A Proof of Lemma 1

Proof. Since each individual function is L-smooth, we have

$$\sum_{t=0}^{N-1} \|\nabla f_{i_t}(w_t^s) - \nabla f_{i_t}(w_0^s)\| \leq L \sum_{t=0}^{N-1} \|w_t^s - w_0^s\| \tag{19}$$

Next, we bound $\sum_{t=0}^{N-1} \|w_t^s - w_0^s\|$ using triangle inequality:

$$\begin{aligned} &\sum_{t=0}^{N-1} \|w_t^s - w_0^s\| \\ &= \sum_{t=1}^{N-1} \|w_t^s - w_{t-1}^s + w_{t-1}^s - \cdots - w_0^s\| \\ &\leq \sum_{t=1}^{N-1} \sum_{m=0}^{t-1} \|w_{m+1}^s - w_m^s\| \end{aligned} \tag{20}$$

We can bound $\|w_{t+1}^s - w_t^s\|$ as follows:

$$\begin{aligned} &\|w_{t+1}^s - w_t^s\| \\ &= \eta \Big\|\nabla f_{i_t}(w_t^s) - \nabla f_{i_t}(w_0^s) + \frac{1}{N}\sum_{i=1}^{N} \nabla f_i(w_0^s)\Big\| \\ &\leq \eta \Big(\big\|\nabla f_{i_t}(w_t^s) - \nabla f_{i_t}(w_0^s)\big\| \\ &\quad + \frac{1}{N}\sum_{i=1}^{N} \big\|\big(\nabla f_i(w_0^s) - \nabla f_i(w^*)\big)\big\| \Big) \\ &\leq \eta L \big(\|w_t^s - w_0^s\| + \|w_0^s - w^*\|\big) \end{aligned} \tag{21}$$

where the first inequality follows from triangle inequality and uses the fact that $\nabla F(w^*) = 0$; and the second inequality holds because each f_i is L-smooth.

Combing (20) and (21), we have

$$\begin{aligned} &\sum_{t=0}^{N-1} \|w_t^s - w_0^s\| \\ &\leq \eta L \sum_{t=1}^{N-1} \sum_{m=0}^{t-1} \Big(\|w_m^s - w_0^s\| + \|w_0^s - w^*\| \Big) \\ &= \eta L \sum_{m=0}^{N-2} \sum_{t=m+1}^{N-1} \Big(\|w_m^s - w_0^s\| + \|w_0^s - w^*\| \Big) \\ &\leq \eta L \sum_{m=0}^{N-2} \sum_{t=m+1}^{N-1} \|w_m^s - w_0^s\|^2 + \frac{\eta L N^2}{2} \|w_0^s - w^*\| \\ &\leq \eta L N \sum_{m=0}^{N-1} \|w_m^s - w_0^s\|^2 + \frac{\eta L N^2}{2} \|w_0^s - w^*\| \end{aligned} \tag{22}$$

The first equality of (22) holds because of the fact that

$$\sum_{t=1}^{N-1} \sum_{m=0}^{t-1} h(m,t) \equiv \sum_{m=0}^{N-2} \sum_{t=m+1}^{N-1} h(m,t),$$

where $h(m,t)$ is an arbitrary function of variables m and t. The second inequality holds because $\sum_{m=0}^{N-2} \sum_{t=m+1}^{N-1} 1 \leq \frac{N^2}{2}$. Rearrange terms of (22), and use the assumption $\eta < \frac{1}{NL}$, we have

$$\sum_{t=0}^{N-1} \|w_t^s - w_0^s\| \leq \frac{\eta L N^2}{2(1 - \eta L N)} \|w_0^s - w^*\|,$$

Therefore, we have

$$\sum_{t=0}^{N-1} \|\nabla f_{i_t}(w_t^s) - \nabla f_{i_t}(w_0^s)\| \leq \frac{\eta L^2 N^2}{2(1 - \eta L N)} \|w_0^s - w^*\|,$$

which is the desired result. □