A PROOF OF THEOREM 1

A.1 Additional Notation

For simplicity of notations, we denote the error of weight quantization $\boldsymbol{r}_n^k \triangleq Q_w\left(\boldsymbol{w}_n^{k+1}\right) - \boldsymbol{w}_n^{k+1}$, and the local "gradient" with weight quantization as $\hat{\boldsymbol{g}}_n^k \triangleq \nabla \widetilde{f}_n(\boldsymbol{w}_n^k) - \boldsymbol{r}_n^k/\eta$.

Inspired by the iterate analysis framework in [24], we define the following virtual sequences:

$$\boldsymbol{u}_n^{k+1} = \boldsymbol{w}_n^k - \eta \hat{\boldsymbol{g}}_n^k, \tag{23}$$

$$\boldsymbol{w}_n^{k+1} = \begin{cases} \boldsymbol{u}_n^{k+1}, & k+1 \notin \mathcal{U}_H, \\ \boldsymbol{u}_n^{k'} - \sum_{n=1}^N p_n Q_g(\Delta_n^{k'}), & k+1 \in \mathcal{U}_H. \end{cases} \tag{24}$$

Here, $k' = k+1-H$ is the last synchronization step and $\Delta_N^k = \boldsymbol{u}_n^{k'} - \boldsymbol{u}_n^{k+1}$ is the differences since the last synchronization. The following short-hand notation will be found useful in the convergence analysis of the proposed FL framework: $\bar{\boldsymbol{u}}^k = \sum_{n=1}^N p_n \boldsymbol{u}_n^k$, $\bar{\boldsymbol{w}}_n^k = \sum_{n=1}^N p_n \boldsymbol{w}_n^k$, and $\hat{\boldsymbol{g}}^k = \sum_{n=1}^N p_n \hat{\boldsymbol{g}}_n^k$. Thus, $\bar{\boldsymbol{u}}^{k+1} = \bar{\boldsymbol{w}}^k - \eta \hat{\boldsymbol{g}}^k$. Note that we can only obtain $\bar{\boldsymbol{w}}^{k+1}$ when $k+1 \in \mathcal{U}_H$. Further, due to the unbiased gradient quantization scheme, Q_g, no matter whether $k+1 \in \mathcal{U}_H$ or $k+1 \notin \mathcal{U}_H$, we always have $\mathbb{E}[\mathbb{E}_Q[\bar{\boldsymbol{w}}^{k+1}]] = \mathbb{E}[\bar{\boldsymbol{u}}^{k+1}]$.

A.2 Key Lemmas

Now, we give four important lemmas to convey our proof.

Lemma 1 (Weight quantization error [21]). *The stochastic rounding error on each iteration can be bounded, in expectation*

$$\mathbb{E}_Q\left[\left\|\boldsymbol{r}_n^k\right\|_2^2\right] \leq \eta\sqrt{d}\delta_w\tau. \tag{25}$$

Lemma 2 (Bounding the divergence). *According to the Alg. 1 the expected inner product between stochastic gradient and full batch gradient is bounded with:*

$$\sum_{k=0}^{K-1}\sum_{n=1}^N p_n \mathbb{E}\left[\mathbb{E}_Q\left[\left\|\bar{\boldsymbol{w}}^k - \boldsymbol{w}_n^k\right\|_2^2\right]\right]$$
$$\leq \frac{\eta^2 KH\sigma^2/M + 3\eta^2 KH^2G^2}{1-3\eta^2L^2H^2} + \frac{\eta KH\sqrt{d}\tau}{1-3\eta^2L^2H^2}\sum_{n=1}^N p_n\delta_{w,n}$$
$$+ \frac{3\eta^2H^2}{1-3\eta^2L^2H^2}\sum_{i=0}^{K-1}\left\|\nabla F(\bar{\boldsymbol{w}}^i)\right\|_2^2 \tag{26}$$

Proof. Recalling that at the synchronization step $k' \in \mathcal{U}_H$, $\boldsymbol{w}_n^{k'} = \bar{\boldsymbol{w}}^{k'}$ for all $n \in \mathcal{N}$. Therefore, for any $k \geq 0$, such that $k' \leq k \leq k' + H$, we get,

