_{t,i} = O(\frac{k_\mu (k_\mu + R_{max})}{c_\mu} \sqrt{d \log(NT)} \sqrt{\frac{NT}{B}}) \|\mathbf{x}_{t,i}\|_{A_{t,i}^{-1}}$$

with probability at least $1 - \delta$.

Therefore, the cumulative regret for the 'good epochs' is $REG_{good} = O(\frac{k_\mu (k_\mu + R_{max})}{c_\mu} d \frac{NT}{\sqrt{B}} \log(NT))$.

Using the same argument as in the proof for FedGLB-UCB$_1$, the cumulative regret for each 'bad ' epoch is upper bounded by $2 k_\mu S \frac{NT}{B}$. Since there can be at most $O(d \log NT)$ 'bad epochs', the cumulative regret for all the 'bad epochs' is upper bounded by

$$REG_{bad} = O(dNT \log(NT) \cdot \frac{k_\mu S}{B})$$

Combining the regret incurred in both 'good' and 'bad' epochs, we have

$$R_T = O\big(\frac{k_\mu (k_\mu + R_{max})}{c_\mu} d \frac{NT}{\sqrt{B}} \log(NT) + dNT \log(NT) \cdot \frac{k_\mu S}{B}\big)$$

To recover the regret in centralized setting, we can $B = NT$, which leads to $R_T = O(\frac{k_\mu (k_\mu + R_{max})}{c_\mu} d \sqrt{NT} \log(NT))$. However, this incurs communication cost $C_T = N^2 T$. Alternatively, if we set $B = \sqrt{NT}$, we have $R_T = O(\frac{k_\mu (k_\mu + R_{max})}{c_\mu} d N^{3/4} T^{3/4} \log(NT))$, and $C_T = O(N^{1.5} \sqrt{T})$. □

F Additional Explanation about Figure 2

In Section 5, we used the scatter plots to present the experiment results. Here we provide more explanation about how to interpret these figures. As mentioned earlier, each dot in Figure 2 denotes the cumulative communication cost (x-axis) and regret (y-axis) that an algorithm (FedGLB-UCB, its variants, or DisLinUCB) with certain threshold value of D or B (labeled next to the dot) has obtained at iteration T.

Here, Figure 3 shows how the cumulative regret/reward and communication cost of five algorithms change over the course of federated bandit learning in our evaluations on synthetic dataset, (their final results at iteration T are used to plot five dots in Figure 2(a)). By carefully examining the relationship between their regret and communication cost, we can see that in Figure 3, FedGLB-UCB ($D = 5.0$), FedGLB-UCB$_1$ ($B = 10.0$), FedGLB-UCB$_2$ ($B = 10.0$), and FedGLB-UCB$_3$ ($B = 5000.0$) incur similar total communication cost, but FedGLB-UCB ($D = 5.0$) attains much smaller regret than the others. Meanwhile, FedGLB-UCB ($D = 5000.0$) attains almost the same regret as FedGLB-UCB$_2$ ($B = 10.0$), but its communication cost is much lower.

Figure 3 also depicts how the communication was controlled in FedGLB-UCB under its event triggered protocol (e.g., generally a decreasing frequency of communication comparing to the scheduled updated in its variants). This shows that FedGLB-UCB strikes the best regret/reward-communication trade-off among the algorithm instances in comparison. However, this line chart can only accommodate a limited range of trade-off settings for these algorithms, to attain a reasonable visibility. In comparison, the scatter plots in Figure 2(a) provide a much more thorough view of how well the algorithms balance regret/reward and communication cost, by covering a large range of trade-off settings.

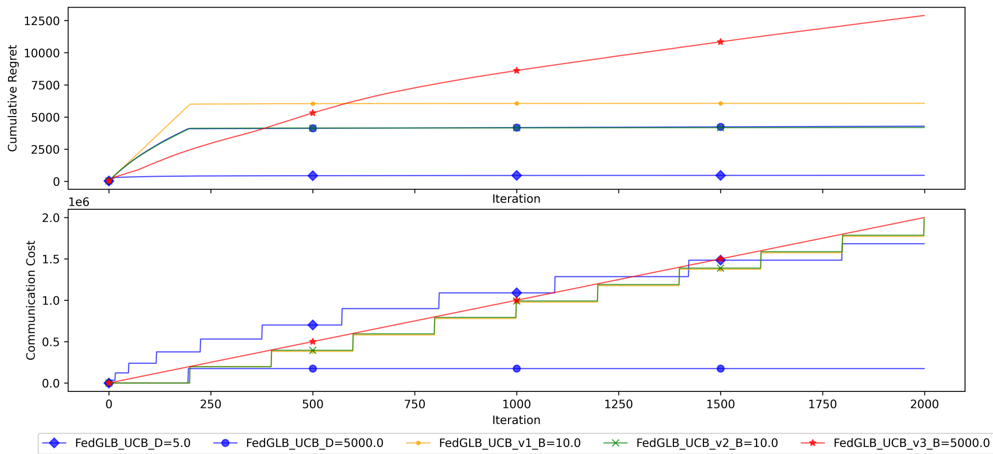

Figure 3: Experiment results showing regret and communication cost over time.