Proof of Theorem 1

Let us introduce useful notations:

$$\tilde{z}^{k+1} = z^{k+1} + \frac{\gamma}{n}\sum_{i=1}^{n} e_i^{k+1}$$

With this notation,

$$\begin{aligned}
\tilde{z}^{k+1} = u_H^k &+ \frac{\gamma}{n}\sum_{i=1}^{n} C(e_i^k + F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)) \\
&+ \frac{\gamma}{n}\sum_{i=1}^{n} e_i^k + F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k) \\
&- \frac{\gamma}{n}\sum_{i=1}^{n} C(e_i^k + F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)) \\
= u_H^k &+ \frac{\gamma}{n}\sum_{i=1}^{n} e_i^k + F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k).
\end{aligned}$$

It follow that

$$\begin{aligned}
\tilde{z}^{k+1} - \tilde{z}^k &= u_H^k + \frac{\gamma}{n}\sum_{i=1}^{n} e_i^k + F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k) - z^k - \frac{\gamma}{n}\sum_{i=1}^{n} e_i^k \\
&= u_H^k + \frac{\gamma}{n}\sum_{i=1}^{n} F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k) - z^k.
\end{aligned} \tag{1}$$

In the algorithm, we run Extra Gradient method for H steps for the local problem

$$\begin{aligned}
&\text{Find} \quad \hat{u}^k \quad \text{such that} \quad \langle T(\hat{u}^k), z - \hat{u}^k \rangle \geq 0, \quad \forall z \in \mathcal{Z} \\
\text{with} \quad & T(z) = F_1(z) + F(m^k) - F_1(m^k) + \frac{1}{\gamma}(z - [z^k + \tau(m^k - z^k)]).
\end{aligned} \tag{2}$$

In the proofs, we use the notation $\hat{u}^k$ as the solution of (2). We proceed with the following lemma.

Lemma 1. *Let $\{z^k\}_{k\geq 0}$ be the sequence that Algorithm 1 generates. Then we have the following estimate on iterations:*

$$\begin{aligned}
&\sum_{k=0}^{K-1} q^k \mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^{k+1} - z^*\|^2 \\
&\quad \leq \left(1 - \frac{\gamma\mu}{2}\right)\sum_{k=0}^{K-1} q^k \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \left(1 - \frac{\gamma\mu}{2}\right)\sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - z^*\|^2 \\
&\quad + \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2 + 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega^2\delta^2\right)\sum_{k=0}^{K-1} q^k \mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\
&\quad - \left(\frac{1}{2} - p - \frac{3\gamma\mu}{4}\right)\sum_{k=0}^{K-1} q^k \mathbb{E}\|z^k - \hat{u}^k\|
\end{aligned}$$

$$-\left(p - \frac{3\gamma\mu}{2} - 12\gamma^2\delta^2 - 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega\delta^2\right)\sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - \hat{u}^k\|^2,$$

where q^k are the weights, with $q^k \le q^j \left(1 + \frac{1}{4\omega}\right)^{k-j}$ for all $j < k$.

Proof: Using basic algebraic operators,

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle \tilde{z}^{k+1} - \tilde{z}^k, \tilde{z}^k - z^*\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \tilde{z}^k\|^2 \\ &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle \tilde{z}^{k+1} - \tilde{z}^k, \hat{u}^k - z^*\rangle + 2\mathbb{E}\langle \tilde{z}^{k+1} - \tilde{z}^k, \tilde{z}^k - \hat{u}^k\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \tilde{z}^k\|^2 \\ &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle \tilde{z}^{k+1} - \tilde{z}^k, \hat{u}^k - z^*\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2.\end{aligned}$$

