Supplemental Material for

Asynchronous Personalized Federated Learning with Irregular Clients

Editors: Emtiyaz Khan and Mehmet Gonen

1. Complete Proofs

1.1. Proof of Theorem 1

Suppose the virtual sequence of global models $\omega^\tau$ for $\tau = 1, 2, ..., T$ is computed by

$$\omega^\tau = \omega^0 - \sum_{l=1}^{M} p^l \sum_{i=0}^{\tau-1} \eta_i \nabla F_k(\omega^i_k, \xi^i_k). \quad (9)$$

We define $g^\tau = \sum_{k=1}^{M} p^k \nabla F_k(\omega^\tau_k, \xi^\tau_k)$ and $\overline{g}^\tau = \sum_{k=1}^{M} p^k \nabla F_k(\omega^* _k)$, and we have $E[g^\tau] = \overline{g}^\tau$.

Consequently, based on Equation 2, we have $\omega^{\tau+1} = \omega^\tau - \eta \overline{g}^\tau$. Then it holds that

$$||\omega^{\tau+1} - \omega^*||^2 = ||\omega^\tau - \eta \overline{g}^\tau - \omega^*||^2$$

$$= ||\omega^\tau - \eta \overline{g}^\tau - \omega^* - \eta \overline{g}^\tau + \eta \overline{g}^\tau||^2$$

$$= ||\omega^\tau - \omega^* - \eta \overline{g}^\tau + \eta (\overline{g}^\tau - \overline{g}^\tau)||^2$$

$$= \underbrace{||\omega^\tau - \omega^*||^2}_{A_1} + \underbrace{\eta^2 ||\overline{g}^\tau - \overline{g}^\tau||^2}_{A_2} + \underbrace{2 \eta < \omega^\tau - \omega^*, \overline{g}^\tau - \overline{g}^\tau >}_{A_3}. \quad (10)$$

To bound Equation (10), we separately bound $A_1$, $A_2$ and $A_3$.

$$A_1 = ||\omega^\tau - \omega^*||^2 + \eta^2 ||\overline{g}^\tau||^2 - 2 \eta < \omega^\tau - \omega^*, \overline{g}^\tau >. \quad (11)$$

Given Assumption 1, for any $\omega$ and $\omega'$, we have

$$F(\omega) = F(\omega') + \int_0^1 dF(\omega' + t(\omega - \omega')) \frac{dt}{dt}$$

$$= F(\omega') + \int_0^1 \nabla F(\omega' + t(\omega - \omega'))^\top (\omega - \omega') dt$$

$$= F(\omega') + \nabla F(\omega')^\top (\omega - \omega') + \int_0^1 [\nabla F(\omega' + t(\omega - \omega')) - \nabla F(\omega')]^\top (\omega - \omega') dt \quad (12)$$

$$\leq F(\omega') + \nabla F(\omega')^\top (\omega - \omega') + \int_0^1 \frac{L}{2} ||t(\omega - \omega')|| ||\omega - \omega'|| dt$$

$$= F(\omega') + \nabla F(\omega')^\top (\omega - \omega') + \frac{1}{2} L ||\omega - \omega'||^2,$$
\[ |\nabla F_k(\omega_k^\tau)|^2 \leq 2L(F_k(\omega_k^\tau) - F_k(\omega^*)) \]  \tag{13}

Thus, \( B_1 \) is bounded by
\[ B_1 \leq \eta_r^2 \sum_{k=1}^{M} p^k |\nabla F_k(\omega_k^\tau)|^2 \leq 2L\eta_r^2 \sum_{k=1}^{M} p^k (F_k(\omega_k^\tau) - F_k(\omega^*)) \]  \tag{14}

