Theoretical Convergence Guaranteed Resource-Adaptive Federated Learning with Mixed Heterogeneity

ABSTRACT

In this paper, we propose an adaptive learning paradigm for resource-constrained cross-device federated learning, in which heterogeneous local submodels with varying resources can be jointly trained to produce a global model. Different from existing studies, the submodel structures of different clients are formed by arbitrarily assigned neurons according to their local resources. Along this line, we first design a general resource-adaptive federated learning algorithm, namely *RA-Fed*, and rigorously prove its convergence with asymptotically optimal rate $O(1/\sqrt{\Gamma^* TQ})$ under loose assumptions. Furthermore, to address both *submodels heterogeneity* and *data heterogeneity* challenges under *non-uniform training*, we come up with a new server aggregation mechanism *RAM-Fed* with the same theoretically proved convergence rate. Moreover, we shed light on several key factors impacting convergence, such as minimum coverage rate, data heterogeneity level, submodel induced noises. Finally, we conduct extensive experiments on two types of tasks with three widely used datasets under different experimental settings. Compared with the state-of-the-arts, our methods improve the accuracy up to 10% on average. Particularly, when submodels jointly train with 50% parameters, *RAM-Fed* achieves comparable accuracy to *FedAvg* trained with the full model.

CCS CONCEPTS

• Computing methodologies → Machine learning; **• Computer systems organization → Distributed architectures**.

KEYWORDS

Federated learning, Limited resources, Heterogeneity, Convergence analysis

ACM Reference Format:

. 2023. Theoretical Convergence Guaranteed Resource-Adaptive Federated Learning with Mixed Heterogeneity. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), August 6–10, 2023, Long Beach, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3580305.3599521

KDD '23, August 6–10, 2023, Long Beach, CA, USA

ACM ISBN 979-8-4007-0103-0/23/08…$15.00

https://doi.org/10.1145/3580305.3599521

A SUPPLEMENT

A.1 Part One

Let us start the proof of RA-Fed from L-Lipschitzian Condition:

$$\mathbb{E}[F(\theta_{q+1})] - \mathbb{E}[F(\theta_q)] \leq \underbrace{\mathbb{E}[< \nabla F(\theta_q), \theta_{q+1} - \theta_q >]}_{U_1} + \frac{L}{2} \underbrace{\mathbb{E}[\|\theta_{q+1} - \theta_q\|^2]}_{U_2}$$

bound U_1:

$$\begin{aligned}
&\mathbb{E} < \nabla F(\theta_q), \theta_{q+1} - \theta_q > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), \theta^i_{q+1} - \theta^i_q > + \sum_{i \in K - S_q} \mathbb{E}[< \nabla F^i(\theta_q), \theta^i_{q+1} - \theta^i_q > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), \theta^i_{q+1} - \theta^i_q > + \sum_{i \in K - S_q} \mathbb{E}[< \nabla F^i(\theta_q), \mathbf{0} > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), \theta^i_{q+1} - \theta^i_q > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), -\frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} (\theta_{q,n,0} - \theta_{q,n,T})^i > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), \\
&- \frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} (\theta_{q,n,0} - (\theta_{q,n,0} - \sum_{t=1}^{T} \gamma \nabla F_n(\theta_{q,n,t-1}, \xi_{n,t-1}) \odot m_{n,q}))^i > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), -\frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} \sum_{t=1}^{T} \gamma \nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1}) > \\
&= \sum_{i \in S_q} \mathbb{E} < \nabla F^i(\theta_q), -\frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} \sum_{t=1}^{T} \gamma \nabla F^i_n(\theta_{q,n,t-1}) > \\
&= \sum_{i \in S_q} \mathbb{E}[< \nabla F^i(\theta_q), -\frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} \sum_{t=1}^{T} \gamma [\nabla F^i_n(\theta_{q,n,t-1}) - \nabla F^i(\theta_q) + \nabla F^i(\theta_q)] > \\
&= \underbrace{- \sum_{i \in S_q} T\gamma \mathbb{E} < \nabla F^i(\theta_q), \nabla F^i(\theta_q) >}_{U_3} \\
&+ \underbrace{\sum_{i \in S_q} \mathbb{E}[< \nabla F^i(\theta_q), -\frac{1}{\Gamma^i_q} \sum_{n \in N^i_q} \sum_{t=1}^{T} \gamma [\nabla F^i_n(\theta_{q,n,t-1}) - \nabla F^i(\theta_q)] >}_{U_4}
\end{aligned}$$

bound U_3:

$$- \sum_{i \in S_q} T\gamma \mathbb{E} < \nabla F^i(\theta_q), \nabla F^i(\theta_q) >= - \sum_{i \in S_q} T\gamma \mathbb{E} \|\nabla F^i(\theta_q)\|^2$$

bound U_4:

$$\sum_{i\in S_q}\mathbb{E}[<\nabla F^i(\theta_q),-\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}\gamma[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F^i(\theta_q)]>$$
$$=\sum_{i\in S_q}T\gamma\mathbb{E}[<\nabla F^i(\theta_q),-\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F^i(\theta_q)]>$$
$$\leq\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+\underbrace{\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)+\nabla F_n^i(\theta_q)-\nabla F^i(\theta_q)]\|^2}_{U_5}$$

bound U_5:

$$\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)+\nabla F_n^i(\theta_q)-\nabla F^i(\theta_q)]\|^2$$
$$\leq\underbrace{T\gamma\sum_{i\in S_q}\mathbb{E}\|\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)]\|^2}_{U_6}$$
$$+\underbrace{T\gamma\sum_{i\in S_q}\mathbb{E}\|\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_q)-\nabla F^i(\theta_q)]\|^2}_{U_7}$$

bound U_6:

$$T\gamma\sum_{i\in S_q}\mathbb{E}\|\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)]\|^2$$
$$\leq T\gamma\sum_{i\in S_q}\frac{1}{T\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}\mathbb{E}\|[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)]\|^2$$
$$\leq T\gamma\frac{1}{T\Gamma^*}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|[\nabla F_n^i(\theta_{q,n,t-1})-\nabla F_n^i(\theta_q)]\|^2$$
$$\leq T\gamma\frac{1}{T\Gamma^*}\sum_{n=1}^{N}\sum_{t=1}^{T}\mathbb{E}\|[\nabla F_n(\theta_{q,n,t-1})-\nabla F_n(\theta_q)]\|^2$$
$$\leq T\gamma\frac{1}{\Gamma^*}\sum_{n=1}^{N}L^2\underbrace{\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1}-\theta_q]\|^2}_{U_8}$$

bound U_8:

