## APPENDIX A

With Equation (14), the second-order Taylor expansion of $H(\boldsymbol{w}^{q+1})$ is

$$\begin{aligned} H(\boldsymbol{w}^{q+1}) &= H(\boldsymbol{w}^{q}) + (\boldsymbol{w}^{q+1}-\boldsymbol{w}^{q})^T \nabla H(\boldsymbol{w}^{q}) \\ &+ \frac{1}{2}(\boldsymbol{w}^{q+1}-\boldsymbol{w}^{q})^T \nabla^2 H(\boldsymbol{w}^{q})(\boldsymbol{w}^{q+1}-\boldsymbol{w}^{q}) \\ &\leq H(\boldsymbol{w}^{q}) + (\boldsymbol{w}^{q+1}-\boldsymbol{w}^{q})^T \nabla H(\boldsymbol{w}^{q}) \\ &+ \frac{l}{2}||\boldsymbol{w}^{q+1}-\boldsymbol{w}^{q}||^2. \end{aligned} \tag{1}$$

Let $\eta = l^{-1}$. According to Equation (11), the expect of loss function is

$$\begin{aligned} \mathbb{E}(H(\boldsymbol{w}^{q+1})) &\leq \mathbb{E}(H(\boldsymbol{w}^{q}) - \eta(\nabla H(\boldsymbol{w}^{q}) - o)^T \nabla H(\boldsymbol{w}^{q}) \\ &+ \frac{l\eta^2}{2}||\nabla H(\boldsymbol{w}^{q}) - o||^2) \\ &= \mathbb{E}(H(\boldsymbol{w}^{q})) - \frac{1}{2l}\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2) \\ &+ \frac{1}{2l}\mathbb{E}(||o||^2), \end{aligned} \tag{2}$$

where the equation is established based on the property of expect. Furthermore,

$$\begin{aligned} &\mathbb{E}(||o||^2) \\ =&\mathbb{E}\left(\|\nabla H(\boldsymbol{w}^{q}) - G/Q\|^2\right) \\ =&\mathbb{E}\left\{\left\| -\frac{(N-Q)\sum\limits_{m\in Y}\sum\limits_{u_m\in X}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})}{NQ} \right.\right. \\ &+ \frac{\sum\limits_{m\in Y}\sum\limits_{u_m\notin X}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \nabla h(\boldsymbol{x}_{u_m i}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})}{N} \\ &\left.\left. + \frac{\sum\limits_{m\notin Y}\sum\limits_{u_m=1}^{U_m}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})}{N}\right\|^2\right\} \\ \leq&\mathbb{E}\left\{\frac{(N-Q)\sum\limits_{m\in Y}\sum\limits_{u_m\in X}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \left\|\nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})\right\|}{NQ}\right. \\ &+ \frac{\sum\limits_{m\in Y}\sum\limits_{u_m\notin X}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \left\|\nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})\right\|}{N} \\ &\left. + \frac{\sum\limits_{m\notin Y}\sum\limits_{u_m=1}^{U_m}\sum\limits_{i_{u_m}=1}^{n_{u_m m}} \nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})}{N}\right\}^2, \end{aligned} \tag{3}$$

where $X = \{u_m | I_{u_m} = 1, \tau_{u_m m} = 1, Z_{u_m} = 1\}$, $Y = \{m | I_m = 1, \pi_m = 1, Z_m^{up} = 1\}$. The inequality is established based on the triangle inequality.

Let $\xi(\boldsymbol{w}^{q}) = \sqrt{\xi_1 + \xi_2||\nabla H(\boldsymbol{w}^{q})||^2}$. According to Equation (15), $\| \nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q}) \| \leq \xi(\boldsymbol{w}^{q})$. Then, $\sum_{m\in Y}\sum_{u_m\in X}\sum_{i=1}^{n_{u_m m}} \left\|\nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})\right\| \leq \xi(\boldsymbol{w}^{q})Q$, $\sum_{m\in Y}\sum_{u_m\notin X}\sum_{i=1}^{n_{u_m m}} \left\|\nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})\right\| \leq \xi(\boldsymbol{w}^{q})z$, $\sum_{m\notin Y}\sum_{u_m=1}^{U_m}\sum_{i=1}^{n_{u_m m}} \left\|\nabla h(\boldsymbol{x}_{u_m i_{u_m}}, y_{u_m i_{u_m}}, \boldsymbol{w}^{q})\right\| \leq \xi(\boldsymbol{w}^{q})(N-Q-z)$, where $z = \sum_{m=1}^{A+1}\left[Z_m^{up}\, \pi_m I_m \sum_{u_m=1}^{U_m}(1 - I_{u_m}\tau_{u_m m}Z_{u_m})\, n_{u_m m}\right]$. So, $\mathbb{E}(||o||^2) \leq \frac{4}{N^2}\left(\xi_1 + \xi_2\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2)\right)\mathbb{E}(N-Q)^2$.

