

APPENDIX E  
COMPLETE PROOF OF THEOREM 2

In this section, we provide the comprehensive proof of Theorem 2. We first present the proofs of Lemma 1, Lemma 2 and Lemma 3. Following that, we leverage these lemmas to derive Theorem 2.

**Lemma 1.** According to Eq. (30), function  $L_p$  has Lipschitz continuous Hessian and let  $L_w$  denote the Lipschitz constant. Based on definition of  $\eta_\mu$ ,  $\eta_v$ ,  $\eta_r$ ,  $\eta_p$  and  $\eta_q$ , we define  $\eta_\mu^k$ ,  $\eta_v^k$ ,  $\eta_r^k$ ,  $\eta_p^k$  and  $\eta_q^k$  as:

$$\begin{aligned} \eta_\mu^k &= \frac{1}{\frac{8\xi}{\eta_\lambda(c_\lambda^k)^2} + \frac{8\xi}{\eta_\alpha(c_\alpha^k)^2} + \frac{8\xi}{\eta_\beta(c_\beta^k)^2}}, & \eta_v^k &= \frac{1}{\frac{8\xi}{\eta_\beta(c_\beta^k)^2} + \frac{8\xi}{\eta_\gamma(c_\gamma^k)^2} \sum_{s=1}^L \|b_s\|^2}, \\ \eta_r^k &= \frac{\eta_\lambda(c_\lambda^k)^2}{32w_r^2\xi}, & \eta_p^k &= \frac{\eta_\alpha(c_\alpha^k)^2}{32w_p^2\xi}, & \eta_q^k &= \frac{\eta_\gamma(c_\gamma^k)^2}{32w_q^2\xi}. \end{aligned} \quad (38)$$

We can obtain that,

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{r_l^{k+1}\}, \{p_l^{k+1}\}, \{q_s^{k+1}\}, v^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & - L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \left(\frac{L_w}{6} + 1 - \frac{1}{\eta_\mu^k}\right) \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \left(-\frac{1}{\eta_r^k} + \frac{w_r L_w M^{\frac{1}{2}}}{3} + w_\lambda\right) \sum_{l=1}^L \|r_l^{k+1} - r_l^k\|^2 + \left(\frac{L_w}{6} - \frac{1}{\eta_v^k}\right) \|v^{k+1} - v^k\|^2 \end{aligned} \quad (39)$$

*Proof.* According to the Lipschitz property of  $L_p$ , we can obtain that,

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) - L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \sum_{l=1}^L \left( \left\langle \nabla_{\mu_l} L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \mu_l^{k+1} - \mu_l^k \right\rangle + \frac{L_w}{6} \|\mu_l^{k+1} - \mu_l^k\|^2 \right) \\ & + \frac{1}{2} \sum_{l=1}^L \left\langle \nabla_{\mu_l}^2 L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) (\mu_l^{k+1} - \mu_l^k), \mu_l^{k+1} - \mu_l^k \right\rangle. \end{aligned} \quad (40)$$

According to (21) and the definition of  $\eta_\mu$ , we have that,

$$\begin{aligned} & \left\langle \nabla_{\mu_l} L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \mu_l^{k+1} - \mu_l^k \right\rangle = -\frac{1}{\eta_\mu} \|\mu_l^{k+1} - \mu_l^k\|^2 \\ & \leq -\frac{1}{\eta_\mu^k} \|\mu_l^{k+1} - \mu_l^k\|^2. \end{aligned} \quad (41)$$

Given that  $\mu_l^k = \mu_l^{\hat{k}_l}$ ,  $\lambda_l^k = \lambda_l^{\hat{k}_l}$ ,  $\alpha_l^k = \alpha_l^{\hat{k}_l}$ , and  $\beta_l^k = \beta_l^{\hat{k}_l}$ , combining (41) with the definition of  $(\nabla \tilde{G}^k)_{\mu_l}$ , we have,

$$\left\langle \nabla_{\mu_l} L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \mu_l^{k+1} - \mu_l^k \right\rangle \leq -\frac{1}{\eta_\mu^k} \|\mu_l^{k+1} - \mu_l^k\|^2. \quad (42)$$

According to the definition of  $L_p$ , we have,

$$\frac{1}{2} \sum_{l=1}^L \left\langle \nabla_{\mu_l}^2 L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) (\mu_l^{k+1} - \mu_l^k), \mu_l^{k+1} - \mu_l^k \right\rangle = \|\mu_l^{k+1} - \mu_l^k\|^2. \quad (43)$$

Combining (40), (42) with (43), we can obtain that,

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) - L_p(\{\mu_l^k\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \left(-\frac{1}{\eta_\mu^k} + \frac{L_w}{6} + 1\right) \sum_{l=1}^L (\|\mu_l^{k+1} - \mu_l^k\|^2). \end{aligned} \quad (44)$$

According to the Lipschitz property of  $L_p$ , we can obtain that,

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{r_l^{k+1}\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) - L_p(\{\mu_l^{k+1}\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \sum_{l=1}^L \left\langle \nabla_{r_l} L_p(\{\mu_l^{k+1}\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), r_l^{k+1} - r_l^k \right\rangle + \frac{L_w}{6} \sum_{l=1}^L \|r_l^{k+1} - r_l^k\|^3 \\ & + \frac{1}{2} \sum_{l=1}^L \left\langle \nabla_{r_l}^2 L_p(\{\mu_l^{k+1}\}, \{r_l^k\}, \{p_l^k\}, \{q_s^k\}, v^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) (r_l^{k+1} - r_l^k), r_l^{k+1} - r_l^k \right\rangle, \end{aligned} \quad (45)$$

Similar to (41)-(42), we have,

$$\left\langle \nabla_{\mathbf{r}_l} L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}), \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \right\rangle \leq -\frac{1}{\eta_{\mathbf{r}}^k} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2. \quad (46)$$

According to Assumption 1 and the definition of  $L_p$ , we can obtain that,

$$\frac{1}{2} \left( \left\langle \nabla_{\mathbf{r}_l}^2 L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\})(\mathbf{r}_l^{k+1} - \mathbf{r}_l^k), \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \right\rangle \right) \leq w_{\lambda} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2. \quad (47)$$

By employing trigonometric inequality and based on Assumption 1, we have that,

$$\frac{Lw}{6} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^3 \leq \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2. \quad (48)$$