\( B_2 \) can be reformulated by
\[ B_2 = -2\eta_r \sum_{k=1}^{M} p^k < \omega^\tau - \omega^*, g_k^\tau > \]  \tag{15}
\[ = -2\eta_r \sum_{k=1}^{M} p^k < \omega^\tau - \omega_k^\tau, g_k^\tau > -2\eta_r \sum_{k=1}^{M} p^k < \omega_k^\tau - \omega^*, g_k^\tau >, \]
and we have
\[ -2\mathbb{E}[< \omega^\tau - \omega_k^\tau, g_k^\tau >] = -2 < \omega^\tau - \omega_k^\tau, \nabla F_k(\omega_k^\tau) > \]
\[ \leq 2||\omega^\tau - \omega_k^\tau||||\nabla F_k(\omega_k^\tau)|| \]
\[ \leq \frac{1}{\eta_r} ||\omega^\tau - \omega_k^\tau||^2 + \eta_r ||\nabla F_k(\omega_k^\tau)||^2. \]  \tag{16}

The first inequality holds for Cauchy-Schwarz inequality and the second holds for AM-GM inequality. According to Equations (13) and (16), we have
\[ -2\mathbb{E}[< \omega^\tau - \omega_k^\tau, g_k^\tau >] \leq \frac{1}{\eta_r} ||\omega^\tau - \omega_k^\tau||^2 + 2L\eta_r (F_k(\omega_k^\tau) - F_k(\omega^*)) \]  \tag{17}

and
\[ -2\mathbb{E}[< \omega_k^\tau - \omega^*, g_k^\tau >] \leq -2(F_k(\omega_k^\tau) - F_k(\omega^*)) - \mu ||\omega_k^\tau - \omega^*||^2. \]  \tag{18}

Thus, \( B_2 \) can be bounded by
\[ \mathbb{E}[B_2] \leq \sum_{k=1}^{M} p^k ||\omega^\tau - \omega_k^\tau||^2 + 2L\eta_r^2 (F_k(\omega_k^\tau) - F_k(\omega^*)) - \eta_r \sum_{k=1}^{M} p^k (2(F_k(\omega_k^\tau) - F_k(\omega^*)) + \mu ||\omega_k^\tau - \omega^*||^2). \]  \tag{19}

According to the convexity property of quadratic function, we have
\[ -\eta_r \mu \sum_{k=1}^{M} p^k ||\omega_k^\tau - \omega^*||^2 \leq -\eta_r \mu ||\omega^\tau - \omega^*||^2. \]  \tag{20}

Substituting Equation (20) into Equation (19), we bound \( B_2 \) by
\[ \mathbb{E}[B_2] \leq \sum_{k=1}^{M} p^k ||\omega^\tau - \omega_k^\tau||^2 + 2L\eta_r^2 (F_k(\omega_k^\tau) - F_k(\omega^*)) - 2\eta_r \sum_{k=1}^{M} p^k (F_k(\omega_k^\tau) - F_k(\omega^*)) - \eta_r \mu ||\omega^\tau - \omega^*||^2. \]  \tag{21}
Based on Equations (14) and (21), $A_1$ can be reformulated as

$$
E[A_1] \leq (1 - \eta_r \mu)||\omega^* - \omega^\tau||^2 + \sum_{k=1}^{M} p^k||\omega^\tau - \omega^\tau_k||^2 \\
+ 4L\eta^2 \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F^\tau_k) - 2\eta_r \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F_k(\omega^*)) + 4L\eta^2 \Gamma.
$$

Let $\gamma = 2\eta_r(1 - 2L\eta_r)$ for $\tau = 1, 2, ..., T$. $C$ can be reformulated as

$$
C = -\gamma \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F(\omega^*)) + (2\eta_r - \gamma) \sum_{k=1}^{M} p^k(F^* - F_k^*) \\
= -\gamma \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F^*) + 4L\eta^2 \Gamma,
$$

where

$$
D = \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F(\omega^*)) + \sum_{k=1}^{M} p^k(F_k(\omega^\tau) - F^*) \\
\geq \sum_{k=1}^{M} p^k < \nabla F_k(\omega^\tau), \omega^\tau_k - \omega^\tau > + F(\omega^\tau) - F^*.
$$

Given Assumption 1, we have

$$
D_1 \geq -\frac{1}{2} \sum_{k=1}^{M} p^k(\eta_r ||\nabla F_k(\omega^\tau)||^2 + \frac{1}{\eta_r} ||\omega^\tau_k - \omega^\tau||^2).
$$