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1}-\theta_q\|^2$$
$$\leq\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1}-\theta_{q,n,0}\|^2+\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,0}-\theta_q\|^2$$
$$=\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\|\sum_{j=0}^{t-2}-\gamma\nabla F_n(\theta_{q,n,j},\xi_{n,j})\odot m_{q,n}\|^2+\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\|\mathsf{C}(\theta_q)\odot m_{n,q}-\theta_q\|^2$$
$$=\frac{2\gamma^2}{T}\sum_{t=1}^{T}\mathbb{E}\|\sum_{j=0}^{t-2}(\nabla F_n(\theta_{q,n,j},\xi_{n,j})-\nabla F_n(\theta_{q,n,j})+\nabla F_n(\theta_{q,n,j}))\odot m_{q,n}\|^2$$
$$+\frac{2}{T}\sum_{t=1}^{T}\mathbb{E}\|\mathsf{C}(\theta_q)\odot m_{n,q}-\mathsf{C}(\theta_q)+\mathsf{C}(\theta_q)-\theta_q\|^2$$
$$\leq\frac{4\gamma^2}{T}\sum_{t=1}^{T}\mathbb{E}\|\sum_{j=0}^{t-2}(\nabla F_n(\theta_{q,n,j},\xi_{n,j})-\nabla F_n(\theta_{q,n,j}))\odot m_{q,n}\|^2$$
$$+\frac{4\gamma^2}{T}\sum_{t=1}^{T}\mathbb{E}\|\sum_{j=0}^{t-2}\nabla F_n(\theta_{q,n,j})\odot m_{q,n}\|^2$$
$$+\frac{4}{T}\sum_{t=1}^{T}\mathbb{E}\|\mathsf{C}(\theta_q)\odot m_{n,q}-\mathsf{C}(\theta_q)\|^2+\frac{4}{T}\sum_{t=1}^{T}\mathbb{E}\|\mathsf{C}(\theta_q)-\theta_q\|^2$$
$$\leq\frac{4\gamma^2}{T}\sum_{t=1}^{T}(t-1)\sigma^2+\frac{4}{T}\sum_{t=1}^{T}w_1^2\mathbb{E}\|\mathsf{C}(\theta_q)\|^2+\frac{4}{T}\sum_{t=1}^{T}w_2^2\mathbb{E}\|\theta_q\|^2$$
$$+\frac{4\gamma^2}{T}\sum_{t=1}^{T}\mathbb{E}\|\sum_{j=0}^{t-2}(\nabla F_n(\theta_{q,n,j})-\nabla F_n(\theta_q)+\nabla F_n(\theta_q))\odot m_{q,n}\|^2$$
$$\leq 2\gamma^2T\sigma^2+\frac{8\gamma^2}{T}\sum_{t=1}^{T}(t-1)\sum_{j=0}^{t-2}\mathbb{E}\|(\nabla F_n(\theta_{q,n,j})-\nabla F_n(\theta_q))\odot m_{q,n}\|^2$$
$$+\frac{8\gamma^2}{T}\sum_{t=1}^{T}(t-1)\sum_{j=0}^{t-2}\mathbb{E}\|\nabla F_n(\theta_q)\odot m_{q,n}\|^2$$
$$+\frac{4}{T}\sum_{t=1}^{T}w_1^2\mathbb{E}\|\mathsf{C}(\theta_q)-\theta_q+\theta_q\|^2+\frac{4}{T}\sum_{t=1}^{T}w_2^2\mathbb{E}\|\theta_q\|^2$$
$$\leq 2\gamma^2T\sigma^2+\frac{8\gamma^2L^2}{T}\sum_{t=1}^{T}(t-1)\sum_{j=0}^{t-2}\mathbb{E}\|\theta_{q,n,j}-\theta_q\|^2$$
$$+8\gamma^2T^2\mathbb{E}\|(\nabla F_n(\theta_q)-\nabla F(\theta_q)+\nabla F(\theta_q))\odot m_{q,n}\|^2$$
$$+\frac{8}{T}\sum_{t=1}^{T}w_1^2\mathbb{E}\|\mathsf{C}(\theta_q)-\theta_q\|^2+\frac{8}{T}\sum_{t=1}^{T}w_1^2\mathbb{E}\|\theta_q\|^2+\frac{4}{T}\sum_{t=1}^{T}w_2^2\mathbb{E}\|\theta_q\|^2$$
$$\leq 2\gamma^2T\sigma^2+8\gamma^2L^2T(T-1)\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1}-\theta_q\|^2$$
$$+16\gamma^2T^2\mathbb{E}\|(\nabla F_n(\theta_q)-\nabla F(\theta_q))\odot m_{q,n}\|^2+16\gamma^2T^2\mathbb{E}\|\nabla F(\theta_q)\odot m_{q,n}\|^2$$
$$+\frac{8}{T}\sum_{t=1}^{T}w_1^2w_2^2\mathbb{E}\|\theta_q\|^2+\frac{8}{T}\sum_{t=1}^{T}w_1^2\mathbb{E}\|\theta_q\|^2+\frac{4}{T}\sum_{t=1}^{T}w_2^2\mathbb{E}\|\theta_q\|^2$$
$$\leq 2\gamma^2T\sigma^2+8\gamma^2L^2T^2\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_q-\theta_{q,n,t-1}\|^2$$
$$+16\gamma^2T^2\delta^2+16\gamma^2T^2\mathbb{E}\|\nabla F(\theta_q)\odot m_{q,n}\|^2$$
$$+4(2w_1^2w_2^2+2w_1^2+w_2^2)\mathbb{E}\|\theta_q\|^2$$