Combining (45)-(48), we can obtain that,

$$\begin{aligned} & L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\ & \leq \left( -\frac{1}{\eta_{\mathbf{r}}^k} + \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} + w_{\lambda} \right) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2. \end{aligned} \quad (49)$$

Likewise, similar results can be obtained for variable  $\mathbf{p}_l$ ,  $q_s$  and  $\mathbf{v}$  :

$$\begin{aligned} & L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\ & \leq \left( -\frac{1}{\eta_{\mathbf{p}}^k} + \frac{w_{\mathbf{p}} L_w M^{\frac{1}{2}}}{3} + w_{\alpha} \right) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2. \end{aligned} \quad (50)$$

$$\begin{aligned} & L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\ & \leq \left( -\frac{1}{\eta_{\mathbf{q}}^k} + \frac{w_{\mathbf{q}} L_w}{3} + w_{\gamma} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2. \end{aligned} \quad (51)$$

$$\begin{aligned} & L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\ & \leq \left( -\frac{1}{\eta_{\mathbf{v}}^k} + \frac{Lw}{6} \right) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2. \end{aligned} \quad (52)$$

By combining (44), (49), (50), (51) and (52), we conclude the proof of Lemma 1.  $\square$

**Lemma 2.**  $\forall k \geq K_1$ , we have:

$$\begin{aligned} & L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^{k+1}\}, \{\gamma_s^{k+1}\}) \\ & - L_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\ & \leq \left( \frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \frac{Lw}{6} + 1 - \frac{1}{\eta_{\boldsymbol{\mu}}^k} \right) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + \left( \frac{4w_{\mathbf{r}}^2}{a_1} + \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} + w_{\lambda} - \frac{1}{\eta_{\mathbf{r}}^k} \right) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\ & + \left( \frac{4w_{\mathbf{p}}^2}{a_2} + \frac{w_{\mathbf{p}} L_w M^{\frac{1}{2}}}{3} + w_{\alpha} - \frac{1}{\eta_{\mathbf{p}}^k} \right) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + \left( \frac{4w_{\mathbf{q}}^2}{a_4} + \frac{w_{\mathbf{q}} L_w}{3} + w_{\gamma} - \frac{1}{\eta_{\mathbf{q}}^k} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 \\ & + \left( -\frac{1}{\eta_{\mathbf{v}}^k} + \frac{Lw}{6} \right) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 + \frac{1}{a_3} \sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2 + \frac{1}{a_4} \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\ & + \left( \frac{a_1}{2} - \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\lambda}}^k} \right) \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\|^2 + \frac{c_{\boldsymbol{\lambda}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\lambda}_l^{k+1}\|^2 - \|\boldsymbol{\lambda}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\lambda}}^k} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}\|^2 \\ & + \left( \frac{a_2}{2} - \frac{c_{\boldsymbol{\alpha}}^{k-1} - c_{\boldsymbol{\alpha}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\alpha}}^k} \right) \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{k+1} - \boldsymbol{\alpha}_l^k\|^2 + \frac{c_{\boldsymbol{\alpha}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\alpha}_l^{k+1}\|^2 - \|\boldsymbol{\alpha}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\alpha}}^k} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^k - \boldsymbol{\alpha}_l^{k-1}\|^2 \\ & + \left( \frac{a_3}{2} - \frac{c_{\boldsymbol{\beta}}^{k-1} - c_{\boldsymbol{\beta}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\beta}}^k} \right) \sum_{l=1}^L \|\boldsymbol{\beta}_l^{k+1} - \boldsymbol{\beta}_l^k\|^2 + \frac{c_{\boldsymbol{\beta}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\beta}_l^{k+1}\|^2 - \|\boldsymbol{\beta}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\beta}}^k} \sum_{l=1}^L \|\boldsymbol{\beta}_l^k - \boldsymbol{\beta}_l^{k-1}\|^2 \\ & + \left( \frac{a_4}{2} - \frac{c_{\gamma}^{k-1} - c_{\gamma}^k}{2} + \frac{1}{2\eta_{\gamma}^k} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + \frac{c_{\gamma}^{k-1}}{2} \sum_{s=1}^{|\mathcal{P}^k|} (\|\gamma_s^{k+1}\|^2 - \|\gamma_s^k\|^2) + \frac{1}{2\eta_{\gamma}^k} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k - \gamma_s^{k-1}\|^2, \end{aligned} \quad (53)$$

where  $a_1 \geq 0$ ,  $a_2 \geq 0$ ,  $a_3 \geq 0$  and  $a_4 \geq 0$  are constants.

*Proof.* According to , in  $(k+1)^{th}$  iteration, for  $\forall \lambda$  and  $\forall l \in \mathcal{Q}^{k+1}$  it follows that,

$$\left\langle \lambda_l^{k+1} - \lambda_l^k - \eta_\lambda \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda - \lambda_l^{k+1} \right\rangle \geq 0. \quad (54)$$

Let  $\lambda = \lambda_l^k$ , we can obtain:

$$\left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) - \frac{1}{\eta_\lambda}(\lambda_l^{k+1} - \lambda_l^k), \lambda_l^k - \lambda_l^{k+1} \right\rangle \leq 0. \quad (55)$$

Likewise, in the  $k^{th}$  iteration, we have that,

$$\left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}) - \frac{1}{\eta_\lambda}(\lambda_l^k - \lambda_l^{k-1}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \leq 0. \quad (56)$$

Since  $\lambda_l^{k+1} - \lambda_l^k = 0, l \notin \mathcal{Q}^{k+1}$ , inequality (56) holds for  $l$ . According to (31),  $\tilde{L}_p$  is concave with respect to  $\lambda_l$ . Based on (56), we can obtain that,

$$\begin{aligned} & \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & - \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \sum_{l=1}^L \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & \leq \sum_{l=1}^L \left( \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle + \frac{1}{\eta_\lambda} \langle \lambda_l^k - \lambda_l^{k-1}, \lambda_l^{k+1} - \lambda_l^k \rangle \right) \\ & - \sum_{l=1}^L \left( \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \right). \end{aligned} \quad (57)$$

And we have that,

$$\begin{aligned} & \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & = \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & + \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1}) \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}), \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1}) \right\rangle \\ & + \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^k - \lambda_l^{k-1} \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}), \lambda_l^k - \lambda_l^{k-1} \right\rangle. \end{aligned} \quad (58)$$

