A. PROOF OF THEOREM 3

For notational simplicity, we define

$$\mathbf{G}^n_{m,t-\tau_{n,m}} = \nabla f_{m,t-\tau_{n,m}}(\mathbf{X}_{n,t-\tau_{n,m}}), \tag{A-1}$$

$$K_n = \max_{m\in\mathcal{W}_n} w_{n,m}L_m, \tag{A-2}$$

$$\tau_{n,\min} = \min_{m\in\mathcal{W}_n} \tau_{n,m}, \tag{A-3}$$

$$\tau_{n,\max} = \max_{m\in\mathcal{W}_n} \tau_{n,m}, \tag{A-4}$$

$$\Delta\tau_n = \tau_{n,\max} - \tau_{n,\min}, \tag{A-5}$$

and denote by $|\mathcal{W}_n|$ the cardinality of $\mathcal{W}_n$.

Recall that the network regret can be written as the sum of local regrets,

$$\mathrm{Reg}(T) = \sum_{n\in\mathcal{N}} \Big(\underbrace{\mathcal{L}^T_n\left(\mathbf{X}^T_n\right) - \mathcal{L}^T_n\left(\mathbf{U}^*_n\right)}_{\mathrm{Reg}_n(T)} \Big) = \sum_{n\in\mathcal{N}} \mathrm{Reg}_n(T). \tag{A-6}$$

We now turn to analyze the regret at agent n. Aggregating the results across all agents then yields the regret bound $\mathrm{Reg}(T)$. For agent n, by applying the convexity property [12, Lemma 2.5], the regret at agent n can be decomposed as

$$\begin{aligned}
\mathrm{Reg}_n(T) &= \sum_{t=1}^{T}\sum_{m\in\mathcal{W}_n} w_{n,m} f_{m,t}(\mathbf{X}_{n,t}) - \sum_{t=1}^{T}\sum_{m\in\mathcal{W}_n} w_{n,m} f_{m,t}(\mathbf{U}^*_n)\\
&\le \sum_{t=1}^{T}\left(\sum_{m\in\mathcal{W}_n}\left\langle w_{n,m}\mathbf{G}^n_{m,t}, \mathbf{X}_{n,t}-\mathbf{U}^*_n\right\rangle\right)\\
&\le \underbrace{\sum_{m\in\mathcal{W}_n}\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}^*_n\right\rangle}_{\mathrm{Reg}^*_n(T)}\\
&\quad + \underbrace{\sum_{m\in\mathcal{W}_n}\sum_{t=T+\tau_{n,m}+1}^{T+\tau_{n,\max}} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}, \mathbf{U}^*_n - \mathbf{X}_{n,t-\tau_{n,m}}\right\rangle}_{\mathrm{Drift}_n(T)}.
\end{aligned} \tag{A-7}$$

Next, we analyze the terms $\mathrm{Reg}^*_n(T)$ and $\mathrm{Drift}_n(T)$ separately.

A.1. Bounding $\mathrm{Reg}^*_n(T)$

To analyze $\mathrm{Reg}^*_n(T)$, we first present following inequity

$$\begin{aligned}
&\frac{1}{2}\left\|\mathbf{X}_{n,t+1}-\mathbf{U}^*_n\right\|^2 - \frac{1}{2}\left\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\right\|^2\\
&= \frac{1}{2}\left\|\Pi_{\mathcal{X}}\left[\mathbf{X}_{n,t} - \eta_{n,t}\sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right] - \mathbf{U}^*_n\right\|^2 - \frac{1}{2}\left\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\right\|^2\\
&\overset{\text{(a)}}{\le} \frac{1}{2}\left\|\mathbf{X}_{n,t} - \eta_{n,t}\sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}} - \mathbf{U}^*_n\right\|^2 - \frac{1}{2}\left\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\right\|^2\\
&= \frac{\eta^2_{n,t}}{2}\left\|\sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right\|^2 - \eta_{n,t}\sum_{m\in\mathcal{W}_n}\left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t}-\mathbf{U}^*_n\right\rangle\\
&= \frac{\eta^2_{n,t}}{2}\left\|\sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right\|^2 - \eta_{n,t}\sum_{m\in\mathcal{W}_n}\left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t-\tau_{n,m}}-\mathbf{U}^*_n\right\rangle\\
&\quad - \eta_{n,t}\sum_{m\in\mathcal{W}_n}\left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t}-\mathbf{X}_{n,t-\tau_{n,m}}\right\rangle
\end{aligned} \tag{A-8}$$