Substituting Equations (25) and (24) into (23), we have

$$
C = \gamma (Ln_r - 1) \sum_{k=1}^{M} p^k(F_k(\omega^\tau_k) - F_k(\omega^*)) + (4L\eta^2 + \gamma \eta_r L) \Gamma + \frac{\gamma}{2\eta_r} \sum_{k=1}^{M} p^k||\omega^\tau - \omega^\tau_k||^2 \\
\leq 6L\eta^2 \Gamma + \sum_{k=1}^{M} p^k||\omega^\tau - \omega^\tau_k||^2.
$$

According to Equations (22) and (26), $A_1$ can be bounded by

$$
E[A_1] \leq (1 - \mu \eta_r)||\omega^* - \omega^\tau||^2 + 2 \sum_{k=1}^{M} p^k||\omega^\tau - \omega^\tau_k||^2 + 6L\eta^2 \Gamma.
$$
To bound $D_2$, we define the auxiliary inequality as
\[
g_k^{\tau+1} = ||g_k^{\tau+1} - g_k^\tau + \bar{g}_k|| \leq ||g_k^{\tau+1} - g_k^\tau|| + ||\bar{g}_k||
\]
\[
= ||\nabla F_k(\omega_k^{\tau} - \eta \tau g_k^{\tau}) - \nabla F_k(\omega_k^{\tau})|| + ||\bar{g}_k||
\]
\[
\leq L||\omega_k^{\tau} - \eta \tau g_k^{\tau} - \omega_k^{\tau}|| + ||\bar{g}_k||
\]
\[
= (1 + L\eta \tau)||\bar{g}_k|| = (1 + L\eta \tau)^{\tau}||\bar{g}_k^0||,
\]
and $\tau_{k_0}$ as the last time step when $k$-th client upload its model to the server. Given the fixed $\eta$, we have
\[
||\bar{\omega}^\tau - \omega_k^{\tau}|| = ||(\omega_k^{\tau} - \omega^{\tau_{k_0}}) - (\bar{\omega}^\tau - \omega^{\tau_{k_0}})|| \leq \sum_{i=\tau_{k_0}}^{\tau} ||\eta (g_k^i - \sum_{j=1}^{M} p^i g_j^i)||
\]
\[
\leq \sum_{i=\tau_{k_0}}^{\tau} (1 + L\eta)^{i-\tau_{k_0}} ||\eta (g_k^{\tau_{k_0}} - \sum_{j=1}^{M} g_j^{\tau_{k_0}})||
\]
\[
\leq ||\sum_{i=\tau_{k_0}}^{\tau} \frac{(1 + L\eta)^{i-\tau_{k_0}} - 1}{L} (g_k^{\tau_{k_0}} - \sum_{j=1}^{M} g_j^{\tau_{k_0}})||.
\]
Equation (29) can be reformulated as
\[
||\bar{\omega}^\tau - \omega_k^{\tau}|| \leq \frac{\sum_{i=\tau_{k_0}}^{\tau} (1 + L\eta)^{i-\tau_{k_0}}}{L} (g_k^{\tau_{k_0}} - \sum_{j=1}^{M} g_j^{\tau_{k_0}})||. \tag{30}
\]
Given Assumption 4, we have
\[
||\bar{\omega}^\tau - \omega_k^{\tau}|| \leq \frac{\sum_{i=\tau_{k_0}}^{\tau} (1 + L\eta)^{i-\tau_{k_0}} - 1}{L} (g_k^{\tau_{k_0}} - \sum_{j=1}^{M} g_j^{\tau_{k_0}})||, \tag{31}
\]
and according to Taylor expansion,
\[
||\bar{\omega}^\tau - \omega_k^{\tau}|| \leq \eta^2 (\pi_k E - 1)^2 \chi^2 \leq \eta^2 E (\pi_k^2 E - \pi_k) \chi^2. \tag{32}
\]
Thus, $D_2$ can be bounded by
\[
D_2 \leq 2\eta^2 E \sum_{k=1}^{M} \sigma_k^2 (\pi_k^2 E - \pi_k) \chi^2. \tag{33}
\]
Substituting Equation (33) into (27), $A_1$ can be bounded by
\[
\mathbb{E}[A_1] \leq (1 - \mu \eta \tau)||\bar{\omega}^\tau - \omega^*||^2 + 2\eta^2 E \sum_{k=1}^{M} \sigma_k^2 (\pi_k^2 E - \pi_k) \chi^2 + 6L\eta \Gamma. \tag{34}
\]
Then, $A_2 = \eta^2 ||g^\tau - \bar{g}^\tau||^2$ in Equation (10) can be bounded by
\[
\mathbb{E}[||g^\tau - \bar{g}^\tau||^2] = \mathbb{E}[||\sum_{k=1}^{M} p^k (\nabla F_k(\omega_k^\tau) - g_k^\tau)||^2]
\]
\[
\leq \sum_{k=1}^{M} p^{2k} \mathbb{E}[||\nabla F_k(\omega_k^\tau) - g_k^\tau||^2] \leq \sum_{k=1}^{M} p^{2k} \sigma_k^2, \tag{35}
\]
and $A_3 = 2\eta_\tau < \omega^\tau - \omega^* - \eta_\tau \bar{g}^\tau, \bar{g}^\tau - g^\tau >$ can be bounded by $E[A_3] = 0$ for $E[g^\tau] = \bar{g}^\tau$.