Letting $2w_1^2w_2^2+2w_1^2+w_2^2=w^2$, $8\gamma^2L^2T^2\leq\frac{1}{2}\Rightarrow\gamma\leq\frac{1}{4LT}$, we can get U_8:

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1}-\theta_q]\|^2\leq 4\gamma^2T\sigma^2+32\gamma^2T^2\delta^2$$
$$+32\gamma^2T^2\mathbb{E}\|\nabla F(\theta_q)\odot m_{q,n}\|^2+4w^2\mathbb{E}\|\theta_q\|^2\leq 4\gamma^2T\sigma^2+32\gamma^2T^2\delta^2$$
$$+32\gamma^2T^2\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2+4w^2\mathbb{E}\|\theta_q\|^2$$

bound U_7:

$$T\gamma \sum_{i\in S_q} \mathbb{E}\| \frac{1}{T\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} [\nabla F_n^i(\theta_q) - \nabla F^i(\theta_q)]\|^2$$
$$\leq T\gamma \frac{1}{T\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \sum_{i\in S_q} \mathbb{E}\|[\nabla F_n^i(\theta_q) - \nabla F^i(\theta_q)]\|^2$$
$$\leq T\gamma \frac{1}{T\Gamma^*} \sum_{n=1}^{N} \sum_{t=1}^{T} \mathbb{E}\|[\nabla F_n(\theta_q) - \nabla F(\theta_q)]\|^2$$
$$\leq T\gamma \frac{N}{\Gamma^*}\delta^2$$

Plugging U_8 into U_6, U_6 and U_7 into U_5, U_5 into U_4, we have:

$$\sum_{i\in S_q} \mathbb{E}[< \nabla F^i(\theta_q), -\frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \gamma[\nabla F_n^i(\theta_{q,n,t-1}) - \nabla F^i(\theta_q)] >$$
$$\leq \frac{T\gamma}{2} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2 \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ 8w^2T\gamma\frac{N}{\Gamma^*}L^2\mathbb{E}\|\theta_q\|^2 + 4\gamma^3T^2\frac{N}{\Gamma^*}L^2\sigma^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2\delta^2 + T\gamma\frac{N}{\Gamma^*}\delta^2$$

Plugging U_3, U_4 into U_1, we have:

$$\mathbb{E} < \nabla F(\theta_q), \theta_{q+1} - \theta_q > \leq -T\gamma \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ \frac{T\gamma}{2} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2 \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ 8w^2T\gamma\frac{N}{\Gamma^*}L^2\mathbb{E}\|\theta_q\|^2 + 4\gamma^3T^2\frac{N}{\Gamma^*}L^2\sigma^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2\delta^2 + T\gamma\frac{N}{\Gamma^*}\delta^2$$

bound U_2:

$$\frac{L}{2}\mathbb{E}\|\theta_{q+1} - \theta_q\|^2$$
$$= \frac{L}{2} \sum_{i\in S_q} \mathbb{E}\|\theta_{q+1}^i - \theta_q^i\|^2 + \frac{L}{2} \sum_{i\in K-S_q} \mathbb{E}\|\theta_{q+1}^i - \theta_q^i\|^2$$
$$= \frac{L}{2} \sum_{i\in S_q} \mathbb{E}\|\theta_{q+1}^i - \theta_q^i\|^2$$
$$= \frac{L}{2} \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} (\theta_{q,n,0} - \theta_{q,n,T})^i\|^2$$
$$= \frac{L}{2} \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} (\theta_{q,n,0}$$
$$- (\theta_{q,n,0} - \sum_{t=1}^{T} \gamma\nabla F_n(\theta_{q,n,t-1}, \xi_{n,t-1}) \odot m_{n,q}))^i\|^2$$
$$= \frac{L}{2} \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \gamma\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1})\|^2$$
$$\leq \frac{3}{2}L \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \gamma[\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1}) - \nabla F_n^i(\theta_{q,n,t-1})]\|^2$$
$$+ \frac{3}{2}L \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \gamma[\nabla F_n^i(\theta_{q,n,t-1}) - \nabla F^i(\theta_q)]\|^2$$
$$+ \frac{3}{2}L \sum_{i\in S_q} \mathbb{E}\| - \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} \sum_{t=1}^{T} \gamma\nabla F^i(\theta_q)\|^2$$
$$\leq \frac{3}{2}LT\gamma^2\frac{N}{\Gamma^*}\sigma^2 + 3L\gamma^2\frac{N}{\Gamma^*}L^2T^2(4\gamma^2T\sigma^2 + 32\gamma^2T^2\delta^2$$
$$+ 32\gamma^2T^2 \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 4w^2\mathbb{E}\|\theta_q\|^2)$$
$$+ 3L\frac{N}{\Gamma^*}\gamma^2T \sum_{t=1}^{T} \delta^2 + \frac{3}{2}L\gamma^2T^2 \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$= \frac{3}{2}LT\gamma^2\frac{N}{\Gamma^*}\sigma^2 + 12L^3\gamma^4\frac{N}{\Gamma^*}T^3\sigma^2 + 96L^3\gamma^4T^4\frac{N}{\Gamma^*}\delta^2$$
$$+ 96L^3\gamma^4T^4\frac{N}{\Gamma^*} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 12L^3\gamma^2T^2\frac{N}{\Gamma^*}w^2\mathbb{E}\|\theta_q\|^2$$
$$+ \frac{3}{2}L\gamma^2T^2 \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 3L\frac{N}{\Gamma^*}\gamma^2T^2\delta^2$$