With $N \geq N - Q \geq 0$, $\mathbb{E}(Z_{u_m}) = 1 - z_{u_m}$, $\mathbb{E}(Z_m^{up}) = 1 - z_m^{up}$, it is known that

$$\begin{aligned} \mathbb{E}(||o||^2) &\leq \frac{4}{N}\left(\xi_1 + \xi_2\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2)\right)\mathbb{E}\left(N-Q\right) \\ &= \frac{4F}{N}\left(\xi_1 + \xi_2\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2)\right), \end{aligned} \tag{4}$$

where $F = \sum\limits_{m=1}^{A+1}\left\{\sum\limits_{u_m=1}^{U_m} n_{u_m m} - I_m\pi_m\left(1 - z_m^{up}\right)\sum\limits_{u_m=1}^{U_m}[I_{u_m}\cdot \tau_{u_m m}\left(1 - z_{u_m}\right)n_{u_m m}]\right\}$.

Based on Equation (2) and (4) in APPENDIX A, we have

$$\begin{aligned} \mathbb{E}(H(\boldsymbol{w}^{q+1})) &\leq \mathbb{E}(H(\boldsymbol{w}^{q})) + \frac{2\xi_1 F}{lN} \\ &- \frac{1}{2l}\left(1 - \frac{4\xi_2 F}{N}\right)\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2). \end{aligned} \tag{5}$$

Then,

$$\begin{aligned} &\mathbb{E}(H(\boldsymbol{w}^{q+1}) - H(\boldsymbol{w}^{*})) \\ \leq&\mathbb{E}(H(\boldsymbol{w}^{q}) - H(\boldsymbol{w}^{*})) + \frac{2\xi_1 F}{lN} \\ -&\frac{1}{2l}\left(1 - \frac{4\xi_2 F}{N}\right)\mathbb{E}(||\nabla H(\boldsymbol{w}^{q})||^2). \end{aligned} \tag{6}$$

According to (17) and (19), it is known that $||\nabla H(\boldsymbol{w}^{q})||^2 \geq 2\zeta(H(\boldsymbol{w}^{q}) - H(\boldsymbol{w}^{*}))$. Then, $\mathbb{E}(H(\boldsymbol{w}^{q+1}) - H(\boldsymbol{w}^{*})) \leq B\mathbb{E}(H(\boldsymbol{w}^{q}) - H(\boldsymbol{w}^{*})) + \frac{2\xi_1 F}{lN}$, where $B = 1 - \frac{\zeta}{l} + \frac{4\zeta\xi_2 F}{lN}$. Furthermore, with the recursive method, we have $\mathbb{E}(H(\boldsymbol{w}^{q+1}) - H(\boldsymbol{w}^{*})) \leq B^{q+1}\mathbb{E}(H(\boldsymbol{w}^{0}) - H(\boldsymbol{w}^{*})) + G$, where $G = \frac{2\xi_1 F}{lN} \times \frac{(1-B^{q+1})}{1-B}$.

## APPENDIX B

With **Theorem 1**, when $B < 1$, $\mathbb{E}(H(\boldsymbol{w}^{q+1}) - H(\boldsymbol{w}^{*})) \leq G$. Meanwhile, the federated learning is convergent. Then, $1 - \frac{\zeta}{l} + \frac{4\zeta\xi_2 F}{lN} < 1$. With Equation (14), $\zeta/l < 1$. Then, $\frac{4\zeta\xi_2 F}{lN} - \frac{\zeta}{l} < 0$. So, $\xi_2 < \frac{N}{4F}$, which is valid for all variables. So, $\xi_2 < \frac{N}{\max\limits_{\boldsymbol{\tau},\boldsymbol{f},\boldsymbol{p},\boldsymbol{t}} 4F}$. Meanwhile, according to Equation (15), it is known the conclusion in **Proposition 1**.