where (a) holds since $\left\|\Pi_{\mathcal{X}}[\mathbf{X}]-\mathbf{Y}\right\|^2 \le \left\|\mathbf{X}-\mathbf{Y}\right\|^2$ for $\mathbf{Y}\in\mathcal{X}$ [11, Proposition 2.11].

Dividing both sides by $\eta_{n,t}$ and rearranging the terms, we obtain

$$\begin{aligned}
&\sum_{m\in\mathcal{W}_n} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}^*_n \right\rangle \\
&\le \frac{\eta_{n,t}}{2}\left\| \sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right\|^2 + \frac{\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}^*_n\|^2}{2\eta_{n,t}} \\
&\quad - \sum_{m\in\mathcal{W}_n} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t} - \mathbf{X}_{n,t-\tau_{n,m}} \right\rangle.
\end{aligned} \tag{A-9}$$

Using Cauchy–Schwarz inequality to upper-bound the last inner product in (A-9), we further have

$$\sum_{m\in\mathcal{W}_n} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}^*_n \right\rangle \tag{A-10a}$$

$$\begin{aligned}
&\le \frac{\eta_{n,t}}{2}\left\| \sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right\|^2 + \frac{\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}^*_n\|^2}{2\eta_{n,t}} \\
&\quad + \sum_{m\in\mathcal{W}_n} w_{n,m}\left\|\mathbf{G}^n_{m,t-\tau_{n,m}}\right\| \left\|\mathbf{X}_{n,t} - \mathbf{X}_{n,t-\tau_{n,m}}\right\| \mathbb{1}_{\{t>\tau_{n,m}\}},
\end{aligned} \tag{A-10b}$$

$$\begin{aligned}
&\overset{(a)}{\le} \frac{\eta_{n,t}}{2}\left\| \sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}\right\|^2 + \frac{\|\mathbf{X}_{n,t}-\mathbf{U}^*_n\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}^*_n\|^2}{2\eta_{n,t}} \\
&\quad + \sum_{m\in\mathcal{W}_n} w_{n,m}L_m \left\|\mathbf{X}_{n,t} - \mathbf{X}_{n,t-\tau_{n,m}}\right\| \mathbb{1}_{\{t>\tau_{n,m}\}}.
\end{aligned} \tag{A-10c}$$

where (a) follows from the bounded gradient assumption.

Unrolling the difference $\mathbf{X}_{n,t} - \mathbf{X}_{n,t-\tau_{n,m}}$ for $t > \tau_{n,m}$ and from the update rule (7), it holds