Last,we have:

$$\mathbb{E}[F(\theta_{Q+1})] - \mathbb{E}[F(\theta_1)] = \sum_{q=1}^{Q} \mathbb{E}[F(\theta_{q+1})] - \sum_{q=1}^{Q} \mathbb{E}[F(\theta_q)]$$
$$\leq \sum_{q=1}^{Q} \mathbb{E}[< \nabla F(\theta_q), \theta_{q+1} - \theta_q >] + \sum_{q=1}^{Q} \frac{L}{2}\mathbb{E}\|\theta_{q+1} - \theta_q\|^2$$

Plugging U_1, U_2 into above equation, we have:

$$\mathbb{E}[F(\theta_{Q+1})] - \mathbb{E}[F(\theta_1)] \leq -T\gamma \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ \frac{T\gamma}{2} \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2 \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ 8w^2T\gamma\frac{N}{\Gamma^*}L^2 \sum_{q=1}^{Q} \mathbb{E}\|\theta_q\|^2 + 4\gamma^3T^2\frac{N}{\Gamma^*}L^2Q\sigma^2 + 32\gamma^3T^3\frac{N}{\Gamma^*}L^2Q\delta^2$$
$$+ T\gamma\frac{N}{\Gamma^*}Q\delta^2 + \frac{3}{2}LT\gamma^2\frac{N}{\Gamma^*}Q\sigma^2 + 12L^3\gamma^4\frac{N}{\Gamma^*}T^3Q\sigma^2 + 96L^3\gamma^4T^4\frac{N}{\Gamma^*}Q\delta^2$$
$$+ 96L^3\gamma^4T^4\frac{N}{\Gamma^*} \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 12L^3\gamma^2T^2\frac{N}{\Gamma^*}w^2 \sum_{q=1}^{Q} \mathbb{E}\|\theta_q\|^2$$
$$+ \frac{3}{2}L\gamma^2T^2 \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2 + 3L\frac{N}{\Gamma^*}\gamma^2T^2Q\delta^2$$
$$= -T\gamma(\frac{1}{2} - 32\gamma^2T^2\frac{N}{\Gamma^*}L^2 - 96L^3\gamma^3T^3\frac{N}{\Gamma^*} - \frac{3}{2}L\gamma T) \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ (8w^2T\gamma\frac{N}{\Gamma^*}L^2 + 12L^3\gamma^2T^2\frac{N}{\Gamma^*}w^2) \sum_{q=1}^{Q} \mathbb{E}\|\theta_q\|^2$$
$$+ T\gamma Q\frac{N}{\Gamma^*}(32\gamma^2T^2L^2 + 1 + 96L^3\gamma^3T^3 + 3L\gamma T)\delta^2$$
$$+ \gamma^2TLQ\frac{N}{\Gamma^*}(4\gamma TL + \frac{3}{2} + 12L^2\gamma^2T^2)\sigma^2$$
$$\overset{a}{\leq} -\frac{T\gamma}{8} \sum_{q=1}^{Q} \sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$+ (8w^2T\gamma\frac{N}{\Gamma^*}L^2 + 12L^3\gamma^2T^2\frac{N}{\Gamma^*}w^2) \sum_{q=1}^{Q} \mathbb{E}\|\theta_q\|^2$$

$$+ T\gamma Q\frac{N}{\Gamma^*}(32\gamma^2T^2L^2 + 1 + 96L^3\gamma^3T^3 + 3L\gamma T)\delta^2$$
$$+ \gamma^2 TLQ\frac{N}{\Gamma^*}(4\gamma TL + \frac{3}{2} + 12L^2\gamma^2T^2)\sigma^2$$

where a follows because:

$$32\gamma^2T^2\frac{N}{\Gamma^*}L^2 \le \frac{1}{8} \Rightarrow \gamma \le \frac{\sqrt{\Gamma^*}}{16TL\sqrt{N}}$$
$$96L^3\gamma^3T^3\frac{N}{\Gamma^*} \le \frac{1}{8} \Rightarrow \gamma \le \frac{(\Gamma^*)^{\frac{1}{3}}}{768^{\frac{1}{3}}LTN^{\frac{1}{3}}}$$
$$\frac{3}{2}L\gamma T \le \frac{1}{8} \Rightarrow \gamma \le \frac{1}{12TL}$$

Therefore, we have:

$$\frac{T\gamma}{8}\sum_{q=1}^{Q}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2 \le \mathbb{E}[F(\theta_1)] - \mathbb{E}[F(\theta_{Q+1})]$$
$$+ (8w^2T\gamma\frac{N}{\Gamma^*}L^2 + 12L^3\gamma^2T^2\frac{N}{\Gamma^*}w^2)\sum_{q=1}^{Q}\mathbb{E}\|\theta_q\|^2$$
$$+ T\gamma Q\frac{N}{\Gamma^*}(32\gamma^2T^2L^2 + 1 + 96L^3\gamma^3T^3 + 3L\gamma T)\delta^2$$
$$+ \gamma^2TLQ\frac{N}{\Gamma^*}(4\gamma TL + \frac{3}{2} + 12L^2\gamma^2T^2)\sigma^2$$

dividing both sides by Q and $\frac{T\gamma}{8}$:

$$\frac{1}{Q}\sum_{q=1}^{Q}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2 \le \frac{8\mathbb{E}[F(\theta_1)]}{T\gamma Q}$$
$$+ (64w^2\frac{N}{\Gamma^*}L^2 + 96L^3\gamma T\frac{N}{\Gamma^*}w^2)\frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}\|\theta_q\|^2$$
$$+ \frac{8N}{\Gamma^*}(32\gamma^2T^2L^2 + 1 + 96L^3\gamma^3T^3 + 3L\gamma T)\delta^2$$
$$+ \gamma L\frac{8N}{\Gamma^*}(4\gamma TL + \frac{3}{2} + 12L^2\gamma^2T^2)\sigma^2$$