According to the definition of  $\tilde{L}_p$  and Cauchy-Schwarz inequality, we can obtain that,

$$\begin{aligned} & \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle \\ & \leq \left\langle \nabla_{\lambda_l} L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) \\ & - \left\langle \nabla_{\lambda_l} L_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}), \lambda_l^{k+1} - \lambda_l^k \right\rangle - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 \\ & \leq \frac{a_1}{2} \|\lambda_l^{k+1} - \lambda_l^k\| + \frac{1}{a_1} \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{1}{a_1} \|\mathbf{r}_l^{k+1} \circ \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \circ \mathbf{r}_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) \\ & - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2, \end{aligned} \quad (59)$$

where  $a_1 > 0$  is a constant. According to the definition of  $\tilde{L}_p$  and Cauchy-Schwarz inequality, we also have that,

$$\begin{aligned} & \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}, \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})) \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}, \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})) \right\rangle \\ & \leq \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})\|^2 + \frac{\eta_\lambda}{2} \|c_\lambda^{k-1}(\lambda_l^k - \lambda_l^{k-1})\|^2. \end{aligned} \quad (60)$$

Following [33], since  $\tilde{L}_p$  is strong concave with respect to  $\lambda$ , we have,

$$\begin{aligned} & \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}, \lambda_l^k - \lambda_l^{k-1}) \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}, \lambda_l^k - \lambda_l^{k-1}) \right\rangle \\ & \leq -\frac{1}{2c_\lambda^{k-1}} \|c_\lambda^{k-1}(\lambda_l^k - \lambda_l^{k-1})\|^2 - \frac{c_\lambda^{k-1}}{2} \|\lambda_l^k - \lambda_l^{k-1}\|^2. \end{aligned} \quad (61)$$

Since  $\frac{\eta_\lambda}{2} \leq \frac{1}{2c_\lambda^{k-1}}$ , combining (58)-(61), we have that,

$$\begin{aligned} & \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}, \lambda_l^{k+1} - \lambda_l^k) \right\rangle \\ & - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}, \lambda_l^{k+1} - \lambda_l^k) \right\rangle \\ & \leq \frac{a_1}{2} \|\lambda_l^{k+1} - \lambda_l^k\| + \frac{1}{a_1} \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{1}{a_1} \|\mathbf{r}_l^{k+1} \circ \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \circ \mathbf{r}_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) \\ & - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})\|^2 - \frac{c_\lambda^{k-1}}{2} \|\lambda_l^k - \lambda_l^{k-1}\|^2. \end{aligned} \quad (62)$$

According to Assumption 1, we have,

$$\frac{1}{a_1} \|\mathbf{r}_l^{k+1} \circ \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \circ \mathbf{r}_l^k\|^2 \leq \frac{4w_r^2}{a_1} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2. \quad (63)$$

In addition, the following equality can be obtained,

$$\frac{1}{\eta_\lambda} \left\langle \lambda_l^k - \lambda_l^{k-1}, \lambda_l^{k+1} - \lambda_l^k \right\rangle = \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{1}{2\eta_\lambda} \|\lambda_l^k - \lambda_l^{k-1}\|^2 - \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})\|^2. \quad (64)$$

Combining (57), (62), (63) with (64), we can obtain,

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & - L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \frac{1}{a_1} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{4w_r^2}{a_1} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \left(\frac{a_1}{2} - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} + \frac{1}{2\eta_\lambda}\right) \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 \\ & + \frac{c_\lambda^{k-1}}{2} \sum_{l=1}^L (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) + \frac{1}{2\eta_\lambda} \sum_{l=1}^L \|\lambda_l^k - \lambda_l^{k-1}\|^2. \end{aligned} \quad (65)$$

Likewise, similar results can be obtained for other variables :

$$\begin{aligned} & L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^{k+1}\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & - L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\ & \leq \frac{1}{a_2} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{4w_p^2}{a_2} \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + \left(\frac{a_2}{2} - \frac{c_\alpha^{k-1} - c_\alpha^k}{2} + \frac{1}{2\eta_\alpha}\right) \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 \\ & + \frac{c_\alpha^{k-1}}{2} \sum_{l=1}^L (\|\alpha_l^{k+1}\|^2 - \|\alpha_l^k\|^2) + \frac{1}{2\eta_\alpha} \sum_{l=1}^L \|\alpha_l^k - \alpha_l^{k-1}\|^2, \end{aligned} \quad (66)$$

$$\begin{aligned}
& L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^{k+1}\}, \{\gamma_s^k\}) \\
& - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\
& \leq \frac{1}{a_3} \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + \frac{1}{a_3} \sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2 + \left(\frac{a_3}{2} - \frac{c_{\boldsymbol{\beta}}^{k-1} - c_{\boldsymbol{\beta}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\beta}}}\right) \sum_{l=1}^L \|\boldsymbol{\beta}_l^{k+1} - \boldsymbol{\beta}_l^k\|^2 \\
& + \frac{c_{\boldsymbol{\beta}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\beta}_l^{k+1}\|^2 - \|\boldsymbol{\beta}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\beta}}} \sum_{l=1}^L \|\boldsymbol{\beta}_l^k - \boldsymbol{\beta}_l^{k-1}\|^2,
\end{aligned} \tag{67}$$

$$\begin{aligned}
& L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^{k+1}\}, \{\gamma_s^{k+1}\}) \\
& - L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^{k+1}\}, \{\gamma_s^k\}) \\
& \leq \frac{1}{a_4} \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 + \frac{4w_q^2}{a_4} \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 + \left(\frac{a_4}{2} - \frac{c_{\gamma}^{k-1} - c_{\gamma}^k}{2} + \frac{1}{2\eta_{\gamma}}\right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 \\
& + \frac{c_{\gamma}^{k-1}}{2} \sum_{s=1}^{|\mathcal{P}^k|} (\|\gamma_s^{k+1}\|^2 - \|\gamma_s^k\|^2) + \frac{1}{2\eta_{\gamma}} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k - \gamma_s^{k-1}\|^2,
\end{aligned} \tag{68}$$

where  $a_2 > 0$ ,  $a_3 > 0$  and  $a_4 > 0$  are constant. Combining Lemma 1 with (65), (66), (67) and (68), we can obtain that,

