APPENDIX A
PROOF OF LEMMA 1

In the Lemma 1, our objective is to establish the validity of the following inequality.

$$(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*) \geq (\mathbf{G}^{t,0})^\top(\Theta^{t,0}-\Theta^*) - 2\eta\beta D\epsilon$$

Next, we detail the specific steps of the proof. Initially, drawing from Assumption 4, we derive the following:

$$\left|\Theta_{k,d}^{t,0}-\Theta_{k,d}^*\right| \leq 2\beta \tag{15}$$

By definition of $\mathbf{G}_{k,d}^{t,\tau}$ and Assumption 3, we get the following inequality:

$$\left|\mathbf{G}_{k,d}^{t,\tau}-\mathbf{G}_{k,d}^{t,0}\right| \leq \left\|\mathbf{G}_{k}^{t,\tau}-\mathbf{G}_{k}^{t,0}\right\| \tag{16}$$

$$\leq \epsilon\left\|\theta_k^{t,\tau}-\theta_k^{t,0}\right\| \tag{17}$$

$$\leq \eta\epsilon\left\|\sum_{\tau=0}^{E-1}\nabla_k F_t(\Phi_k^{t,\tau})\right\| \tag{18}$$

$$\leq \eta\epsilon EL \tag{19}$$

Where the last inequality is due to the Assumption 2. By integrating the aforementioned inequality, we can deduce the following:

$$\left|(\mathbf{G}_{k,d}^{t,\tau}-\mathbf{G}_{k,d}^{t,0})(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*)\right| \leq 2\eta\beta\epsilon EL \tag{20}$$

Where Eq. (20) is hold due to the fact $|AB| \leq |A|\,|B|$. Then we can get the lower bound of $\mathbf{G}_{k,d}^{t,\tau}(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*)$ as:

$$\mathbf{G}_{k,d}^{t,\tau}(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*) \geq \mathbf{G}_{k,d}^{t,0}(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*) - 2\eta\beta\epsilon EL \tag{21}$$

Finally, we can get the lower bound of $(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)$:

$$(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*) = \sum_{d=1}^{D}\mathbf{G}_{k,d}^{t,\tau}(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*) \tag{22}$$

$$\geq \sum_{d=1}^{D}\mathbf{G}_{k,d}^{t,0}(\Theta_{k,d}^{t,0}-\Theta_{k,d}^*) - 2\eta\beta D\epsilon EL \tag{23}$$

$$\geq (\mathbf{G}^{t,0})^\top(\Theta^{t,0}-\Theta^*) - 2\eta\beta D\epsilon EL \tag{24}$$

The end of the proof of Lemma 1.

APPENDIX B
PROOF OF THEOREM 1

In this section, we will present the detailed proof procedure for Theorem 1, which addresses the scenario of OVFL without the application of quantization methods. According to the overall model update rule of OVFL, we can get $\Theta^{t+1,0} = \Theta^{t,0} - \eta\sum_{\tau=0}^{E-1}\mathbf{G}^{t,\tau}$. Then we obtain the subsequent inequality:

$$\begin{aligned}
&\left[\left\|\Theta^{t+1,0}-\Theta^*\right\|^2\right] \\
&= \left\|\Theta^{t,0}-\Theta^*\right\|^2 + \eta^2\left[\left\|\sum_{\tau=0}^{E-1}\mathbf{G}^{t,\tau}\right\|^2\right] \\
&-2\eta\left[\sum_{\tau=0}^{E-1}(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)\right]
\end{aligned} \tag{25}$$

Then we can further get the bound of $\left\|\sum_{\tau=0}^{E-1}\mathbf{G}^{t,\tau}\right\|^2$ as:

$$\left\|\sum_{\tau=0}^{E-1}\mathbf{G}^{t,\tau}\right\|^2 \leq E\sum_{\tau=0}^{E-1}\left\|\mathbf{G}^{t,\tau}\right\|^2 \tag{26}$$

Where Eq. (26) is due to triangle inequality. Plugging Eq. (26) into Eq. (25), we can obtain the upper-bound as:

$$\begin{aligned}
&\left[\left\|\Theta^{t+1,0}-\Theta^*\right\|^2\right] \\
&\leq \left\|\Theta^{t,0}-\Theta^*\right\|^2 - 2\eta\left[\sum_{\tau=0}^{E-1}(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)\right] \\
&+\eta^2 E\sum_{\tau=0}^{E-1}\left\|\mathbf{G}^{t,\tau}\right\|^2
\end{aligned} \tag{27}$$