Supposing that the step size $\gamma = O(\sqrt{\frac{\Gamma^*}{TQ}})$ and that $\delta = O(\frac{1}{\sqrt{TQ}})$, when the constant $C > 0$ exists, the convergence rate can be expressed as follows:

$$\frac{1}{Q}\sum_{q=1}^{Q}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2 \le C(\frac{1}{\sqrt{\Gamma^*TQ}} + \frac{1}{Q} + \frac{1}{\Gamma^*TQ} + \frac{1}{Q^{1.5}} + \frac{1}{Q^2} + \frac{1}{Q^{2.5}})$$

A.2 Part Two

Let us start the proof of RAM-Fed from L-Lipschitzian Condition:

$$\sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), \theta^i_{q+1} - \theta^i_q > = \sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), -\gamma \mathrm{v}^{\mathrm{i}}_{\mathrm{q}} >$$
$$= \sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), -\gamma(\frac{1}{N}\sum_{n=1}^{N}u^i_{q,n} + \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}(\Delta^i_{q,n} - u^i_{q,n})) >$$
$$= \sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), -\gamma(\frac{1}{N}\sum_{n=1}^{N}\Delta^i_{q-\tau_q,n} + \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}(\Delta^i_{q,n} - \Delta^i_{q-\tau_q,n})) >$$
$$= \sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), -\gamma(\frac{1}{N}\sum_{n=1}^{N}\frac{(\theta_{q-\tau_q,n,0} - \theta_{q-\tau_q,n,T})^i}{\gamma}$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}(\frac{(\theta_{q,n,0} - \theta_{q,n,T})^i}{\gamma} - \frac{(\theta_{q-\tau_q,n,0} - \theta_{q-\tau_q,n,T})^i}{\gamma})) >$$
$$= \sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), -\gamma(\frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1}) - \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})) >$$
$$= -T\gamma\sum_{i\in S_q}\mathbb{E} < \nabla F^i(\theta_q), (\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})) >$$
$$= -T\gamma\sum_{i\in S_q}[\frac{1}{2}\mathbb{E}\|\nabla F^i(\theta_q)\|^2 + \frac{1}{2}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$- \frac{1}{2}\mathbb{E}\|\nabla F^i(\theta_q) - \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2]$$
$$= -\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2$$
$$- \frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$+ \frac{T\gamma}{2}\underbrace{\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q) - \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})}_{T_1}$$

$$\underbrace{-\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q,n,t-1},\xi_{n,t-1})+\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2}_{T_1}+4\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\underbrace{\mathbb{E}\|\theta_q-\theta_{q-\tau_q}\|^2}_{T_4}+4\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\underbrace{\sum_{t=1}^{T}\mathbb{E}\|\theta_{q-\tau_q}-\theta_{q-\tau_q,n,t-1}\|^2}_{T_5}+2\sigma^2$$

bound T_1:

$$\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})$$
$$-\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q,n,t-1},\xi_{n,t-1})$$
$$+\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2$$
$$\le 2\underbrace{\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2}_{T_2}$$
$$+2\underbrace{\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q,n,t-1},\xi_{n,t-1})}_{T_3}$$
$$\underbrace{-\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2}_{T_3}$$

bound T_2:

$$\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2$$
$$=\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(\nabla F_n^i(\theta_q)-\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1}))\|^2$$
$$=\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(\nabla F_n^i(\theta_q)-\nabla F_n^i(\theta_{q-\tau_q,n,t-1})$$
$$+\nabla F_n^i(\theta_{q-\tau_q,n,t-1})-\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1}))\|^2$$
$$\le 2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(\nabla F_n^i(\theta_q)-\nabla F_n^i(\theta_{q-\tau_q,n,t-1}))\|^2$$
$$+2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}(\nabla F_n^i(\theta_{q-\tau_q,n,t-1})-\nabla F_n^i(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1}))\|^2$$
$$\le 2\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\nabla F_n(\theta_q)-\nabla F_n(\theta_{q-\tau_q,n,t-1})\|^2$$
$$+2\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\nabla F_n(\theta_{q-\tau_q,n,t-1})-\nabla F_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2$$
$$\le 2\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_q-\theta_{q-\tau_q,n,t-1}\|^2+2\sigma^2$$

bound T_5:

Letting $2w_1^2w_2^2+2w_1^2+w_2^2=w^2=w^2$, $8\gamma^2L^2T^2\le\frac{1}{2}\Rightarrow\gamma\le\frac{1}{4LT}$, we have:

$$\sum_{t=1}^{T}\mathbb{E}\|\theta_{q-\tau_q}-\theta_{q-\tau_q,n,t-1}\|^2$$
$$\le 4\gamma^2T\sigma^2+32\gamma^2T^2\delta^2+32\gamma^2T^2G+4w^2\mathbb{E}\|\theta_{q-\tau_q}\|^2$$

bound T_4:

$$\mathbb{E}\|\theta_q-\theta_{q-\tau_q}\|^2=\sum_{i\in S_q}\mathbb{E}\|\theta_q^i-\theta_{q-\tau_q}^i\|^2+\sum_{i\in K-S_q}\mathbb{E}\|\theta_q^i-\theta_{q-\tau_q}^i\|^2$$
$$\le\sum_{i\in S_q}\tau_q\sum_{l=0}^{\tau_q-1}\mathbb{E}\|\theta_{q-l}^i-\theta_{q-(l+1)}^i\|^2+\sum_{i\in K-S_q}\tau_q\sum_{l=0}^{\tau_q-1}\mathbb{E}\|\theta_{q-l}^i-\theta_{q-(l+1)}^i\|^2$$
$$\le\tau_q\sum_{l=0}^{\tau_q-1}\underbrace{\sum_{i\in S_q}\mathbb{E}\|\theta_{q-l}^i-\theta_{q-(l+1)}^i\|^2}_{T_6}+\tau_q\sum_{l=0}^{\tau_q-1}\underbrace{\sum_{i\in K-S_q}\mathbb{E}\|\theta_{q-l}^i-\theta_{q-(l+1)}^i\|^2}_{T_7}$$