With (1), we can write

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 \\ &\quad + 2\mathbb{E}\langle u_H^k + \gamma \cdot \frac{1}{n}\sum_{i=1}^n F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k) - z^k, \hat{u}^k - z^*\rangle \\ &\quad + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\ &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 \\ &\quad + 2\mathbb{E}\langle u_H^k + \gamma \cdot \left[\frac{1}{n}\sum_{i=1}^n F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)\right] - z^k, \hat{u}^k - z^*\rangle \\ &\quad + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\ &= \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle u_H^k + \gamma \cdot (F(m^k) - F_1(m^k)) - z^k, \hat{u}^k - z^*\rangle \\ &\quad - 2\gamma\mathbb{E}\langle F(u_H^k) - F_1(u_H^k), \hat{u}^k - z^*\rangle \\ &\quad + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2.\end{aligned}$$

Introducing the notation $v^k = z^k + \tau(m^k - z^k) - \gamma \cdot \big(F(m^k) - F_1(m^k)\big)$, we have

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \le& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle u_H^k - v^k, \hat{u}^k - z^*\rangle + \tau\mathbb{E}\langle m^k - z^k, \hat{u}^k - z^*\rangle \\ &- 2\gamma\mathbb{E}\langle F(u_H^k) - F_1(u_H^k), \hat{u}^k - z^*\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\ =& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + 2\mathbb{E}\langle \hat{u}^k - v^k, \hat{u}^k - z^*\rangle + \tau\mathbb{E}\langle m^k - z^k, \hat{u}^k - z^*\rangle \\ &- 2\gamma\mathbb{E}\langle F(u_H^k) - F_1(u_H^k), \hat{u}^k - z^*\rangle + 2\mathbb{E}\langle u_H^k - \hat{u}^k, \hat{u}^k - z^*\rangle \\ &+ \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2.\end{aligned}$$

Since $\hat{u}^k$ is the optimal solution of (2), then $\langle \gamma F_1(\hat{u}^k) + \hat{u}^k - v^k, \hat{u}^k - z\rangle \le 0$. We can use it in $2\mathbb{E}\langle \hat{u}^k - v^k, \hat{u}^k - z^*\rangle$ to get

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \le& \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mathbb{E}\langle F_1(\hat{u}^k), \hat{u}^k - z^*\rangle - 2\gamma\mathbb{E}\langle F(u_H^k) - F_1(u_H^k), \hat{u}^k - z^*\rangle \\ &+ 2\mathbb{E}\langle u_H^k - \hat{u}^k, \hat{u}^k - z^*\rangle + \tau\mathbb{E}\langle m^k - z^k, \hat{u}^k - z^*\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\ =& \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mathbb{E}\langle F_1(\hat{u}^k), \hat{u}^k - z^*\rangle - 2\gamma\mathbb{E}\langle F(\hat{u}^k) - F_1(\hat{u}^k), \hat{u}^k - z^*\rangle \\ &+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^*\rangle \\ &+ \tau\mathbb{E}\langle m^k - z^k, \hat{u}^k - z^*\rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2.\end{aligned}$$

By using the optimality of the solution z^* and μ-strong monotonicity of F (Assumption 2), we get

$$\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \le \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mathbb{E}\langle F(\hat{u}^k) - F(z^*), \hat{u}^k - z^*\rangle$$

$$
\begin{aligned}
&+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^* \rangle \\
&+ \tau \mathbb{E}\langle m^k - z^k, \hat{u}^k - z^* \rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\
\le& \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mu\mathbb{E}\|\hat{u}^k - z^*\|^2 \\
&+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^* \rangle \\
&+ \tau \mathbb{E}\langle m^k - z^k, \hat{u}^k - z^* \rangle + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2.
\end{aligned}
$$