$$\begin{aligned}
\left\|\mathbf{X}_{n,t} - \mathbf{X}_{n,t-\tau_{n,m}}\right\| &= \left\| \sum_{j=1}^{\tau_{n,m}} (\mathbf{X}_{n,t-j+1} - \mathbf{X}_{n,t-j})\right\| \overset{(a)}{\le} \sum_{j=1}^{\tau_{n,m}} \|\mathbf{X}_{n,t-j+1} - \mathbf{X}_{n,t-j}\| \\
&\overset{(b)}{\le} \sum_{j=1}^{\tau_{n,m}} \eta_{n,t-j} \left\| \sum_{s\in\mathcal{W}_n} w_{n,s}\mathbf{G}^n_{s,t-j-\tau_{n,s}}\mathbb{1}_{\{t-j\ge\tau_{n,s}\}}\right\| \\
&\overset{(c)}{\le} \sum_{j=1}^{\tau_{n,m}} \eta_{n,t-j} \sum_{s\in\mathcal{W}_n} w_{n,s}\left\|\mathbf{G}^n_{s,t-j-\tau_{n,s}}\right\|\mathbb{1}_{\{t-j\ge\tau_{n,s}\}} \\
&\overset{(d)}{\le} \sum_{s\in\mathcal{W}_n} \sum_{j=1}^{\min\{\tau_{n,m},t-\tau_{n,s}-1\}} \eta_{n,t-j} w_{n,s} L_s.
\end{aligned} \tag{A-11}$$

where (a) and (c) hold by the triangle inequality, (b) follows from $\|\Pi_{\mathcal{X}}[\mathbf{X}] - \mathbf{Y}\|^2 \le \|\mathbf{X}-\mathbf{Y}\|^2$ for $\mathbf{Y}\in\mathcal{X}$ [11, Proposition 2.11], and (d) holds by the bounded gradient assumption.

By plugging (A-11) into the last term of (A-10c), and summing over $t = \tau_{n,\min}+1$ to $T+\tau_{n,\max}$, we obtain the following upper bound on $\mathrm{Reg}^*_n(T)$

$$\begin{aligned}
\mathrm{Reg}^*_n(T) &\le \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \sum_{m\in\mathcal{W}_n} \left\langle w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\mathbb{1}_{\{t>\tau_{n,m}\}}, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}^*_n \right\rangle \\
&\le \underbrace{\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\eta_{n,t}}{2}\left\| \sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\right\|^2}_{\text{Part 1}} \\
&\quad + \underbrace{\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \sum_{m\in\mathcal{W}_n}\sum_{s\in\mathcal{W}_n} \sum_{j=1}^{\min\{\tau_{n,m},t-\tau_{n,s}-1\}} \eta_{n,t-j} w_{n,m} w_{n,s} L_m L_s \mathbb{1}_{\{t>\tau_{n,m}\}}}_{\text{Part 2}}
\end{aligned}$$

$$+\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\|\mathbf{X}_{n,t}-\mathbf{U}_n^*\|^2-\|\mathbf{X}_{n,t+1}-\mathbf{U}_n^*\|^2}{2\eta_{n,t}}. \tag{A-12}$$

Next, we bound Part 1 and Part 2. Specifically, for Part 1, we have

$$\begin{aligned}\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\eta_{n,t}}{2}\left\|\sum_{m\in\mathcal{W}_n} w_{n,m}\mathbf{G}^n_{m,t-\tau_{n,m}}\right\|^2 &\overset{\text{(a)}}{\leq} \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\eta_{n,t}}{2}\sum_{m\in\mathcal{W}_n} w_{n,m}\left\|\mathbf{G}^n_{m,t-\tau_{n,m}}\right\|^2\\ &\overset{\text{(b)}}{\leq}\left(\sum_{m\in\mathcal{W}_n} w_{n,m}L_m^2\right)\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}}\frac{\eta_{n,t}}{2}\\ &\overset{\text{(c)}}{\leq} K_n\left(\sum_{m\in\mathcal{W}_n} L_m\right)\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}}\frac{\eta_{n,t}}{2},\end{aligned} \tag{A-13}$$

where (a) follows from Jensen's inequality, (b) from the bounded gradient assumption, and (c) from the definition of $K_n = \max_{m\in\mathcal{W}_n} w_{n,m}L_m$.