bound T_6:

$$\sum_{i\in S_q}\mathbb{E}\|\theta_{q-l}^i-\theta_{q-(l+1)}^i\|^2=\sum_{i\in S_q}\mathbb{E}\|-\gamma v_{q-(l+1)}^i\|^2$$
$$=\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}u_{q-(l+1),n}^i+\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}(\Delta_{q-(l+1),n}^i-u_{q-(l+1),n}^i)\|^2$$
$$=\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\Delta_{q-(l+1)-\tau_{q-(l+1)},n}^i$$
$$+\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}(\Delta_{q-(l+1),n}^i-\Delta_{q-(l+1)-\tau_{q-(l+1)},n}^i)\|^2$$
$$\le 3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\Delta_{q-(l+1)-\tau_{q-(l+1)},n}^i\|^2+3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\Delta_{q-(l+1),n}^i\|^2$$
$$+3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\Delta_{q-(l+1)-\tau_{q-(l+1)},n}^i\|^2$$
$$=3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1},\xi_{n,t-1})\|^2$$
$$+3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-(l+1),n,t-1},\xi_{n,t-1})\|^2$$
$$+3\gamma^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1},\xi_{n,t-1})\|^2$$
$$\le 3\gamma^2\frac{1}{N}\sum_{n=1}^{N}T\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1},\xi_{n,t-1})\|^2$$
$$+3\gamma^2\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}T\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q-(l+1),n,t-1},\xi_{n,t-1})\|^2$$

$$+ 3\gamma^2 \frac{1}{\Gamma_q^i} \sum_{n\in N_q^i} T \sum_{t=1}^{T} \sum_{i\in S_q} \mathbb{E}\|\nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1}, \xi_{n,t-1})\|^2$$
$$\le 3\gamma^2T^2G + 3\gamma^2T^2\frac{N}{\Gamma^*}G + 3\gamma^2T^2\frac{N}{\Gamma^*}G = 3\gamma^2T^2(1+\frac{2N}{\Gamma^*})G$$

bound T_7:

$$\sum_{i\in K-S_q} \mathbb{E}\|\theta^i_{q-l} - \theta^i_{q-(l+1)}\|^2 = \sum_{i\in K-S_q} \mathbb{E}\| - \gamma v^i_{q-(l+1)}\|^2$$
$$= \gamma^2 \sum_{i\in K-S_q} \mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N} u^i_{q-(l+1),n}\|^2$$
$$= \gamma^2 \sum_{i\in K-S_q} \mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N} \Delta^i_{q-(l+1)-\tau_{q-(l+1)},n}\|^2$$
$$= \gamma^2 \sum_{i\in K-S_q} \mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T} \nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1}, \xi_{n,t-1})\|^2$$
$$\le \gamma^2 \frac{1}{N}\sum_{n=1}^{N} T\sum_{t=1}^{T}\sum_{i\in S_q} \mathbb{E}\|\nabla F_n^i(\theta_{q-(l+1)-\tau_{q-(l+1)},n,t-1}, \xi_{n,t-1})\|^2$$
$$\le \gamma^2T^2G$$

Plugging T_6, T_7 into T_4, we have:

$$\mathbb{E}\|\theta_q - \theta_{q-\tau_q}\|^2 \le 3(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + (\tau_q)^2\gamma^2T^2G$$

Plugging T_4, T_5 into T_2, we have:

$$\sum_{i\in S_q} \mathbb{E}\|\nabla F^i(\theta_q) - \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$\le 12(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + 4(\tau_q)^2\gamma^2T^2G$$
$$+ 16\gamma^2T\sigma^2 + 128\gamma^2T^2\delta^2 + 128\gamma^2T^2G + 16w^2\mathbb{E}\|\theta_{q-\tau_q}\|^2$$

bound T_3:

$$\sum_{i\in S_q}\mathbb{E}\|\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$\le \sum_{i\in S_q}\frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1}) - \nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$\le \frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1}) - \nabla F_n^i(\theta_{q,n,t-1})$$
$$+ \nabla F_n^i(\theta_{q,n,t-1}) + \nabla F_n^i(\theta_{q-\tau_q,n,t-1})$$
$$- \nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1}) - \nabla F_n^i(\theta_{q-\tau_q,n,t-1})\|^2$$
$$\le 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1}) - \nabla F_n^i(\theta_{q,n,t-1})$$
$$+ 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q-\tau_q,n,t-1}) - \nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$+ 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in S_q}\mathbb{E}\|\nabla F_n^i(\theta_{q,n,t-1}) - \nabla F_n^i(\theta_{q-\tau_q,n,t-1})\|^2$$
$$\le 6\frac{N}{\Gamma^*}\sigma^2 + 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\nabla F_n(\theta_{q,n,t-1}) - \nabla F_n(\theta_{q-\tau_q,n,t-1})\|^2$$
$$\le 6\frac{N}{\Gamma^*}\sigma^2 + 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}L^2\mathbb{E}\|\theta_{q,n,t-1} - \theta_{q-\tau_q,n,t-1}\|^2$$
$$= 6\frac{N}{\Gamma^*}\sigma^2 + 3\frac{N}{\Gamma^*}\frac{1}{T}\sum_{t=1}^{T}L^2\mathbb{E}\|\theta_{q,n,t-1} - \theta_q + \theta_q - \theta_{q-\tau_q} + \theta_{q-\tau_q} - \theta_{q-\tau_q,n,t-1}\|^2$$
$$\le 6\frac{N}{\Gamma^*}\sigma^2 + 9\frac{N}{\Gamma^*}L^2\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q,n,t-1} - \theta_q\|^2 + 9\frac{N}{\Gamma^*}L^2\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_q - \theta_{q-\tau_q}\|^2$$
$$+ 9\frac{N}{\Gamma^*}L^2\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\|\theta_{q-\tau_q} - \theta_{q-\tau_q,n,t-1}\|^2$$
$$\overset{T4,T5}{\le} 6\frac{N}{\Gamma^*}\sigma^2 + 9\frac{N}{\Gamma^*}L^2(4\gamma^2T\sigma^2 + 32\gamma^2T^2\delta^2 + 32\gamma^2T^2G + 4w^2\mathbb{E}\|\theta_q\|^2)$$
$$+ 9\frac{N}{\Gamma^*}L^2(3(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + (\tau_q)^2\gamma^2T^2G)$$
$$+ 9\frac{N}{\Gamma^*}L^2(4\gamma^2T\sigma^2 + 32\gamma^2T^2\delta^2 + 32\gamma^2T^2G + 4w^2\mathbb{E}\|\theta_{q-\tau_q}\|^2)$$
$$= 6\frac{N}{\Gamma^*}\sigma^2 + 18\frac{N}{\Gamma^*}L^2(4\gamma^2T\sigma^2 + 32\gamma^2T^2\delta^2 + 32\gamma^2T^2G)$$
$$+ 36\frac{N}{\Gamma^*}L^2w^2(\mathbb{E}\|\theta_q\|^2 + \mathbb{E}\|\theta_{q-\tau_q}\|^2)$$
$$+ 9\frac{N}{\Gamma^*}L^2(3(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + (\tau_q)^2\gamma^2T^2G)$$