The equality $2\langle a, b\rangle = \|a+b\|^2 - \|a\|^2 - \|b\|^2$ gives

$$
\begin{aligned}
\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \le& \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mu\mathbb{E}\|\hat{u}^k - z^*\|^2 \\
&+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^* \rangle \\
&+ \tau\mathbb{E}\langle m^k - \hat{u}^k, \hat{u}^k - z^* \rangle + \tau\mathbb{E}\langle \hat{u}^k - z^k, \hat{u}^k - z^* \rangle \\
&+ \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\
=& \mathbb{E}\|\tilde{z}^k - z^*\|^2 - 2\gamma\mu\mathbb{E}\|\hat{u}^k - z^*\|^2 \\
&+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^* \rangle \\
&+ \tau\mathbb{E}\|m^k - z^*\|^2 - \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 - \tau\mathbb{E}\|\hat{u}^k - z^*\|^2 \\
&+ \tau\mathbb{E}\|\hat{u}^k - z^k\|^2 + \tau\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\
&+ \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\
=& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - 2\gamma\mu\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\
&+ 2\mathbb{E}\langle \gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k, \hat{u}^k - z^* \rangle \\
&+ \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2.
\end{aligned}
$$

By the Young's inequalities $2\langle a, b\rangle \le \|a\|^2 + \|b\|^2$ and $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$, we have

$$
\begin{aligned}
\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \le& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - 2\gamma\mu\mathbb{E}\|\hat{u}^k - z^*\|^2 \\
&+ \frac{2}{\gamma\mu}\mathbb{E}\|\gamma(F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)) + u_H^k - \hat{u}^k\|^2 \\
&+ \frac{\gamma\mu}{2}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 + \mathbb{E}\|\tilde{z}^{k+1} - \hat{u}^k\|^2 - \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 \\
&- \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 \\
=& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - \frac{3\gamma\mu}{2}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\
&+ \frac{4\gamma}{\mu}\mathbb{E}\|F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)\|^2 + \frac{4}{\gamma\mu}\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\
&+ \mathbb{E}\|u_H^k + \gamma \cdot \frac{1}{n}\sum_{i=1}^n F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k) - \hat{u}^k + \frac{\gamma}{n}\sum_{i=1}^n e_i^k\|^2 \\
&- \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 \\
=& \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - \frac{3\gamma\mu}{2}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\
&+ \frac{4\gamma}{\mu}\mathbb{E}\|F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)\|^2 + \frac{4}{\gamma\mu}\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\
&+ 3\mathbb{E}\|u_H^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_{i=1}^n F_i(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)\|^2
\end{aligned}
$$

$$- \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_i^n e_i^k\|. \tag{3}$$

Under Assumption 3, we can write 2 bounds below:

$$\begin{aligned}\|\frac{1}{n}\sum_{i=1}^n F(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)\|^2 &\leq \frac{1}{n}\sum_{i=1}^n \|F(m^k) - F_1(m^k) - F_i(u_H^k) + F_1(u_H^k)\| \\ &\leq 2\delta^2\|m^k - u_H^k\|^2 \end{aligned} \tag{4}$$

$$\|F(\hat{u}^k) - F_1(\hat{u}^k) - F(u_H^k) + F_1(u_H^k)\|^2 \leq \delta^2\|\hat{u}^k - u_H^k\|^2. \tag{5}$$

We substitute (4) and (5) in (3) to obtain

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 \leq &\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - \frac{3\gamma\mu}{2}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\ &+ \frac{4\gamma\delta^2}{\mu}\mathbb{E}\|u_H^k - \hat{u}^k\|^2 + \frac{4}{\gamma\mu}\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\ &+ 3\mathbb{E}\|u_H^k - \hat{u}^k\|^2 + 6\gamma^2\delta^2\mathbb{E}\|m^k - u_H^k\|^2 \\ &- \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - \tau\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_{i=1}^n e_i^k\|^2 \\ \leq &\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - \frac{3\gamma\mu}{2}\mathbb{E}\|\hat{u}^k - z^*\|^2 \\ &+ \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2\right)\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\ &- \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - (\tau - 12\gamma^2\delta^2)\mathbb{E}\|m^k - \hat{u}^k\|^2 - \tau\mathbb{E}\|z^k - z^*\|^2 \\ &+ \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_i^n e_i^k\|^2. \\ = &\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \tau\mathbb{E}\|m^k - z^*\|^2 - \frac{3\gamma\mu}{4}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \frac{3\gamma\mu}{4}\mathbb{E}\|\hat{u}^k - z^*\|^2 \\ &- \tau\mathbb{E}\|z^k - z^*\|^2 + \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2\right)\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\ &- \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - (\tau - 12\gamma^2\delta^2)\mathbb{E}\|m^k - \hat{u}^k\|^2 + \tau\mathbb{E}\|z^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_i^n e_i^k\|. \end{aligned} \tag{6}$$