$$\begin{aligned}
& L_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^{k+1}\}, \{\boldsymbol{\beta}_l^{k+1}\}, \{\gamma_s^{k+1}\}) \\
& - L_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\
& \leq \left(\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \frac{L_w}{6} + 1 - \frac{1}{\eta_{\boldsymbol{\mu}}^k}\right) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + \left(\frac{4w_{\mathbf{r}}^2}{a_1} + \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} + w_{\boldsymbol{\lambda}} - \frac{1}{\eta_{\mathbf{r}}^k}\right) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + \left(\frac{4w_{\mathbf{p}}^2}{a_2} + \frac{w_{\mathbf{p}} L_w M^{\frac{1}{2}}}{3} + w_{\boldsymbol{\alpha}} - \frac{1}{\eta_{\mathbf{p}}^k}\right) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + \left(\frac{4w_q^2}{a_4} + \frac{w_q L_w}{3} + w_{\gamma} - \frac{1}{\eta_q^k}\right) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 \\
& + \left(-\frac{1}{\eta_{\mathbf{v}}^k} + \frac{L_w}{6}\right) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 + \frac{1}{a_3} \sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2 + \frac{1}{a_4} \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\
& + \left(\frac{a_1}{2} - \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\lambda}}}\right) \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\|^2 + \frac{c_{\boldsymbol{\lambda}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\lambda}_l^{k+1}\|^2 - \|\boldsymbol{\lambda}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\lambda}}} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}\|^2 \\
& + \left(\frac{a_2}{2} - \frac{c_{\boldsymbol{\alpha}}^{k-1} - c_{\boldsymbol{\alpha}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\alpha}}}\right) \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{k+1} - \boldsymbol{\alpha}_l^k\|^2 + \frac{c_{\boldsymbol{\alpha}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\alpha}_l^{k+1}\|^2 - \|\boldsymbol{\alpha}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\alpha}}} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^k - \boldsymbol{\alpha}_l^{k-1}\|^2 \\
& + \left(\frac{a_3}{2} - \frac{c_{\boldsymbol{\beta}}^{k-1} - c_{\boldsymbol{\beta}}^k}{2} + \frac{1}{2\eta_{\boldsymbol{\beta}}}\right) \sum_{l=1}^L \|\boldsymbol{\beta}_l^{k+1} - \boldsymbol{\beta}_l^k\|^2 + \frac{c_{\boldsymbol{\beta}}^{k-1}}{2} \sum_{l=1}^L (\|\boldsymbol{\beta}_l^{k+1}\|^2 - \|\boldsymbol{\beta}_l^k\|^2) + \frac{1}{2\eta_{\boldsymbol{\beta}}} \sum_{l=1}^L \|\boldsymbol{\beta}_l^k - \boldsymbol{\beta}_l^{k-1}\|^2 \\
& + \left(\frac{a_4}{2} - \frac{c_{\gamma}^{k-1} - c_{\gamma}^k}{2} + \frac{1}{2\eta_{\gamma}}\right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + \frac{c_{\gamma}^{k-1}}{2} \sum_{s=1}^{|\mathcal{P}^k|} (\|\gamma_s^{k+1}\|^2 - \|\gamma_s^k\|^2) + \frac{1}{2\eta_{\gamma}} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k - \gamma_s^{k-1}\|^2,
\end{aligned} \tag{69}$$

which conclude the proof of Lemma 2.  $\square$

**Lemma 3.** Define  $S_1^{k+1}$ ,  $S_2^{k+1}$ ,  $S_3^{k+1}$ ,  $S_4^{k+1}$ ,  $F^{k+1}$  and  $a_5$  as :

$$\begin{aligned}
S_1^{k+1} &= \frac{4}{\eta_\lambda^2 c_\lambda^{k+1}} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-1}}{c_\lambda^k} - 1 \right) \sum_{l=1}^L \|\lambda_l^{k+1}\|^2, \\
S_2^{k+1} &= \frac{4}{\eta_\alpha^2 c_\alpha^{k+1}} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 - \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-1}}{c_\alpha^k} - 1 \right) \sum_{l=1}^L \|\alpha_l^{k+1}\|^2, \\
S_3^{k+1} &= \frac{4}{\eta_\beta^2 c_\beta^{k+1}} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 - \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-1}}{c_\beta^k} - 1 \right) \sum_{l=1}^L \|\beta_l^{k+1}\|^2, \\
S_4^{k+1} &= \frac{4}{\eta_\gamma^2 c_\gamma^{k+1}} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 - \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-1}}{c_\gamma^k} - 1 \right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1}\|^2.
\end{aligned} \tag{70}$$

$$\begin{aligned}
F^{k+1} &= L_P(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^{k+1}\}, \{\beta_l^{k+1}\}, \{\gamma_s^{k+1}\}) \\
&\quad + S_1^{k+1} + S_2^{k+1} + S_3^{k+1} + S_4^{k+1} \\
&\quad - \frac{7}{2\eta_\lambda} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{c_\lambda^k}{2} \sum_{l=1}^L \|\lambda_l^{k+1}\|^2 \\
&\quad - \frac{7}{2\eta_\alpha} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 - \frac{c_\alpha^k}{2} \sum_{l=1}^L \|\alpha_l^{k+1}\|^2 \\
&\quad - \frac{7}{2\eta_\beta} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 - \frac{c_\beta^k}{2} \sum_{l=1}^L \|\beta_l^{k+1}\|^2 \\
&\quad - \frac{7}{2\eta_\gamma} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 - \frac{c_\gamma^k}{2} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1}\|^2.
\end{aligned} \tag{71}$$

$$a_5 = \max \left\{ \eta_\lambda + \eta_\alpha + \eta_\beta + \frac{Lw}{6} + 1, 4w_{\mathbf{r}}^2\eta_\lambda + \frac{w_{\mathbf{r}}LwM^{\frac{1}{2}}}{3} + w_\lambda, 4w_{\mathbf{p}}^2\eta_\lambda + \frac{w_{\mathbf{p}}LwM^{\frac{1}{2}}}{3} + w_\alpha, 4w_q^2\eta_\gamma + \frac{w_qLw}{3} + w_\gamma, \eta_\beta + \frac{Lw}{6} + \eta_\gamma \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \right\}, \tag{72}$$