$$A1_k := \sum_{n=1}^N p_n \mathbb{E}\left[\mathbb{E}_Q\left[||\bar{\boldsymbol{w}}^k - \boldsymbol{w}_n^k||_2^2\right]\right]$$
$$= \sum_{n=1}^N p_n \mathbb{E}\left[\mathbb{E}_Q\left[||(\bar{\boldsymbol{w}}^k - \bar{\boldsymbol{w}}^{k'}) - (\boldsymbol{w}_n^k - \bar{\boldsymbol{w}}^{k'})||_2^2\right]\right]$$
$$\overset{(a)}{\leq} \sum_{n=1}^N p_n \mathbb{E}\left[\mathbb{E}_Q\left[||\boldsymbol{w}_n^k - \boldsymbol{w}_n^{k'}||_2^2\right]\right]$$
$$= \sum_{n=1}^N p_n \mathbb{E}\left[\left\|\sum_{i=k'}^k\left(\eta\nabla\widetilde{f}_n(\boldsymbol{w}_n^i) - \boldsymbol{r}_n^i\right)\right\|_2^2\right]$$
$$= \sum_{n=1}^N p_n \mathbb{E}\left[\left\|\sum_{i=k'}^k \eta\nabla\widetilde{f}_n(\boldsymbol{w}_n^i)\right\|_2^2 + \sum_{i=k'}^k \mathbb{E}_Q\left[||\boldsymbol{r}_n^i||_2^2\right]\right]$$
$$\leq \eta^2 \sum_{n=1}^N p_n \mathbb{E}\left[\left\|\sum_{i=k'}^{k'_H}\nabla\widetilde{f}_n(\boldsymbol{w}_n^i)\right\|_2^2\right] + H\eta\sqrt{d}\tau\sum_{n=1}^N p_n\delta_{w,n}, \tag{27}$$

where $k'_H = k' + H - 1$, (a) holds due to $\mathbb{E}[\left\|\sum_{i=1}^n a_i\right\|_2^2] = \sum_{i=1}^n \mathbb{E}[\|a_i\|_2^2]$ if $\mathbb{E}[a_i] = 0$. The last equality is due to Lemma 1.

We generalize the result from [19] to upper-bound the first term in RHS of (27), (see the of Theorem 3 and its proof in appendix for the special case of $p_n = \frac{1}{N}$):

$$\eta^2\sum_{n=1}^N p_n\mathbb{E}\left[\left\|\sum_{i=k'}^{k'_H}(\nabla\widetilde{f}_n(\boldsymbol{w}_n^i) - \nabla F_n(\boldsymbol{w}_n^i) + \nabla F_n(\boldsymbol{w}_n^i))\right\|_2^2\right]$$
$$\leq \eta^2 H\frac{\sigma^2}{M} + 3\eta^2H^2G^2 + 3\eta^2 H\sum_{i=k'}^{k'_H}(L^2 A1_i + \left\|\nabla F(\bar{\boldsymbol{w}}^i)\right\|_2^2). \tag{28}$$

It follows that

$$\sum_{i=0}^{K-1} A1_i \leq \eta^2 KH\frac{\sigma^2}{M} + 3\eta^2KH^2G^2 + 3\eta^2L^2H^2\sum_{i=0}^{K-1}A1_i$$
$$+ 3\eta^2H^2\sum_{i=0}^{K-1}\left\|\nabla F(\bar{\boldsymbol{w}}^i)\right\|_2^2 + \eta KH\sqrt{d}\tau\sum_{n=1}^N p_n\delta_{w,n}. \tag{29}$$

Suppose $1 - 3\eta^2L^2H^2 \geq 0$, we have

$$\sum_{i=0}^{K-1}A1_i \leq \frac{\eta^2 KH\sigma^2/M + 3\eta^2KH^2G^2}{1-3\eta^2L^2H^2} + \frac{\eta KH\sqrt{d}\tau}{1-3\eta^2L^2H^2}\sum_{n=1}^N p_n\delta_{w,n}$$
$$+ \frac{3\eta^2H^2}{1-3\eta^2L^2H^2}\sum_{i=0}^{K-1}\left\|\nabla F(\bar{\boldsymbol{w}}^i)\right\|_2^2, \tag{30}$$

and the proof complete. □

Lemma 3. *Under the Assumptions 1-3, the following conditions hold*

$$\frac{L}{2}\mathbb{E}\left[\mathbb{E}_Q\left[\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2\right]\right] \leq \frac{\eta L}{2}\sum_{n=1}^{N}\left(\eta H\tau^2 p_n\delta_{g,n}+H\sqrt{d}\tau p_n\delta_{g,n}\delta_{w,n}+\eta\frac{\sigma^2}{M}p_n^2\right) + \frac{\eta^2 L}{2}\left\|\sum_{n=1}^{N}p_n\nabla F_n(\boldsymbol{w}_n^k)\right\|_2^2+\frac{\eta LKH\sqrt{d}\tau}{2}\sum_{n=1}^{N}p_n^2\delta_{w,n}. \quad (31)$$

Proof.