The Young's inequality $\|a+b\|^2 \leq 2\|a\|^2 + 2\|b\|^2$ was also used here. The update of m^{k+1} gives

$$\mathbb{E}_{m^{k+1}}\left[\|m^{k+1} - z^*\|^2\right] = (1-p)\|m^k - z^*\|^2 + p\|z^k - z^*\|^2. \tag{7}$$

Then, summing up (6) and (7) with $\tau = p$, we obtain

$$\begin{aligned}\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \mathbb{E}\|m^{k+1} - z^*\|^2 \leq &\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \mathbb{E}\|m^k - z^*\|^2 \\ &- \frac{3\gamma\mu}{4}\mathbb{E}\|\hat{u}^k - z^*\|^2 - \frac{3\gamma\mu}{4}\mathbb{E}\|\hat{u}^k - z^*\|^2\end{aligned}$$

$$
\begin{aligned}
&+ \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2\right)\mathbb{E}\|u_H^k - \hat{u}^k\|^2 \\
&- \mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 - (\tau - 12\gamma^2\delta^2)\mathbb{E}\|m^k - \hat{u}^k\|^2 \\
&+ p\mathbb{E}\|z^k - \hat{u}^k\| + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_{i=1}^n e_i^k\|.
\end{aligned}
$$

Using $\|a+b\|^2 \geq \frac{2}{3}\|a\|^2 - 2\|b\|^2$, one can get

$$
\begin{aligned}
\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \mathbb{E}\|m^{k+1} - z^*\|^2 \leq& \left(1 - \frac{\gamma\mu}{2}\right)\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \left(1 - \frac{\gamma\mu}{2}\right)\mathbb{E}\|m^k - z^*\|^2 \\
&+ \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2\right)\mathbb{E}\|u_H^k - \hat{u}^k\| \\
&- \left(1 - \frac{3\gamma\mu}{2}\right)\mathbb{E}\|\tilde{z}^k - \hat{u}^k\| \\
&- \left(p - \frac{3\gamma\mu}{2} - 12\gamma^2\delta^2\right)\mathbb{E}\|m^k - \hat{u}^k\|^2 \\
&+ p\mathbb{E}\|z^k - \hat{u}^k\|^2 + 3\gamma^2\mathbb{E}\|\frac{1}{n}\sum_{i=0}^n e_i^k\|^2.
\end{aligned}
$$

We can bound the term $\mathbb{E}\|z^k - \hat{u}^k\|^2$:

$$
\mathbb{E}\|z^k - \hat{u}^k\|^2 = \mathbb{E}\|\tilde{z}^k - \frac{\gamma}{n}\sum_{i=1}^n e_i^k - \hat{u}^k\|^2 \leq 2\mathbb{E}\|\tilde{z}^k - \hat{u}^k\|^2 + 2\gamma^2\mathbb{E}\|\frac{1}{n}\sum_{i=1}^n e_i^k\|^2.
$$