Then we can further obtain the following upper-bound:

$$\sum_{\tau=0}^{E-1}(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*) \tag{28}$$

$$\begin{aligned}
&\leq \frac{1}{2\eta}\left(\left\|\Theta^{t,0}-\Theta^*\right\|^2 - \left[\left\|\Theta^{t+1,0}-\Theta^*\right\|^2\right]\right) \\
&+\frac{\eta E}{2}\sum_{\tau=0}^{E-1}\left\|\mathbf{G}^{t,\tau}\right\|^2
\end{aligned} \tag{29}$$

Subsequently, it is necessary to establish the lower bound of Eq. (28). Then with the convexity of loss function in Assumption 1, we can get the lower-bound as:

$$(\mathbf{G}^{t,0})^\top(\Theta^{t,0}-\Theta^*) \geq \left[F_t(\Theta^{t,0}) - F_t(\Theta^*)\right] \tag{30}$$

Then with the help of Lemma 1, we can further get:

$$\begin{aligned}
&(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*) \\
&\geq (\mathbf{G}^{t,0})^\top(\Theta^{t,0}-\Theta^*) - 2\eta\beta D\epsilon EL
\end{aligned} \tag{31}$$

We derive the lower-bound of Eq.(28) by integrating Eq.(30) and Eq. (31), as follows:

$$\begin{aligned}
&\sum_{\tau=0}^{E-1}(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*) \\
&\geq E\left[F_t(\Theta^{t,0}) - F_t(\Theta^*)\right] - 2\eta\beta D\epsilon E^2 L
\end{aligned} \tag{32}$$

Utilizing $\sum_{\tau=0}^{E-1}(\mathbf{G}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)$ as a transitional term and combining both the upper and lower bounds, then taking the expectation $\mathbb{E}_t$ respect to t on both sides and summing over $t = 1, 2, \ldots, T$, we derive the subsequent regret bound:

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{E}_t \left[F_t(\Theta^{t,0})\right] - \sum_{t=1}^{T} F_t(\Theta^*) \\
\leq& \frac{1}{2\eta E} \sum_{t=1}^{T} (\mathbb{E}_t \|\Theta^{t,0} - \Theta^*\|^2 - \mathbb{E}_{t+1} \left[\left\|\Theta^{t+1,0} - \Theta^*\right\|^2\right]) \\
&+ \frac{\eta}{2} \sum_{t=1}^{T} \sum_{\tau=0}^{E-1} \mathbb{E}_t \left\|\mathbf{G}^{t,\tau}\right\|^2 + 2T\eta\beta D\epsilon EL \qquad (33) \\
\overset{(a)}{\leq}& \frac{1}{2\eta E} \left\|\Theta^{1,0} - \Theta^*\right\|^2 + 2T\eta\beta D\epsilon EL \\
&+ \frac{\eta}{2} \sum_{t=1}^{T} \sum_{\tau=0}^{E-1} \left\| \left[(\nabla_0 F_t(\Phi_0^{t,\tau}))^\top, \ldots, (\nabla_K F_t(\Phi_K^{t,\tau}))^\top\right]^\top \right\|^2 \qquad (34) \\
\overset{(b)}{\leq}& \frac{\left\|\Theta^{1,0} - \Theta^*\right\|^2}{2\eta E} + \frac{\eta TEKL^2}{2} + 2\eta T\beta D\epsilon EL \qquad (35)
\end{aligned}$$

Where (a) results from the summation eliminating the intermediate term, and (b) holds when Assumption 2 is applied.

Appendix C
Proof of Lemma 2

In the Lemma 2, our objective is to establish the validity of the following inequality.

$$(\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*) \geq (\mathbf{G}^{t,0})^\top (\Theta^{t,0} - \Theta^*) - 2\beta D(\eta\epsilon EL + \rho)$$