For the Part 2, we have

$$\begin{aligned}&\sum_{s\in\mathcal{W}_n}\sum_{j=1}^{\min\{\tau_{n,m},t-\tau_{n,s}-1\}} \eta_{n,t-j}w_{n,m}w_{n,s}L_mL_s\mathbb{1}_{\{t>\tau_{n,m}\}}\\ &\overset{\text{(a)}}{\leq}\eta_{n,t-1}\sum_{s\in\mathcal{W}_n} w_{n,m}w_{n,s}L_mL_s\min\{\tau_{n,m},t-\tau_{n,s}-1\}\mathbb{1}_{\{t>\tau_{n,m}\}}\\ &\overset{\text{(b)}}{\leq}K_n^2\eta_{n,t-1}\sum_{s\in\mathcal{W}_n}\min\{\tau_{n,m},t-\tau_{n,s}-1\}\mathbb{1}_{\{t>\tau_{n,m}\}},\end{aligned} \tag{A-14}$$

where (a) follows from $\sum_{j=1}^{\min\{a,b\}} c = \min\{a,b\}\times c$, with c being a constant independent of j, and (b) from the definition of $K_n = \max_{m\in\mathcal{W}_n} w_{n,m}L_m$.

Under the non-increasing learning rate assumption, it follows

$$\begin{aligned}&\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}}\sum_{m\in\mathcal{W}_n}K_n^2\eta_{n,t-1}\sum_{s\in\mathcal{W}_n}\min\{\tau_{n,m},t-\tau_{n,s}-1\}\mathbb{1}_{\{t>\tau_{n,m}\}}\\ &=\sum_{m\in\mathcal{W}_n}\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,\min}+1}^{\tau_{n,s}+\tau_{n,m}}K_n^2\eta_{n,t-1}(t-\tau_{n,s}-1)\mathbb{1}_{\{t>\tau_{n,m}\}}+\sum_{m\in\mathcal{W}_n}\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,s}+\tau_{n,m}+1}^{T+\tau_{n,\max}}K_n^2\eta_{n,t-1}\tau_{n,m}\\ &=\sum_{m\in\mathcal{W}_n}\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,m}+1}^{\tau_{n,s}+\tau_{n,m}}K_n^2\eta_{n,t-1}(t-\tau_{n,s}-1)+\sum_{m\in\mathcal{W}_n}\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,s}+\tau_{n,m}+1}^{T+\tau_{n,\max}}K_n^2\eta_{n,t-1}\tau_{n,m}\\ &=\sum_{m\in\mathcal{W}_n}\eta_{n,\tau_{n,m}}K_n^2\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,m}+1}^{\tau_{n,s}+\tau_{n,m}}(t-\tau_{n,s}-1)+\sum_{m\in\mathcal{W}_n}\tau_{n,m}K_n^2\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,s}+\tau_{n,m}+1}^{T+\tau_{n,\max}}\eta_{n,t-1}\\ &\overset{\text{(a)}}{\leq}|\mathcal{W}_n|^2K_n^2\tau_{n,\max}^2\eta_{n,\tau_{n,\min}+1}+|\mathcal{W}_n|K_n^2\sum_{m\in\mathcal{W}_n}\tau_{n,m}\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}}\eta_{n,t}\end{aligned} \tag{A-15}$$

where (a) follows from

$$\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,m}+1}^{\tau_{n,s}+\tau_{n,m}}(t-\tau_{n,s}-1)=\sum_{s\in\mathcal{W}_n}\frac{2\tau_{n,m}-\tau_{n,s}-1}{2}\tau_{n,s}\leq|\mathcal{W}_n|\tau_{n,\max}^2, \tag{A-16}$$

and

$$\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,s}+\tau_{n,m}+1}^{T+\tau_{n,\max}}\eta_{n,t-1}=\sum_{s\in\mathcal{W}_n}\sum_{t=\tau_{n,m}+\tau_{n,s}}^{T+\tau_{n,\max}-1}\eta_{n,t}\leq|\mathcal{W}_n|\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}}\eta_{n,t}. \tag{A-17}$$