This leads to the following estimate:

$$
\begin{aligned}
\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \mathbb{E}\|m^{k+1} - z^*\|^2 \leq& \left(1 - \frac{\gamma\mu}{2}\right)\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \left(1 - \frac{\gamma\mu}{2}\right)\mathbb{E}\|m^k - z^*\|^2 \\
&+ \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2\right)\mathbb{E}\|u_H^k - \hat{u}^k\| \\
&- \left(\frac{1}{2} - \frac{3\gamma\mu}{4} - p\right)\mathbb{E}\|z^k - \hat{u}^k\| \\
&- \left(p - \frac{3\gamma\mu}{2} - 12\gamma^2\delta^2\right)\mathbb{E}\|m^k - \hat{u}^k\|^2 \\
&+ \left(4 - \frac{3\gamma\mu}{2}\right)\mathbb{E}\|\frac{\gamma}{n}\sum_{i=0}^n e_i^k\|^2.
\end{aligned}
$$

The Young's inequality on error part gives

$$
\begin{aligned}
\mathbb{E}\|\frac{1}{n}\sum_{i=0}^n e_i^k\|^2 \leq& \frac{1}{n}\sum_{i=0}^n \mathbb{E}\|e_i^k\|^2 \\
\leq& \frac{1}{n}\sum_{i=1}^n \mathbb{E}\|e_i^{k-1} + F_i(m^{k-1}) - F_1(m^{k-1}) - F_i(u_H^{k-1}) + F_1(u_H^{k-1}) \\
&- C(e_i^{k-1} + F_i(m^{k-1}) - F_1(m^{k-1}) - F_i(u_H^{k-1}) + F_1(u_H^{k-1})\|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{1}{n}\left(1-\frac{1}{\omega}\right)\sum_{i=1}^{n}\mathbb{E}\|e_i^{k-1}+F_i(m^{k-1})-F_1(m^{k-1})-F_i(u_H^{k-1})+F_1(u_H^{k-1})\|^2\\
&\leq \frac{1}{n}\left(1-\frac{1}{\omega}\right)(1+c)\sum_{i}^{n}\mathbb{E}\|e_i^{k-1}\|^2\\
&\quad+\frac{1}{n}\left(1-\frac{1}{\omega}\right)(1+\frac{1}{c})\sum_{i=1}^{n}\mathbb{E}\|F_i(m^{k-1})-F_1(m^{k-1})-F_i(u_H^{k-1})+F_1(u_H^{k-1}))\|^2\\
&\leq \frac{1}{n}\left(1-\frac{1}{2\omega}\right)\sum_{i=1}^{n}\mathbb{E}\|e_i^{k-1}\|^2\\
&\quad+\frac{1}{n}2\omega\sum_{i=1}^{n}\mathbb{E}\|F_i(m^{k-1})-F_1(m^{k-1})-F_i(u_H^{k-1})+F_1(u_H^{k-1})\|^2\\
&\leq \frac{1}{n}\left(1-\frac{1}{2\omega}\right)\sum_{i=1}^{n}\mathbb{E}\|e_i^{k-1}\|^2\\
&\quad+\frac{2\omega}{n}\sum_{i=1}^{n}\mathbb{E}\|F_i(m^{k-1})-F_1(m^{k-1})-F_i(u_H^{k-1})+F_1(u_H^{k-1})\|^2\\
&\leq \frac{1}{n}\left(1-\frac{1}{2\omega}\right)\sum_{i=1}^{n}\mathbb{E}\|e_i^{k-1}\|^2\\
&\quad+2\omega\delta^2\mathbb{E}\|m^{k-1}-u_H^{k-1}\|^2.
\end{aligned}
$$

Running the reccurence, we get

$$
\begin{aligned}
\mathbb{E}\|\frac{1}{n}\sum_{i=0}^{n}e_i^k\|^2 &\leq \frac{1}{n}\sum_{i=0}^{n}\mathbb{E}\|e_i^k\|^2\\
&\leq 2\omega\delta^2\sum_{j=0}^{k-1}\left(1-\frac{1}{2\omega}\right)^{k-1-j}\mathbb{E}\|m^j-u_H^j\|^2.
\end{aligned}
$$