Plugging T_2, T_3 into T_1, we have:

$$\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q) - \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})$$
$$- \frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q,n,t-1}, \xi_{n,t-1})$$
$$+ \frac{1}{\Gamma_q^i}\sum_{n\in N_q^i}\frac{1}{T}\sum_{t=1}^{T}\nabla F_n^i(\theta_{q-\tau_q,n,t-1}, \xi_{n,t-1})\|^2$$
$$\le 24(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + 8(\tau_q)^2\gamma^2T^2G$$
$$+ 32\gamma^2T\sigma^2 + 256\gamma^2T^2\delta^2 + 256\gamma^2T^2G + 32w^2\mathbb{E}\|\theta_{q-\tau_q}\|^2$$
$$+ 4\sigma^2 + 12\frac{N}{\Gamma^*}\sigma^2 + 36\frac{N}{\Gamma^*}L^2(4\gamma^2T\sigma^2 + 32\gamma^2T^2\delta^2 + 32\gamma^2T^2G)$$
$$+ 72\frac{N}{\Gamma^*}L^2w^2(\mathbb{E}\|\theta_q\|^2 + \mathbb{E}\|\theta_{q-\tau_q}\|^2)$$
$$+ 18\frac{N}{\Gamma^*}L^2(3(\tau_q)^2\gamma^2T^2G(1+\frac{2N}{\Gamma^*}) + (\tau_q)^2\gamma^2T^2G)$$
$$= (32(\tau_q)^2 + 256 + 1152\frac{N}{\Gamma^*}L^2 + 48\frac{N}{\Gamma^*}(\tau_q)^2 + 72\frac{N}{\Gamma^*}L^2(\tau_q)^2$$
$$+ 108\frac{N^2}{(\Gamma^*)^2}L^2(\tau_q)^2)\gamma^2T^2G$$
$$+ (32\gamma^2T + 12\frac{N}{\Gamma^*} + 4 + 144\frac{N}{\Gamma^*}L^2\gamma^2T)\sigma^2$$
$$+ 128\gamma^2T^2\delta^2(2 + 9\frac{N}{\Gamma^*}L^2)$$
$$+ 8w^2(4 + 9\frac{N}{\Gamma^*}L^2)\mathbb{E}\|\theta_{q-\tau_q}\|^2 + 72\frac{N}{\Gamma^*}L^2w^2\mathbb{E}\|\theta_q\|^2$$

For another term in L-Lipschitzian condition, we have:

$$\begin{aligned}
&\frac{L}{2}\sum_{i\in S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&=\frac{L}{2}\gamma^2T^2\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\\
&+\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1},\xi_{n,t-1})\\
&-\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2
\end{aligned}$$

For $i\in S_q$, we get:

$$\begin{aligned}
&\sum_{i\in S_q}\mathbb{E}<\nabla F^i(\theta_q),\theta^i_{q+1}-\theta^i_q>+\frac{L}{2}\sum_{i\in S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&=-\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2\\
&+(\frac{L}{2}\gamma^2T^2-\frac{T\gamma}{2})\sum_{i\in S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\\
&+\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1},\xi_{n,t-1})\\
&-\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&+\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\\
&-\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q,n,t-1},\xi_{n,t-1})\\
&+\frac{1}{\Gamma^i_q}\sum_{n\in N^i_q}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&\overset{b}{\le}-\frac{T\gamma}{2}\sum_{i\in S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2+\frac{T\gamma}{2}(T_1)
\end{aligned}$$

where b follows because: $\frac{L}{2}\gamma^2T^2-\frac{T\gamma}{2}<0\Rightarrow\gamma<\frac{1}{LT}$