And then  $\forall k \geq K_1$ , we have that,

$$\begin{aligned}
F^{k+1} - F^k &\leq (a_5 - \frac{1}{\eta_\mu^k} + \frac{16}{\eta_\lambda(c_\lambda^k)^2} + \frac{16}{\eta_\alpha(c_\alpha^k)^2} + \frac{16}{\eta_\beta(c_\beta^k)^2}) \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 \\
&\quad + (a_5 - \frac{1}{\eta_{\mathbf{r}}^k} + \frac{64w_{\mathbf{r}}^2}{\eta_\lambda(c_\lambda^k)^2}) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
&\quad + (a_5 - \frac{1}{\eta_{\mathbf{p}}^k} + \frac{64w_{\mathbf{p}}^2}{\eta_\alpha(c_\alpha^k)^2}) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
&\quad + (a_5 - \frac{1}{\eta_q^k} + \frac{64w_q^2}{\eta_\gamma(c_\gamma^k)^2}) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 \\
&\quad + (a_5 - \frac{1}{\eta_{\mathbf{v}}^k} + \frac{16}{\eta_\beta(c_\beta^k)^2} + \frac{16}{\eta_\beta(c_\beta^k)^2} \sum_{s=1}^L \|b_s\|^2) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\
&\quad - \frac{1}{10\eta_\lambda} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \sum_{l=1}^L \|\lambda_l^{k+1}\|^2 + \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-2}}{c_\lambda^{k-1}} - \frac{c_\lambda^{k-1}}{c_\lambda^k} \right) \sum_{l=1}^L \|\lambda_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\alpha} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 + \frac{c_\alpha^{k-1} - c_\alpha^k}{2} \sum_{l=1}^L \|\alpha_l^{k+1}\|^2 + \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-2}}{c_\alpha^{k-1}} - \frac{c_\alpha^{k-1}}{c_\alpha^k} \right) \sum_{l=1}^L \|\alpha_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\beta} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 + \frac{c_\beta^{k-1} - c_\beta^k}{2} \sum_{l=1}^L \|\beta_l^{k+1}\|^2 + \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-2}}{c_\beta^{k-1}} - \frac{c_\beta^{k-1}}{c_\beta^k} \right) \sum_{l=1}^L \|\beta_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\gamma} \sum_{l=1}^L \|\gamma_l^{k+1} - \gamma_l^k\|^2 + \frac{c_\gamma^{k-1} - c_\gamma^k}{2} \sum_{l=1}^L \|\gamma_l^{k+1}\|^2 + \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k} \right) \sum_{l=1}^L \|\gamma_l^k\|^2.
\end{aligned} \tag{73}$$

*Proof.* Let  $a_1 = \frac{1}{\eta_\lambda}, a_2 = \frac{1}{\eta_\alpha}, a_3 = \frac{1}{\eta_\beta}, a_4 = \frac{1}{\eta_\gamma}$ , and substitute them into the Lemma 2,  $\forall k \geq K_1$ , we have,

$$\begin{aligned}
& L_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^{k+1}\}, \{\alpha_l^{k+1}\}, \{\beta_l^{k+1}\}, \{\gamma_s^{k+1}\}) \\
& - L_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}) \\
& \leq (\eta_\lambda + \eta_\alpha + \eta_\beta + \frac{Lw}{6} + 1 - \frac{1}{\eta_\mu^k}) \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + (4w_r^2\eta_\lambda + \frac{w_r Lw M^{\frac{1}{2}}}{3} + w_\lambda - \frac{1}{\eta_r^k}) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + (4w_p^2\eta_\alpha + \frac{w_p Lw M^{\frac{1}{2}}}{3} + w_\alpha - \frac{1}{\eta_p^k}) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + (4w_q^2\eta_\gamma + \frac{w_q Lw}{3} + w_\gamma - \frac{1}{\eta_q^k}) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 \\
& + (-\frac{1}{\eta_v^k} + \frac{Lw}{6}) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 + \eta_\beta \sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2 + \eta_\gamma \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\
& + (-\frac{c_\lambda^{k-1} - c_\lambda^k}{2} + \frac{1}{\eta_\lambda}) \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{c_\lambda^{k-1}}{2} \sum_{l=1}^L (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) + \frac{1}{2\eta_\lambda} \sum_{l=1}^L \|\lambda_l^k - \lambda_l^{k-1}\|^2 \\
& + (-\frac{c_\alpha^{k-1} - c_\alpha^k}{2} + \frac{1}{\eta_\alpha}) \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 + \frac{c_\alpha^{k-1}}{2} \sum_{l=1}^L (\|\alpha_l^{k+1}\|^2 - \|\alpha_l^k\|^2) + \frac{1}{2\eta_\alpha} \sum_{l=1}^L \|\alpha_l^k - \alpha_l^{k-1}\|^2 \\
& + (-\frac{c_\beta^{k-1} - c_\beta^k}{2} + \frac{1}{\eta_\beta}) \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 + \frac{c_\beta^{k-1}}{2} \sum_{l=1}^L (-\|\beta_l^k\|^2) + \frac{1}{2\eta_\beta} \sum_{l=1}^L \|\beta_l^k - \beta_l^{k-1}\|^2 \\
& + (-\frac{c_\gamma^{k-1} - c_\gamma^k}{2} + \frac{1}{\eta_\gamma}) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + \frac{c_\gamma^{k-1}}{2} \sum_{s=1}^{|\mathcal{P}^k|} (\|\gamma_s^{k+1}\|^2 - \|\gamma_s^k\|^2) + \frac{1}{2\eta_\gamma} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k - \gamma_s^{k-1}\|^2.
\end{aligned} \tag{74}$$

According to (55) and (56), in the  $(k+1)^{th}$  iteration, we can obtain,

$$\begin{aligned}
& \frac{1}{\eta_\lambda} \langle \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1}), \lambda_l^{k+1} - \lambda_l^k \rangle \\
& \leq \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\lambda_l^k\}, \{\alpha_l^k\}, \{\beta_l^k\}, \{\gamma_s^k\}, \lambda_l^{k+1} - \lambda_l^k) \right\rangle \\
& - \left\langle \nabla_{\lambda_l} \tilde{L}_p(\{\mu_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\lambda_l^{k-1}\}, \{\alpha_l^{k-1}\}, \{\beta_l^{k-1}\}, \{\gamma_s^{k-1}\}, \lambda_l^{k+1} - \lambda_l^k) \right\rangle \\
& \leq \frac{1}{b_1^k} \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{4w_r^2}{b_1^k} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \frac{b_1^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) \\
& - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})\|^2 - \frac{c_\lambda^{k-1}}{2} \|\lambda_l^k - \lambda_l^{k-1}\|^2.
\end{aligned} \tag{75}$$

where  $b_1^k > 0$ . And we have,

$$\frac{1}{\eta_\lambda} \langle \lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1}), \lambda_l^{k+1} - \lambda_l^k \rangle = \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{1}{2\eta_\lambda} \|\lambda_l^k - \lambda_l^{k-1}\|^2 + \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k - (\lambda_l^k - \lambda_l^{k-1})\|^2. \tag{76}$$