Weighting the previous expression by q^k, where $q^k \leq q^j\left(1+\frac{1}{4\omega}\right)^{k-j}$, leads to

$$
\begin{aligned}
\sum_{k=0}^{K-1}q^k\frac{1}{n}\sum_{i=0}^{n}\mathbb{E}\|e_i^k\|^2 \leq& 2\omega\delta^2\sum_{k=0}^{K-1}q^k\sum_{j=0}^{k-1}\left(1-\frac{1}{2\omega}\right)^{k-1-j}\mathbb{E}\|m^j-u_H^j\|^2\\
\leq&\frac{2\omega n\delta^2}{1-\frac{1}{2\omega}}\sum_{k=0}^{K-1}\sum_{j=0}^{K-1}q^j\left(1+\frac{1}{4\omega}\right)^{k-j}\left(1-\frac{1}{2\omega}\right)^{k-j}\mathbb{E}\|m^j-u_H^j\|^2\\
\leq&\frac{2\omega\delta^2}{1-\frac{1}{2\omega}}\sum_{k=0}^{K-1}\sum_{j=0}^{k-1}q^j\left(1-\frac{1}{4\omega}\right)^{k-j}\mathbb{E}\|m^j-u_H^j\|^2\\
\leq&\frac{2\omega\delta^2}{1-\frac{1}{2\omega}}\sum_{k=0}^{K-1}q^k\mathbb{E}\|m^k-u_H^k\|^2\sum_{j=0}^{\infty}\left(1-\frac{1}{4\omega}\right)^{j}\\
\leq&16\omega^2\delta^2\sum_{k=0}^{K-1}q^k\mathbb{E}\|m^k-u_H^k\|^2
\end{aligned}
$$

$$\leq 32\omega^2\delta^2 \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - \hat{u}^k\|^2 + 32\omega^2\delta^2 \sum_{k=0}^{K-1} q^k \mathbb{E}\|u_H^k - \hat{u}^k\|^2.$$

After weighting the recurrence we get what lemma says.

□

Next, it is time to finish with the theorem. In Line 3 we run Extra Gradient algorithm. The operator in (2) is $\left(L + \frac{1}{\gamma}\right)$-Lipschitz continuous (Assumption 1) and $\frac{1}{\gamma}$-strongly monotone (Assumption 2). For this case, there are convergence guarantees from [Alacaoglu, Malitsky, 2021]:

$$\text{with} \quad \eta = \frac{1}{4\left(L + \frac{1}{\gamma}\right)} \quad \text{it holds that} \quad \|u_H^k - \hat{u}^k\|^2 \leq \exp\left(-\frac{H}{4\left(\gamma L + 1\right)}\right) \|z^k - \hat{u}^k\|^2.$$

If $\gamma \leq \frac{1}{L} \cdot \left(\frac{H}{4 \log \frac{220L\omega}{\mu\sqrt{p}}} - 1\right)$, then it holds that

$$\|u_H^k - \hat{u}^k\|^2 \leq \frac{\mu p}{220L\omega} \|z^k - \hat{u}^k\|^2.$$

This fact, in addition to Lemma 1, gives

$$\begin{aligned}
&\sum_{k=0}^{K-1} q^k \mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^{k+1} - z^*\|^2 \\
&\quad\leq \left(1 - \frac{\gamma\mu}{2}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|\tilde{z}^k - z^*\|^2 + \left(1 - \frac{\gamma\mu}{2}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - z^*\|^2 \\
&\quad\quad + \left(3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2 + 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega^2\delta^2\right) \frac{\mu p}{220L} \sum_{k=0}^{K-1} q^k \mathbb{E}\|z^k - \hat{u}^k\|^2 \\
&\quad\quad - \left(\frac{1}{2} - p - \frac{3\gamma\mu}{4}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|z^k - \hat{u}^k\| \\
&\quad\quad - \left(p - \frac{3\gamma\mu}{2} - 12\gamma^2\delta^2 - 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega^2\delta^2\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - \hat{u}^k\|^2.
\end{aligned}$$