$$\begin{aligned}
&\frac{L}{2}\mathbb{E}\left[\mathbb{E}_Q\left[\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2\right]\right]\\
&=\frac{L}{2}\mathbb{E}\left[\mathbb{E}_Q\left[\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{u}}^{k+1}+\bar{\boldsymbol{u}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2\right]\right]\\
&=\mathbb{E}\left[\frac{L}{2}\mathbb{E}_Q\left[\left\|\sum_{n=1}^{N}p_n(Q(\widetilde{\Delta}_n^k)-\widetilde{\Delta}_n^k)\right\|_2^2\right]+\frac{L}{2}||\eta\hat{\boldsymbol{g}}^k||_2^2\right]\\
&\leq\frac{L}{2}\sum_{n=1}^{N}p_n^2\delta_{g,n}\mathbb{E}\left[||\boldsymbol{w}_n^k-\bar{\boldsymbol{w}}^{k'}||_2^2\right]+\frac{\eta^2 L\sigma^2}{2M}\sum_{n=1}^{N}p_n^2\\
&\quad+\frac{\eta^2 L}{2}\left(\left\|\sum_{n=1}^{N}p_n\nabla F_n(\boldsymbol{w}_n^k)\right\|_2^2+\mathbb{E}\left[\left\|\sum_{n=1}^{N}p_n\boldsymbol{r}_n^k/\eta\right\|_2^2\right]\right)\\
&\leq\frac{\eta^2 L}{2}\sum_{n=1}^{N}p_n^2\delta_{g,n}\left(H\frac{\sigma^2}{M}+3HG^2\right)+3\frac{\eta^2L^3}{2}H\sum_{i=k'}^{k'_H}A2_i\\
&\quad+\frac{3\eta^2 L}{2}H\sum_{n=1}^{N}p_n^2\delta_{g,n}\sum_{i=k'}^{k'_H}\left\|\nabla F(\bar{\boldsymbol{w}}^i)\right\|_2^2+\frac{\eta^2L\sigma^2/M\sum_{n=1}^{N}p_n^2}{2}\\
&+\frac{\eta^2 L}{2}\left\|\sum_{n=1}^{N}p_n\nabla F_n(\boldsymbol{w}_n^k)\right\|_2^2+\frac{\eta L\sqrt{d}\tau}{2}\sum_{n=1}^{N}p_n^2\delta_{w,n}(1+H\delta_{g,n}).
\end{aligned}\quad (32)$$

where $A2_i=\sum_{n=1}^{N}p_n\delta_{g,n}\mathbb{E}[\mathbb{E}_Q[||\bar{\boldsymbol{w}}^k-\boldsymbol{w}_n^k||_2^2]]$. The result of $A2_i$ can be bounded using similar analysis of (29).

□

A.3 Main Results

Under the L-smooth assumption of F, we have,

$$\begin{aligned}
&\mathbb{E}\left[F(\bar{\boldsymbol{w}}^{k+1})-F(\bar{\boldsymbol{w}}^{k})\right]\\
&\leq\mathbb{E}\left[\langle\nabla F(\bar{\boldsymbol{w}}^k),\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{u}}^{k+1}+\bar{\boldsymbol{u}}^{k+1}-\bar{\boldsymbol{w}}^{k}\rangle\right]+\frac{L}{2}\mathbb{E}\left[\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2\right]\\
&\overset{(a)}{\leq}\mathbb{E}\left[\langle\nabla F(\bar{\boldsymbol{w}}^k),\bar{\boldsymbol{u}}^{k+1}-\bar{\boldsymbol{w}}^{k}\rangle+\frac{L}{2}\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2\right]\\
&\overset{(b)}{\leq}\frac{\eta}{2}\mathbb{E}_Q\left[\sum_{n=1}^{N}L^2p_n\left\|\bar{\boldsymbol{w}}^k-\boldsymbol{w}_n^k\right\|_2^2-\left\|\nabla F(\bar{\boldsymbol{w}}^k)\right\|_2^2-\left\|\sum_{n=1}^{N}p_n\nabla F_n(\boldsymbol{w}_n^k)\right\|_2^2\right]\\
&\quad+\frac{L}{2}\left\|\bar{\boldsymbol{w}}^{k+1}-\bar{\boldsymbol{w}}^{k}\right\|_2^2,
\end{aligned}\quad (33)$$

where (a) is $\mathbb{E}[\mathbb{E}_Q[\bar{\boldsymbol{w}}^{k+1}]]=\mathbb{E}[\bar{\boldsymbol{u}}^{k+1}]$, (b) is due to $2<a,b>=||a||^2+||b||^2+||a-b||^2$ and $\mathbb{E}[\hat{\boldsymbol{g}}_n^k]=\nabla F_n(\boldsymbol{w}_n^k)$, and follows from L-smoothness assumption.. We use Lemma 1-3 to upper bound the RHS of (33) and set $\eta L\leq 1$, which gets,