Next, we detail the specific steps of the proof. Firstly, according to Assumption 4, we can further obtain:

$$\left|\Theta_{k,d}^{t,0} - \Theta_{k,d}^*\right| \leq 2\beta \qquad (36)$$

By integrating the aforementioned inequality with Assumption 3, we can deduce the following:

$$\begin{aligned}
&\left|(\hat{\mathbf{G}}_{k,d}^{t,\tau} - \mathbf{G}_{k,d}^{t,0})(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*)\right| \\
=& \left|(\hat{\mathbf{G}}_{k,d}^{t,\tau} - \mathbf{G}_{k,d}^{t,\tau} + \mathbf{G}_{k,d}^{t,\tau} - \mathbf{G}_{k,d}^{t,0})(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*)\right| \qquad (37) \\
\leq& \left|(\hat{\mathbf{G}}_{k,d}^{t,\tau} - \mathbf{G}_{k,d}^{t,\tau})(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*)\right| \\
&+ \left|(\mathbf{G}_{k,d}^{t,\tau} - \mathbf{G}_{k,d}^{t,0})(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*)\right| \qquad (38) \\
\leq& 2\eta\beta\epsilon EL + 2\beta\rho \qquad (39)
\end{aligned}$$

The derivation of Eq. (39) is partly based on the previous results presented in Lemma 1 and partly on the stipulations of Assumption 5. Then we can get the lower bound as:

$$\begin{aligned}
&\hat{\mathbf{G}}_{k,d}^{t,\tau}(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*) \\
&\geq \mathbf{G}_{k,d}^{t,0}(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*) - 2\beta(\eta\epsilon EL + \rho) \qquad (40)
\end{aligned}$$

Finally, we can get the lower bound of $(\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*)$:

$$\begin{aligned}
(\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*) &= \sum_{d=1}^{D} \hat{\mathbf{G}}_{k,d}^{t,\tau}(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*) \qquad (41) \\
&\geq \sum_{d=1}^{D} \mathbf{G}_{k,d}^{t,0}(\Theta_{k,d}^{t,0} - \Theta_{k,d}^*) - 2\beta D(\eta\epsilon EL + \rho) \qquad (42) \\
&\geq (\mathbf{G}^{t,0})^\top (\Theta^{t,0} - \Theta^*) - 2\beta D(\eta\epsilon EL + \rho) \qquad (43)
\end{aligned}$$

Here we finish the proof of Lemma 2.

Appendix D
Proof of Theorem 2

In this section, we will present the detailed proof procedure for Theorem 2, which addresses the scenario of OVFL with the application of quantization methods. According to the overall model update rule of OVFL, we can get $\Theta^{t+1,0} = \Theta^{t,0} - \eta \sum_{\tau=0}^{E-1} \hat{\mathbf{G}}^{t,\tau}$. Then we obtain the subsequent inequality:

$$\begin{aligned}
&\left[\left\|\Theta^{t+1,0} - \Theta^*\right\|^2\right] \\
&= \left\|\Theta^{t,0} - \Theta^*\right\|^2 + \eta^2 \left[\left\|\sum_{\tau=0}^{E-1} \hat{\mathbf{G}}^{t,\tau}\right\|^2\right] \\
&-2\eta \left[\sum_{\tau=0}^{E-1} (\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*)\right] \qquad (44)
\end{aligned}$$

Then we can further get the bound of $\left\|\sum_{\tau=0}^{E-1} \hat{\mathbf{G}}^{t,\tau}\right\|^2$ as:

$$\begin{aligned}
\left\|\sum_{\tau=0}^{E-1} \hat{\mathbf{G}}^{t,\tau}\right\|^2 &\leq E \sum_{\tau=0}^{E-1} \left\|\hat{\mathbf{G}}^{t,\tau}\right\|^2 \qquad (45) \\
&\leq 2E \sum_{\tau=0}^{E-1} \left\|\hat{\mathbf{G}}^{t,\tau} - \mathbf{G}^{t,\tau}\right\|^2 + 2E \sum_{\tau=0}^{E-1} \left\|\mathbf{G}^{t,\tau}\right\|^2 \qquad (46)
\end{aligned}$$

Where Eq. (45) is due to triangle inequality, and Eq. (46) is due to the face $\|A + B\|^2 \leq 2\|A\|^2 + 2\|B\|^2$. Plugging Eq. (46) into Eq. (44), we can obtain the upper-bound as:

$$\begin{aligned}
&\left[\left\|\Theta^{t+1,0} - \Theta^*\right\|^2\right] \\
&\leq \left\|\Theta^{t,0} - \Theta^*\right\|^2 - 2\eta \left[\sum_{\tau=0}^{E-1} (\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*)\right] \\
&+ 2\eta^2 E \sum_{\tau=0}^{E-1} \left\|\hat{\mathbf{G}}^{t,\tau} - \mathbf{G}^{t,\tau}\right\|^2 + 2\eta^2 E \sum_{\tau=0}^{E-1} \left\|\mathbf{G}^{t,\tau}\right\|^2 \qquad (47)
\end{aligned}$$