Combining (75) with (76), we have

$$\begin{aligned}
& \frac{1}{2\eta_\lambda} \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{1}{2\eta_\lambda} \|\lambda_l^k - \lambda_l^{k-1}\|^2 \leq \frac{1}{b_1^k} \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{4w_r^2}{b_1^k} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \frac{b_1^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 \\
& + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} (\|\lambda_l^{k+1}\|^2 - \|\lambda_l^k\|^2) - \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{c_\lambda^{k-1}}{2} \|\lambda_l^k - \lambda_l^{k-1}\|^2.
\end{aligned} \tag{77}$$

Multiplying both sides of by  $\frac{8}{\eta_\lambda c_\lambda^k}$ , we have,

$$\begin{aligned}
& \frac{4}{\eta_\lambda^2 c_\lambda^k} \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-1} - c_\lambda^k}{c_\lambda^k} \right) \|\lambda_l^{k+1}\|^2 \\
& \leq \frac{4}{\eta_\lambda^2 c_\lambda^k} \|\lambda_l^k - \lambda_l^{k-1}\|^2 - \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-1} - c_\lambda^k}{c_\lambda^k} \right) \|\lambda_l^k\|^2 + \frac{4b_1^k}{\eta_\lambda c_\lambda^k} \|\lambda_l^{k+1} - \lambda_l^k\|^2 - \frac{4}{\eta_\lambda} \|\lambda_l^k - \lambda_l^{k-1}\|^2 \\
& + \frac{8}{b_1^k \eta_\lambda c_\lambda^k} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{32w_r^2}{b_1^k \eta_\lambda c_\lambda^k} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2.
\end{aligned} \tag{78}$$

Setting  $b_1^k = \frac{c_\lambda^k}{2}$  in (78) and combine it with the definition of  $S_1^k$ , we have,

$$\begin{aligned}
S_1^{k+1} - S_1^k &\leq \sum_{l=1}^L \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-2}}{c_\lambda^{k-1}} - \frac{c_\lambda^{k-1}}{c_\lambda^k} \right) \|\lambda_l^k\|^2 + \sum_{l=1}^L \left( \frac{2}{\eta_\lambda} + \frac{4}{\eta_\lambda^2} \left( \frac{1}{c_\lambda^{k+1}} - \frac{1}{c_\lambda^k} \right) \right) \|\lambda_l^{k+1} - \lambda_l^k\|^2 \\
&\quad - \sum_{l=1}^L \frac{4}{\eta_\lambda} \|\lambda_l^k - \lambda_l^{k-1}\|^2 + \frac{16}{\eta_\lambda (c_\lambda^k)^2} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{64w_r^2}{\eta_\lambda (c_\lambda^k)^2} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2.
\end{aligned} \tag{79}$$

Similar results can be obtained for  $S_2^k, S_3^k, S_4^k$ ,

$$\begin{aligned}
S_2^{k+1} - S_2^k &\leq \sum_{l=1}^L \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-2}}{c_\alpha^{k-1}} - \frac{c_\alpha^{k-1}}{c_\alpha^k} \right) \|\alpha_l^k\|^2 + \sum_{l=1}^L \left( \frac{2}{\eta_\alpha} + \frac{4}{\eta_\alpha^2} \left( \frac{1}{c_\alpha^{k+1}} - \frac{1}{c_\alpha^k} \right) \right) \|\alpha_l^{k+1} - \alpha_l^k\|^2 \\
&\quad - \sum_{l=1}^L \frac{4}{\eta_\alpha} \|\alpha_l^k - \alpha_l^{k-1}\|^2 + \frac{16}{\eta_\alpha (c_\alpha^k)^2} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{64w_p^2}{\eta_\alpha (c_\alpha^k)^2} \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2.
\end{aligned} \tag{80}$$

$$\begin{aligned}
S_3^{k+1} - S_3^k &\leq \sum_{l=1}^L \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-2}}{c_\beta^{k-1}} - \frac{c_\beta^{k-1}}{c_\beta^k} \right) \|\beta_l^k\|^2 + \sum_{l=1}^L \left( \frac{2}{\eta_\beta} + \frac{4}{\eta_\beta^2} \left( \frac{1}{c_\beta^{k+1}} - \frac{1}{c_\beta^k} \right) \right) \|\beta_l^{k+1} - \beta_l^k\|^2 \\
&\quad - \sum_{l=1}^L \frac{4}{\eta_\beta} \|\beta_l^k - \beta_l^{k-1}\|^2 + \frac{16}{\eta_\beta (c_\beta^k)^2} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \frac{16}{\eta_\beta (c_\beta^k)^2} \sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2.
\end{aligned} \tag{81}$$

$$\begin{aligned}
S_4^{k+1} - S_4^k &\leq \sum_{s=1}^{|\mathcal{P}^k|} \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k} \right) \|\gamma_s^k\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \left( \frac{2}{\eta_\gamma} + \frac{4}{\eta_\gamma^2} \left( \frac{1}{c_\gamma^{k+1}} - \frac{1}{c_\gamma^k} \right) \right) \|\gamma_s^{k+1} - \gamma_s^k\|^2 \\
&\quad - \sum_{s=1}^{|\mathcal{P}^k|} \frac{4}{\eta_\gamma} \|\gamma_s^k - \gamma_s^{k-1}\|^2 + \frac{16}{\eta_\gamma (c_\gamma^k)^2} \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \|\mathbf{v}_l^{k+1} - \mathbf{v}_l^k\|^2 + \frac{64w_q^2}{\eta_\gamma (c_\gamma^k)^2} \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2.
\end{aligned} \tag{82}$$

Based on the setting of  $c_\lambda^k, c_\alpha^k, c_\beta^k$  and  $c_\gamma^k$ , we can obtain that,  $\frac{\eta_\lambda}{10} \geq \frac{1}{c_\lambda^{k+1}} - \frac{1}{c_\lambda^k}$ ,  $\frac{\eta_\alpha}{10} \geq \frac{1}{c_\alpha^{k+1}} - \frac{1}{c_\alpha^k}$ ,  $\frac{\eta_\beta}{10} \geq \frac{1}{c_\beta^{k+1}} - \frac{1}{c_\beta^k}$ ,  $\frac{\eta_\gamma}{10} \geq \frac{1}{c_\gamma^{k+1}} - \frac{1}{c_\gamma^k}$ ,  $\forall k \geq K_1$ . In addition, according to the definition of  $\mathbf{B}_l$ , the following inequality can be obtained,

$$\sum_{l=1}^L \|\mathbf{B}_l \mathbf{v}^{k+1} - \mathbf{B}_l \mathbf{v}^k\|^2 = \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2. \tag{83}$$