Let $\Delta_{\tau+1} = E[||\omega^\tau_{\tau+1} - \omega^*||^2]$ and combining the obtained bounds of $A_1$, $A_2$ and $A_3$, we have
\[
\Delta_{\tau+1} \leq (1 - \mu \eta_\tau)\Delta_\tau + \eta_\tau^2 A',
\]
(36)
where $A' = \sum_{k=1}^M p^2 k^2 \sigma_k^2 + 6LE\sum_{k=1}^M p^k (\pi_k^2 E - \pi_k)^2$. Equation (36) can be further reformulated as
\[
\Delta_{\tau+1} - \frac{\eta_\tau A'}{\mu} \leq (1 - \mu \eta_\tau) (\Delta_\tau - \frac{\eta_\tau A'}{\mu}).
\]
(37)
Thus we have
\[
\Delta_{\tau+1} - \frac{\eta_\tau A'}{\mu} \leq (1 - \mu \eta_\tau)^{\tau+1} (\Delta_0 - \frac{\eta_\tau A'}{\mu}).
\]
(38)
1.2. Proof of Theorem 2
The proof of Theorem 2 follows from the proof of Theorem 1. Starting from inequality (28), using the decayed learning rate in $D^2_2$, we have
\[
||\omega^\tau - \omega^\tau_k|| = ||(\omega^\tau_k - \omega^\tau_{k0}) - (\omega^\tau - \omega^\tau_{k0})|| \leq \sum_{i=\tau_{k0}}^\tau ||\eta_\tau (g^i_k - \sum_{j=1}^M g^i_j)|| \leq \left(1 + \frac{L \eta_\tau}{\mu}\right)^{\tau - \tau_{k0}} E \left(\pi_{k0} - \sum_{j=1}^M \pi_{j0}\right) \leq \left(1 + \frac{L \eta_\tau}{\mu}\right)^{\tau - \tau_{k0}} E \left(\pi_{k0} - \sum_{j=1}^M \pi_{j0}\right).\]
(40)
Given Assumption 4, we have
\[
||\omega^\tau - \omega^\tau_k||^2 \leq \left(1 + \frac{L \eta_\tau}{\mu}\right)^{\tau - \tau_{k0}} E \left(\pi_{k0} - \sum_{j=1}^M \pi_{j0}\right)^2 \leq \frac{(1 + L \eta_\tau)^{\tau - \tau_{k0}} - 1}{L} E \left(\pi_{k0} - \sum_{j=1}^M \pi_{j0}\right)^2 \chi^2.
\]
(41)
According to the Taylor expansion, Equation (41) can be reformulated as
\[
||\omega^\tau - \omega^\tau_k||^2 \leq 4\eta^2_\tau (\pi_{k0} E - 1)^2 \chi^2 \leq 4\eta^2_\tau E(\pi_{k0}^2 E - \pi_{k}) \chi^2.
\]
(42)
Thus $D_2$ can be bounded by
\[
D_2 \leq 8\eta^2_\tau E \sum_{k=1}^M p^k (\pi_{k0}^2 E - \pi_{k}) \chi^2.
\]
(43)
Substituting the bound of $D_2$ into Equation (34), we have
\[
\Delta_{\tau+1} \leq (1 - \mu \eta_\tau)\Delta_\tau + \eta_\tau^2 A'.
\]
(44)
1.3. Proof of Theorem 3