Combining the above equations, the term $\mathrm{Reg}_n^*(T)$ can be bounded as

$$\begin{aligned}\mathrm{Reg}_n^*(T) \leq &\left(\frac{K_n}{2}\sum_{m\in\mathcal{W}_n} L_m + |\mathcal{W}_n| K_n^2 \sum_{m\in\mathcal{W}_n}\tau_{n,m}\right)\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \eta_{n,t} \\ &+ |\mathcal{W}_n|^2 K_n^2 \tau_{n,\max}^2 \eta_{n,\tau_{n,\min}+1} + \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\|\mathbf{X}_{n,t}-\mathbf{U}_n^*\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}_n^*\|^2}{2\eta_{n,t}}.\end{aligned} \tag{A-18}$$

A.2. Bounding $\mathrm{Drift}_n(T)$

For the term $\mathrm{Drift}_n(T)$, we have

$$-\sum_{m\in\mathcal{W}_n}\sum_{t=T+\tau_{n,m}+1}^{T+\tau_{n,\max}} \left\langle w_{n,m}\mathbf{G}_{m,t-\tau_{n,m}}^n, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}_n^* \right\rangle \tag{A-19a}$$

$$\overset{(a)}{\leq} \sum_{m\in\mathcal{W}_n}\sum_{t=T+\tau_{n,m}+1}^{T+\tau_{n,\max}} \left|\left\langle w_{n,m}\mathbf{G}_{m,t-\tau_{n,m}}^n, \mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}_n^* \right\rangle\right| \tag{A-19b}$$

$$\overset{(b)}{\leq} \sum_{m\in\mathcal{W}_n}\sum_{t=T+\tau_{n,m}+1}^{T+\tau_{n,\max}} \left(\left\|w_{n,m}\mathbf{G}_{m,t-\tau_{n,m}}^n\right\|^2 + \left\|\mathbf{X}_{n,t-\tau_{n,m}} - \mathbf{U}_n^*\right\|^2\right) \tag{A-19c}$$

$$\leq \frac{\Delta\tau_n}{2}\left(K_n\Big(\sum_{m\in\mathcal{W}_n} L_m\Big) + |\mathcal{W}_n| B^2\right) \tag{A-19d}$$

where (a) follows from the inequality $-\langle a,b\rangle \leq |\langle a,b\rangle|$, and (b) from Young's inequality $2\,|\langle a,b\rangle| \leq |a|^2 + |b|^2$.

A.3. Regret Bound

Combining the above equations, the regret of user n can be bounded as

$$\begin{aligned}\mathrm{Reg}_n^*(T) \leq &\left(\frac{K_n}{2}\sum_{m\in\mathcal{W}_n} L_m + |\mathcal{W}_n| K_n^2 \sum_{m\in\mathcal{W}_n}\tau_{n,m}\right)\sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \eta_{n,t} \\ &+ |\mathcal{W}_n|^2 K_n^2 \tau_{n,\max}^2 \eta_{n,\tau_{n,\min}+1} + \frac{\Delta\tau_n}{2}\left(K_n\Big(\sum_{m\in\mathcal{W}_n} L_m\Big) + |\mathcal{W}_n| B^2\right) \\ &+ \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\|\mathbf{X}_{n,t}-\mathbf{U}_n^*\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}_n^*\|^2}{2\eta_{n,t}}.\end{aligned} \tag{A-20}$$