If we pick γ, H and p from Theorem 1, we get

$$\begin{aligned}
3 + \frac{4\gamma\delta^2}{\mu} + \frac{4}{\gamma\mu} + 12\gamma^2\delta^2 + 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega^2\delta^2 \leq& 3 + \frac{4\delta^2}{\mu} \cdot \frac{\sqrt{p}}{12\delta} \\
&+ \frac{4}{\mu} \cdot \max\left\{\frac{6\mu}{p}, \frac{12\delta}{\sqrt{p}}, L\left(\frac{H}{4\log\frac{220L}{\mu p}} - 1\right)^{-1}, \frac{\alpha\omega\delta}{\sqrt{p}}\right\} \\
&+ 12\delta^2 \cdot \left(\frac{\sqrt{p}}{12\delta}\right)^2 + 32\left(4 - \frac{3\gamma\mu}{2}\right)\omega^2\delta^2 \frac{p}{\alpha^2\omega^2\delta^2} \\
\leq& 3 + \frac{\delta\sqrt{p}}{3\mu} + \frac{4}{\mu} \cdot \max\left\{\frac{6\mu}{p}, \frac{12\delta}{\sqrt{p}}, L, \frac{\alpha\delta\omega}{\sqrt{p}}\right\} \\
&+ \frac{p}{2} + \frac{128p}{\alpha^2}
\end{aligned}$$

$$\leq 4 + \frac{128p}{\alpha^2} + \max\{48, 4\alpha + 1\} \max\left\{\frac{1}{p}, \frac{\delta\omega}{\mu\sqrt{p}}, \frac{L}{\mu}\right\}$$
$$\leq \frac{55L\omega}{\mu p}.$$

Next, if we pick $\alpha = \sqrt{128}$, we get

$$\sum_{k=0}^{K-1} q^k (\mathbb{E}\|\tilde{z}^{k+1} - z^*\|^2 + \mathbb{E}\|m^{k+1} - z^*\|^2)$$
$$\leq \left(1 - \frac{\gamma\mu}{2}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|\tilde{z}^k - z^*\|^2$$
$$+ \left(1 - \frac{\gamma\mu}{2}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - z^*\|^2$$
$$- \left(\frac{1}{4} - p - \frac{3\gamma\mu}{4}\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|z^k - \hat{u}^k\|$$
$$- \left(p - \frac{3\gamma\mu}{2} - 12\gamma^2\delta^2 - 32\left(4 - \frac{3\gamma\mu}{2}\right)\gamma^2\omega\delta^2\right) \sum_{k=0}^{K-1} q^k \mathbb{E}\|m^k - \hat{u}^k\|^2$$
$$\leq \left(1 - \frac{\gamma\mu}{2}\right) \sum_{k=0}^{K-1} q^k (\mathbb{E}\|\tilde{z}^k - z^*\|^2 + \mathbb{E}\|m^k - z^*\|^2).$$

With $q = \frac{1}{1-\mu\gamma/2}$ (for which $q^k \leq q^j \left(1 + \frac{1}{4\omega}\right)^{k-j}$ holds with chosen above learning rate), we have

$$\mathbb{E}\Big(\|\tilde{z}^{k+1} - z^*\|^2 + \|m^{k+1} - z^*\|^2\Big) \leq \left(1 - \frac{\gamma\mu}{2}\right)^K \mathbb{E}\big(\|\tilde{z}^0 - z^*\|^2 + \|m^0 - z^*\|^2\big)$$

This ends the proof.

□

References

Alacaoglu Ahmet, Malitsky Yura. Stochastic variance reduction for variational inequality methods. // arXiv preprint arXiv:2102.08352. 2021.