Then we can further obtain the following upper-bound:

$$\begin{aligned}
&\sum_{\tau=0}^{E-1} (\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*) \qquad (48) \\
\leq& \frac{1}{2\eta} (\left\|\Theta^{t,0} - \Theta^*\right\|^2 - \left[\left\|\Theta^{t+1,0} - \Theta^*\right\|^2\right]) \\
&+ \eta E \sum_{\tau=0}^{E-1} \left\|\hat{\mathbf{G}}^{t,\tau} - \mathbf{G}^{t,\tau}\right\|^2 + \eta E \sum_{\tau=0}^{E-1} \left\|\mathbf{G}^{t,\tau}\right\|^2 \qquad (49)
\end{aligned}$$

Subsequently, it is necessary to establish the lower bound of Eq. (48). Then with the Assumption 1, we can get the lower-bound with the convexity of the loss function as:

$$(\mathbf{G}^{t,0})^\top (\Theta^{t,0} - \Theta^*) \geq \left[F_t(\Theta^{t,0}) - F_t(\Theta^*)\right] \qquad (50)$$

Then based on Lemma 2, we can further get:

$$\begin{aligned}
&(\hat{\mathbf{G}}^{t,\tau})^\top (\Theta^{t,0} - \Theta^*) \\
&\geq (\mathbf{G}^{t,0})^\top (\Theta^{t,0} - \Theta^*) - 2\beta D(\eta\epsilon EL + \rho) \qquad (51)
\end{aligned}$$

Then we can get the lower-bound of Eq. (48) as:

$$\begin{aligned}&\sum_{\tau=0}^{E-1}(\hat{\mathbf{G}}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)\\&\geq E\left[F_t(\Theta^{t,0})-F_t(\Theta^*)\right]-2E\beta D(\eta\epsilon EL+\rho)\end{aligned}\tag{52}$$

Utilizing $\sum_{\tau=0}^{E-1}(\hat{\mathbf{G}}^{t,\tau})^\top(\Theta^{t,0}-\Theta^*)$ as a transitional term and combining both the upper and lower bounds, then taking the expectation $\mathbb{E}_t$ respect to t on both sides and summing over $t=1,2,\ldots,T$, we derive the subsequent regret bound:

$$\begin{aligned}&\sum_{t=1}^{T}\mathbb{E}_t\left[F_t(\Theta^{t,0})\right]-\sum_{t=1}^{T}F_t(\Theta^*)\\
\leq&\frac{1}{2\eta E}\sum_{t=1}^{T}(\mathbb{E}_t\|\Theta^{t,0}-\Theta^*\|^2-\mathbb{E}_{t+1}\left[\left\|\Theta^{t+1,0}-\Theta^*\right\|^2\right])\\
&+\eta\sum_{t=1}^{T}\sum_{\tau=0}^{E-1}\left\|\hat{\mathbf{G}}^{t,\tau}-\mathbf{G}^{t,\tau}\right\|^2+\eta\sum_{t=1}^{T}\sum_{\tau=0}^{E-1}\left\|\mathbf{G}^{t,\tau}\right\|^2\\
&+2T\beta D\eta\epsilon EL+2T\beta D\rho\end{aligned}\tag{53}$$

$$\begin{aligned}\overset{(c)}{\leq}&\frac{1}{2\eta E}\left\|\Theta^{1,0}-\Theta^*\right\|^2+\eta TED\rho^2\\
&+\eta\sum_{t=1}^{T}\sum_{\tau=0}^{E-1}\left\|\left[(\nabla_0F_t(\Phi_0^{t,\tau}))^\top,\ldots,(\nabla_KF_t(\Phi_K^{t,\tau}))^\top\right]^\top\right\|^2\\
&+2T\beta D\eta\epsilon EL+2T\beta D\rho\end{aligned}\tag{54}$$

$$\begin{aligned}\overset{(d)}{\leq}&\frac{\left\|\Theta^{1,0}-\Theta^*\right\|^2}{2\eta E}+\eta TEKL^2+\eta TED\rho^2\\
&+2\eta T\beta D\epsilon EL+2T\beta D\rho\end{aligned}\tag{55}$$

Where (c) results from the summation eliminating the intermediate term and the application of Assumption 5. (d) holds when Assumption 2 is applied.