According to the definition of  $a_5$ , combining (79)-(83) with (74), we can obtain that,

$$\begin{aligned}
F^{k+1} - F^k &\leq \left( a_5 - \frac{1}{\eta_\mu^k} + \frac{16}{\eta_\lambda (c_\lambda^k)^2} + \frac{16}{\eta_\alpha (c_\alpha^k)^2} + \frac{16}{\eta_\beta (c_\beta^k)^2} \right) \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 \\
&\quad + \left( a_5 - \frac{1}{\eta_r^k} + \frac{64w_r^2}{\eta_\lambda (c_\lambda^k)^2} \right) \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
&\quad + \left( a_5 - \frac{1}{\eta_p^k} + \frac{64w_p^2}{\eta_\alpha (c_\alpha^k)^2} \right) \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
&\quad + \left( a_5 - \frac{1}{\eta_q^k} + \frac{64w_q^2}{\eta_\gamma (c_\gamma^k)^2} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 \\
&\quad + \left( a_5 - \frac{1}{\eta_v^k} + \frac{16}{\eta_\beta (c_\beta^k)^2} + \frac{16}{\eta_\gamma (c_\gamma^k)^2} \sum_{s=1}^{|\mathcal{P}^k|} \|\mathbf{b}_s\|^2 \right) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\
&\quad - \frac{1}{10\eta_\lambda} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \sum_{l=1}^L \|\lambda_l^{k+1}\|^2 + \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-2}}{c_\lambda^{k-1}} - \frac{c_\lambda^{k-1}}{c_\lambda^k} \right) \sum_{l=1}^L \|\lambda_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\alpha} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 + \frac{c_\alpha^{k-1} - c_\alpha^k}{2} \sum_{l=1}^L \|\alpha_l^{k+1}\|^2 + \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-2}}{c_\alpha^{k-1}} - \frac{c_\alpha^{k-1}}{c_\alpha^k} \right) \sum_{l=1}^L \|\alpha_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\beta} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 + \frac{c_\beta^{k-1} - c_\beta^k}{2} \sum_{l=1}^L \|\beta_l^{k+1}\|^2 + \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-2}}{c_\beta^{k-1}} - \frac{c_\beta^{k-1}}{c_\beta^k} \right) \sum_{l=1}^L \|\beta_l^k\|^2 \\
&\quad - \frac{1}{10\eta_\gamma} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + \frac{c_\gamma^{k-1} - c_\gamma^k}{2} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1}\|^2 + \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k\|^2,
\end{aligned} \tag{84}$$



which concludes the proof of Lemma 3.  $\square$

Finally, we provide the proof of Theorem 2.

*Proof.* First, we set

$$a_6^k = \min\left\{\frac{16}{\eta_\lambda(c_\lambda^k)^2} + \frac{16}{\eta_\alpha(c_\alpha^k)^2} + \frac{16}{\eta_\beta(c_\beta^k)^2}, \frac{64w_r^2}{\eta_\lambda(c_\lambda^k)^2}, \frac{64w_p^2}{\eta_\alpha(c_\alpha^k)^2}, \frac{64w_q^2}{\eta_\gamma(c_\gamma^k)^2}, \frac{16}{\eta_\beta(c_\beta^k)^2} + \frac{16}{\eta_\gamma(c_\gamma^k)^2} \sum_{s=1}^L \|b_s\|^2\right\} \frac{\xi-2}{2} - a_5. \quad (85)$$

where constant  $\xi > 2$  and satisfies

$$\min\left\{\frac{16}{\eta_\lambda(c_\lambda^0)^2} + \frac{16}{\eta_\alpha(c_\alpha^0)^2} + \frac{16}{\eta_\beta(c_\beta^0)^2}, \frac{64w_r^2}{\eta_\lambda(c_\lambda^0)^2}, \frac{64w_p^2}{\eta_\alpha(c_\alpha^0)^2}, \frac{64w_q^2}{\eta_\gamma(c_\gamma^0)^2}, \frac{16}{\eta_\beta(c_\beta^0)^2} + \frac{16}{\eta_\gamma(c_\gamma^0)^2} \sum_{s=1}^L \|b_s\|^2\right\} \frac{\xi-2}{2} > a_5. \quad (86)$$

Thus, we have  $a_6^k > 0, \forall k$ . According to the setting of  $\eta_\lambda^k, \eta_\alpha^k, \eta_\beta^k, \eta_\gamma^k$  and  $c_\lambda^k, c_\alpha^k, c_\beta^k, c_\gamma^k$ , we have,

$$\begin{aligned} a_5 - \frac{1}{\eta_\mu^k} + \frac{16}{\eta_\lambda(c_\lambda^k)^2} + \frac{16}{\eta_\alpha(c_\alpha^k)^2} + \frac{16}{\eta_\beta(c_\beta^k)^2} &\leq -a_6^k, \\ a_5 - \frac{1}{\eta_r^k} + \frac{64w_r^2}{\eta_\lambda(c_\lambda^k)^2} &\leq -a_6^k, \\ a_5 - \frac{1}{\eta_p^k} + \frac{64w_p^2}{\eta_\alpha(c_\alpha^k)^2} &\leq -a_6^k, \\ (a_5 - \frac{1}{\eta_q^k} + \frac{64w_q^2}{\eta_\gamma(c_\gamma^k)^2}) &\leq -a_6^k, \\ a_5 - \frac{1}{\eta_v^k} + \frac{16}{\eta_\beta(c_\beta^k)^2} + \frac{16}{\eta_\gamma(c_\gamma^k)^2} \sum_{s=1}^L \|b_s\|^2 &\leq -a_6^k. \end{aligned} \quad (87)$$