Defining

$$Q_n = \frac{K_n}{2}\sum_{m\in\mathcal{W}_n} L_m + |\mathcal{W}_n| K_n^2 \sum_{m\in\mathcal{W}_n}\tau_{n,m}, \tag{A-21}$$

$$P_n = |\mathcal{W}_n|^2 K_n^2 \tau_{n,\max}^2, \tag{A-22}$$

$$H_n = \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\|\mathbf{X}_{n,t}-\mathbf{U}_n^*\|^2 - \|\mathbf{X}_{n,t+1}-\mathbf{U}_n^*\|^2}{2\eta_{n,t}}, \tag{A-23}$$

$$C_n = \frac{\Delta\tau_n}{2}\left(K_n\Big(\sum_{m\in\mathcal{W}_n} L_m\Big) + |\mathcal{W}_n| B^2\right), \tag{A-24}$$

and summing over all agents, the network regret can be bounded as

$$\mathrm{Reg}(T) \leq \sum_{n\in\mathcal{W}_n}\left(Q_n \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \eta_{n,t} + P_n \eta_{n,\tau_{n,\min}+1} + H_n + C_n\right). \tag{A-25}$$

B. PROOF OF COROLLARY 1

We first assume that a learning rate of the form $\eta_{n,t} = c(t - \tau_{n,\min})^{-\beta}$ with $\beta \in (0, 1)$. For this choice, we will establish a bound on the regret, and then verify that a sublinear regret can be achieved only for $\beta = 0.5$.

For the first term in the regret bound (A-25), we have

$$Q_n \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \eta_{n,t} = Q_n c \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} (t - \tau_{n,\min})^{-\beta}. \tag{B-1}$$

Substituting $s = t - \tau_{n,\min}$, we get

$$Q_n c \sum_{s=1}^{T+\Delta\tau_n} s^{-\beta} \le Q_n c \int_1^{T+\Delta\tau_n} s^{-\beta} ds = \frac{Q_n c}{1-\beta}(T + \Delta\tau_n)^{1-\beta}. \tag{B-2}$$

For the second term in the regret bound (A-25), we have

$$P_n \eta_{n,\tau_{n,\min}+1} = P_n c \times 1^{-\beta} = P_n c, \tag{B-3}$$

which is constant with respect to T.

For the third term in the regret bound (A-25), we have

$$\begin{aligned} H_n &= \sum_{t=\tau_{n,\min}+1}^{T+\tau_{n,\max}} \frac{\|\mathbf{X}_{n,t} - \mathbf{U}_n^*\|^2 - \|\mathbf{X}_{n,t+1} - \mathbf{U}_n^*\|^2}{2\eta_{n,t}} \\ &\le \frac{\left\|\mathbf{X}_{n,\tau_{n,\min}+1} - \mathbf{U}_n^*\right\|^2}{2\eta_{n,\tau_{n,\min}+1}} - \frac{\left\|\mathbf{X}_{n,T+\tau_{n,\max}+1} - \mathbf{U}_n^*\right\|^2}{2\eta_{n,T+\tau_{n,\max}+1}} \\ &\le \frac{B^2}{2\eta_{n,\tau_{n,\min}+1}} = \frac{B^2}{2T^{-\beta}c} = \frac{B^2}{2c}T^\beta \end{aligned} \tag{B-4}$$

Overall, combining the three terms, we obtain

$$\mathrm{Reg}(T) \le \sum_{n\in\mathcal{N}} \left(\frac{c}{1-\beta} Q_n (T + \Delta\tau_n)^{1-\beta} + \frac{B^2}{2c} T^\beta + P_n c + C_n \right) = \mathcal{O}\left((T + \Delta\tau_n)^{1-\beta} + T^\beta \right). \tag{B-5}$$

From (B-5), it follows that $\mathrm{Reg}(T) = \mathcal{O}\big(T^{\max\{\beta,\,1-\beta\}}\big)$, hence the regret is sublinear for all $\beta \in (0, 1)$. Choosing $\beta = \frac{1}{2}$ balances the two terms and yields

$$\mathrm{Reg}(T) \le \sum_{n\in\mathcal{N}} \left(2cQ_n \sqrt{T + \Delta\tau_n} + \frac{B^2}{2c}\sqrt{T} + P_n c + C_n \right) = \mathcal{O}\left(\sqrt{T + \Delta\tau_n} + \sqrt{T} \right). \tag{B-6}$$