Then:

$$\begin{aligned}
&\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),\theta^i_{q+1}-\theta^i_q>=\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),-\gamma \mathrm{v}^{\mathrm{i}}_{\mathrm{q}}>\\
&=\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),-\gamma\frac{1}{N}\sum_{n=1}^{N}u^i_{q,n}>\\
&=\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),-\gamma\frac{1}{N}\sum_{n=1}^{N}\Delta^i_{q-\tau_q,n}>\\
&=\sum_{i\in S_q}\mathbb{E}<\nabla F^i(\theta_q),-\gamma\frac{1}{N}\sum_{n=1}^{N}\frac{(\theta_{q-\tau_q,n,0}-\theta_{q-\tau_q,n,T})^i}{\gamma}\\
&=\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),-\gamma\frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})>\\
&=-T\gamma\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})>\\
&=-T\gamma\sum_{i\in K-S_q}[\frac{1}{2}\mathbb{E}\|\nabla F^i(\theta_q)\|^2\\
&+\frac{1}{2}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&-\frac{1}{2}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2]\\
&=-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2\\
&-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&+\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\nabla F^i(\theta_q)-\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&=-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2\\
&-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2\\
&+\frac{T\gamma}{2}\sum_{i\in K-S_q}(T_2)
\end{aligned}$$

For another term in L-Lipschitzian condition, we have:

$$\begin{aligned}
&\frac{L}{2}\sum_{i\in K-S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&=\frac{L}{2}\gamma^2T^2\sum_{i\in K-S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2
\end{aligned}$$

For $i\in K-S_q$, we get:

$$\begin{aligned}
&\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),\theta^i_{q+1}-\theta^i_q>+\frac{L}{2}\sum_{i\in K-S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&\le-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2\\
&+(\frac{L}{2}\gamma^2T^2-\frac{T\gamma}{2})\sum_{i\in K-S_q}\mathbb{E}\|\frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\nabla F^i_n(\theta_{q-\tau_q,n,t-1},\xi_{n,t-1})\|^2+\frac{T\gamma}{2}(T_2)\\
&\overset{b}{\le}-\frac{T\gamma}{2}\sum_{i\in K-S_q}\mathbb{E}\|\nabla F^i(\theta_q)\|^2+\frac{T\gamma}{2}(T_2)
\end{aligned}$$

Combining $i\in S_q$ and $i\in K-S_q$:

$$\begin{aligned}
&\mathbb{E}<\nabla F(\theta_q),\theta_{q+1}-\theta_q>+\frac{L}{2}\mathbb{E}\|\theta_{q+1}-\theta_q\|^2\\
&=\sum_{i\in S_q}\mathbb{E}<\nabla F^i(\theta_q),\theta^i_{q+1}-\theta^i_q>+\frac{L}{2}\sum_{i\in S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&+\sum_{i\in K-S_q}\mathbb{E}<\nabla F^i(\theta_q),\theta^i_{q+1}-\theta^i_q>+\frac{L}{2}\sum_{i\in K-S_q}\mathbb{E}\|\theta^i_{q+1}-\theta^i_q\|^2\\
&\le-\frac{T\gamma}{2}\mathbb{E}\|\nabla F(\theta_q)\|^2+\frac{T\gamma}{2}(T_1+T_2)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[F(\theta_{q+1})] - \mathbb{E}[F(\theta_q)] \\
&\leq \mathbb{E} < \nabla F(\theta_q), \theta_{q+1} - \theta_q > + \frac{L}{2}\mathbb{E}\|\theta_{q+1} - \theta_q\|^2 \\
&\leq -\frac{T\gamma}{2}\mathbb{E}\|\nabla F(\theta_q)\|^2 + \frac{T\gamma}{2}(T_1 + T_2)
\end{aligned}$$

Then, we can obtain:

$$\begin{aligned}
&\mathbb{E}[F(\theta_{Q+1})] - \mathbb{E}[F(\theta_1)] = \sum_{q=1}^{Q}\mathbb{E}[F(\theta_{q+1})] - \sum_{q=1}^{Q}\mathbb{E}[F(\theta_q)] \\
&\leq -\frac{T\gamma}{2}\sum_{q=1}^{Q}\mathbb{E}\|\nabla F(\theta_q)\|^2 + \frac{T\gamma}{2}\sum_{q=1}^{Q}(T_1 + T_2)
\end{aligned}$$

Re-arranging the terms:

$$\frac{T\gamma}{2}\sum_{q=1}^{Q}\mathbb{E}\|\nabla F(\theta_q)\|^2 \leq \mathbb{E}[F(\theta_1)] - \mathbb{E}[F(\theta_{Q+1})] + \frac{T\gamma}{2}\sum_{q=1}^{Q}(T_1 + T_2)$$

Letting $\frac{1}{Q}\sum_{q=1}^{Q}(\tau_q)^2 = \tau$ and dividing both sides by $\frac{T\gamma Q}{2}$

$$\begin{aligned}
&\frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}\|\nabla F(\theta_q)\|^2 \leq \frac{2\mathbb{E}[F(\theta_1)]}{T\gamma Q} \\
&+ (48\tau + 384 + 1152\frac{N}{\Gamma^*}L^2 + 72\frac{N}{\Gamma^*}\tau + 72\frac{N}{\Gamma^*}L^2\tau + 108(\frac{N}{\Gamma^*})^2L^2\tau)\gamma^2T^2G \\
&+ (48\gamma^2T + 12\frac{N}{\Gamma^*} + 6 + 144\frac{N}{\Gamma^*}L^2\gamma^2T)\sigma^2 + 128(3 + 9\frac{N}{\Gamma^*}L^2)\gamma^2T^2\delta^2 \\
&+ 8w^2(6 + 9\frac{N}{\Gamma^*}L^2)\frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}\|\theta_{q-(\tau_q)}\|^2 + 72\frac{N}{\Gamma^*}L^2w^2\frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}\|\theta_q\|^2
\end{aligned}$$

Supposing that the step size $\gamma = O(\sqrt{\frac{\Gamma^*}{TQ}})$ and σ is sufficiently small, when the constant $C > 0$ exists, the convergence rate can be expressed as follows:

$$\frac{1}{Q}\sum_{q=1}^{Q}\mathbb{E}\|\nabla F(\theta_q)\|^2 \leq C(\frac{1}{\sqrt{\Gamma^*TQ}} + \frac{1}{Q} + \frac{1}{\Gamma^*Q} + \frac{1}{Q^2})$$