$$\mathbb{E}[F(\bar{\boldsymbol{w}}^{k+1})-F(\bar{\boldsymbol{w}}^{k})]\leq-\frac{\eta}{2}\mathbb{E}\left[\left\|\nabla F(\bar{\boldsymbol{w}}^k)\right\|_2^2\right]+\frac{\eta L^2}{2}A1_k+\frac{L}{2}A2_k+\frac{\eta^2L\sigma^2/M}{2}\sum_{n=1}^{N}p_n^2+\frac{\eta L\sqrt{d}\tau}{2}\sum_{n=1}^{N}p_n^2\delta_{w,n}. \quad (34)$$

Summing up for all K communication rounds and we get,

$$\begin{aligned}
&\mathbb{E}\left[F(\bar{\boldsymbol{w}}^{K})-F(\bar{\boldsymbol{w}}^{0})\right]\\
&\leq-\frac{\eta C_1}{2}\sum_{k=0}^{K-1}\mathbb{E}\left[\left\|\nabla F(\bar{\boldsymbol{w}}^k)\right\|_2^2\right]+\frac{\eta L^2}{2}\sum_{k=0}^{K-1}A1_k+\frac{3\eta^2L^3}{2}H\sum_{k=0}^{K-1}A2_k\\
&+\frac{\eta LK\sqrt{d}\tau}{2}\sum_{n=1}^{N}p_n^2\delta_{w,n}(1+H\delta_{g,n})\\
&+\frac{\eta^2LK\sigma^2/M}{2}\sum_{n=1}^{N}p_n^2(1+H\delta_{g,n})+\frac{3\eta^2LKH^2G^2\bar{\delta}_g}{2}.
\end{aligned}\quad (35)$$

where $\bar{\delta}_g=\sum_{n=1}^{N}p_n^2\delta_{g,n}$, $C_1=1-3\eta LH\bar{\delta}_g$ and $C_2=1+3\eta LH\bar{\delta}_g$. Plugging Lemma 4 into (35), if $C_1'=C_1-\frac{3\eta^2H^2(L^2+3\eta L^3\bar{\delta}_gH)}{1-3\eta^2L^2H^2}\geq 0$,

$$\begin{aligned}
&\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\left\|\nabla F(\bar{\boldsymbol{w}}^k)\right\|_2^2\right]\\
&\leq\frac{2\mathbb{E}\left[F(\bar{\boldsymbol{w}}^{0})-F(\bar{\boldsymbol{w}}^{K})\right]}{\eta C_1'K}+\frac{\eta^2L^2C_2H\sigma^2}{C_1'M(1-3\eta^2L^2H^2)}+\frac{3\eta LKH}{C_1'}G^2\\
&+\frac{3\eta^2L^2C_2H^2G^2}{C_1'(1-3\eta^2L^2H^2)}+\frac{\eta L^2H\sqrt{d}\tau}{C_1'(1-3\eta^2L^2H^2)}\sum_{n=1}^{N}p_n\delta_{w,n}C_{3,n}\\
&+\frac{\eta L\sigma^2}{C_1'M}\sum_{n=1}^{N}p_n^2(1+H\delta_{g,n})+\frac{L\sqrt{d}\tau}{C_1'}\sum_{n=1}^{N}p_n^2\delta_{w,n}(1+H\delta_{g,n}),
\end{aligned}\quad (36)$$

where $C_{3,n} = 1 + 3\eta LHp_n\delta_{g,n}$. If we set $\eta = \sqrt{MN/K}$ and $\eta LH\bar{\delta}_g \geq \frac{\eta^2 H^2(L^2+3\eta L^3\bar{\delta}_g H)}{1-3\eta^2L^2H^2}$, we can get the $1/C_1' \leq 2$. Thus,

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\left\|\nabla F(\bar{\boldsymbol{w}}^k)\right\|_2^2\right]$$
$$\leq \frac{4\mathbb{E}\left[F(\bar{\boldsymbol{w}}^0) - F(\bar{\boldsymbol{w}}^K)\right]}{\sqrt{MNK}} + \frac{2LH\sigma^2}{\sqrt{MNK}}(2\bar{\delta}_g + p) + \frac{12MLH\bar{\delta}_g G^2}{\sqrt{MNK}}$$
$$+ 2LH\sqrt{d}\tau\sum_{n=1}^{N} p_n^2\delta_{w,n}(1 + 2\delta_{g,n}). \tag{37}$$