Suppose the error bound of AsyPFL satisfies
\[ \mathbb{E}[\omega^{T_{\text{min}}}] - F^* \leq \frac{2L}{\mu(\gamma + T_{\text{min}})^{2/3}} \left( \frac{A}{\mu} + 2L||\omega^0 - \omega^*|| \right) \leq \epsilon. \] (45)

Equation (45) can be reformulated as
\[ T_{\text{min}} + \gamma \geq \frac{2L}{\mu \epsilon} \left( \frac{A}{\mu} + 2L||\omega^0 - \omega^*|| \right). \] (46)

which is equivalent to
\[ T_{\text{min}} \geq \frac{2L}{\mu \epsilon} \left( \sum_{k=1}^{M} p^k \frac{2k \sigma_k^2}{\mu} + 6L \Gamma \right) + 2L||\omega^0 - \omega^*|| + \frac{16L}{\mu^2 \epsilon} \sum_{k=1}^{M} p^k (\pi_k^2 E - \pi_k) \chi^2 - \gamma. \] (47)

Thus, the required communication rounds are
\[ R_{\text{min}}(E) = \frac{T_{\text{min}}}{E} \geq \frac{2L}{E \mu \epsilon} \left( \sum_{k=1}^{M} p^k \frac{2k \sigma_k^2}{\mu} + 6L \Gamma \right) + 2L||\omega^0 - \omega^*|| + \frac{16L}{\mu^2 \epsilon} \sum_{k=1}^{M} p^k (\pi_k^2 E - \pi_k) \chi^2 - \gamma. \] (48)

According to Equation (48), we define the required communication rounds as a function of local epochs \( E \) by
\[ R(E) = \left( \frac{16L}{\epsilon \mu^2} \chi^2 \sum_{k=1}^{M} p^k \pi_k^2 E + \frac{4L^2}{\epsilon \mu} ||\omega^0 - \omega^*|| \right) E \left( \frac{16L}{\epsilon \mu^2} \sum_{k=1}^{M} p^k \pi_k \chi^2 + \gamma \right), \] (49)

which is in the form of
\[ R(E) = aE + \frac{b}{E} - c. \] (50)

Taking the derivative w.r.t. \( E \) and let the resulting equation equal to zero, we have
\[ \frac{dR(E)}{dE} = a - \frac{b}{E^2} = 0, \] (51)

and \( E = \sqrt{\frac{b}{a}}. \)

1.4. Proof of Theorem 4

\( \Delta T \) is minimized if \( C_{\pi_k} \) achieves its minimum. Thus, according to the definition of \( C_{\pi_k} \) and equation (47), we have
\[ C_{\pi_k} = \Delta T_k^2 Z_k + \frac{16L}{\epsilon \mu^2} \Delta T_k^2 \chi^2 p^k (E \pi_k)^2 + \Delta T_k^2 \frac{Z_k}{E \pi_k} - \frac{16L}{\epsilon \mu^2} \Delta T_k^2 \chi^2 p^k. \] (52)

Taking the derivative w.r.t. \( \pi_k \) and let the resulting equation equal to zero, we have
\[ 2 \frac{\Delta T_k^2}{\Delta T_k} \frac{16L}{\epsilon \mu^2} E^2 \chi^2 p^k \pi_k = \frac{Z_k}{E \pi_k^2}. \] (53)

Rearranging Equation (53) to obtain the result in Theorem 4.