Combining it with Lemma 3,  $\forall k \geq K_1$ , we can obtain that,

$$\begin{aligned} a_6^k &\left( \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 + \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \right) \\ &\leq F^k - F^{k+1} - \frac{1}{10\eta_\lambda} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} \sum_{l=1}^L \|\lambda_l^{k+1}\|^2 + \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-2}}{c_\lambda^{k-1}} - \frac{c_\lambda^{k-1}}{c_\lambda^k} \right) \sum_{l=1}^L \|\lambda_l^k\|^2 \\ &\quad - \frac{1}{10\eta_\alpha} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 + \frac{c_\alpha^{k-1} - c_\alpha^k}{2} \sum_{l=1}^L \|\alpha_l^{k+1}\|^2 + \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-2}}{c_\alpha^{k-1}} - \frac{c_\alpha^{k-1}}{c_\alpha^k} \right) \sum_{l=1}^L \|\alpha_l^k\|^2 \\ &\quad - \frac{1}{10\eta_\beta} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 + \frac{c_\beta^{k-1} - c_\beta^k}{2} \sum_{l=1}^L \|\beta_l^{k+1}\|^2 + \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-2}}{c_\beta^{k-1}} - \frac{c_\beta^{k-1}}{c_\beta^k} \right) \sum_{l=1}^L \|\beta_l^k\|^2 \\ &\quad - \frac{1}{10\eta_\gamma} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + \frac{c_\gamma^{k-1} - c_\gamma^k}{2} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1}\|^2 + \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k\|^2. \end{aligned} \quad (88)$$

Given the definition of  $\nabla \tilde{G}^k$ , we have that,

$$\begin{aligned} (\nabla \tilde{G}^k)_{\mu_l} &= \nabla_{\mu_l} \tilde{L}_p(\{\mu_l^{\hat{k}_l}\}, \{\mathbf{r}_l^{\hat{k}_l}\}, \{\mathbf{p}_l^{\hat{k}_l}\}, \{q_s^{\hat{k}_l}\}, \mathbf{v}^{\hat{k}_l}, \{\lambda_l^{\hat{k}_l}\}, \{\alpha_l^{\hat{k}_l}\}, \{\beta_l^{\hat{k}_l}\}, \{\gamma_s^{\hat{k}_l}\}) \\ &\quad + \nabla_{\mu_l} \tilde{L}_p(\{\mu_l^{\hat{k}_l}\}, \{\mathbf{r}_l^{\hat{k}_l}\}, \{\mathbf{p}_l^{\hat{k}_l}\}, \{q_s^{\hat{k}_l}\}, \mathbf{v}^{\hat{k}_l}, \{\lambda_l^{\hat{k}_l}\}, \{\alpha_l^{\hat{k}_l}\}, \{\beta_l^{\hat{k}_l}\}, \{\gamma_s^{\hat{k}_l}\}) \\ &\quad - \nabla_{\mu_l} \tilde{L}_p(\{\mu_l^{\hat{k}_l}\}, \{\mathbf{r}_l^{\hat{k}_l}\}, \{\mathbf{p}_l^{\hat{k}_l}\}, \{q_s^{\hat{k}_l}\}, \mathbf{v}^{\hat{k}_l}, \{\lambda_l^{\hat{k}_l}\}, \{\alpha_l^{\hat{k}_l}\}, \{\beta_l^{\hat{k}_l}\}, \{\gamma_s^{\hat{k}_l}\}). \end{aligned} \quad (89)$$

Combining it with (41), we have that,

$$\|(\nabla \tilde{G}^k)_{\mu_l}\|^2 \leq \frac{1}{\eta_\mu^2} \|\mu_l^{\hat{k}_l} - \mu_l^k\|^2. \quad (90)$$

Similar results can be derived for other variables,

$$\begin{aligned}
\|(\nabla \tilde{G}^k)_{\mathbf{r}_l}\|^2 &\leq \frac{1}{\eta_{\mathbf{r}}^2} \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2. \\
\|(\nabla \tilde{G}^k)_{\mathbf{p}_l}\|^2 &\leq \frac{1}{\eta_{\mathbf{p}}^2} \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2. \\
\|(\nabla \tilde{G}^k)_{q_s}\|^2 &\leq \frac{1}{\eta_q^2} \|q_s^{k+1} - q_s^k\|^2. \\
\|(\nabla \tilde{G}^k)_{\mathbf{v}}\|^2 &\leq \frac{1}{\eta_{\mathbf{v}}^2} \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2.
\end{aligned} \tag{91}$$

According to the definition 3, we have,

$$\begin{aligned}
(\nabla \tilde{G}^k)_{\lambda_l} &= \nabla_{\lambda_l} \tilde{L}_p(\{\boldsymbol{\mu}_l^{\bar{k}_l}\}, \{\mathbf{r}_l^{\bar{k}_l}\}, \{\mathbf{p}_l^{\bar{k}_l}\}, \{q_s^{\bar{k}_l}\}, \mathbf{v}^{\bar{k}_l}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\
&\quad + \nabla_{\lambda_l} \tilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{\bar{k}_l}\}, \{\boldsymbol{\alpha}_l^{\bar{k}_l}\}, \{\boldsymbol{\beta}_l^{\bar{k}_l}\}, \{\gamma_s^{\bar{k}_l}\}) \\
&\quad - \nabla_{\lambda_l} \tilde{L}_p(\{\boldsymbol{\mu}_l^{\bar{k}_l}\}, \{\mathbf{r}_l^{\bar{k}_l}\}, \{\mathbf{p}_l^{\bar{k}_l}\}, \{q_s^{\bar{k}_l}\}, \mathbf{v}^{\bar{k}_l}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}).
\end{aligned} \tag{92}$$

Combining trigonometric inequality, (26) with Assumption 1, we can obtain

$$\begin{aligned}
\|(\nabla \tilde{G}^k)_{\lambda_l}\|^2 &\leq 3\|\nabla_{\lambda_l} \tilde{L}_p(\{\boldsymbol{\mu}_l^{\bar{k}_l}\}, \{\mathbf{r}_l^{\bar{k}_l}\}, \{\mathbf{p}_l^{\bar{k}_l}\}, \{q_s^{\bar{k}_l}\}, \mathbf{v}^{\bar{k}_l}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\})\|^2 \\
&\quad + 3((c_{\lambda}^{\bar{k}_l-1})^2 - (c_{\lambda}^{\bar{k}_l-1})^2)\|\boldsymbol{\lambda}_l^k\|^2 + 3\|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k + \mathbf{r}_l^k \circ \mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^{\bar{k}_l} \circ \mathbf{r}_l^k\|^2 \\
&\leq \frac{3}{\eta_{\lambda}^2} \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 + 3((c_{\lambda}^{\bar{k}_l-1})^2 - (c_{\lambda}^{\bar{k}_l-1})^2)\|\boldsymbol{\lambda}_l^k\|^2 + 6\|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 + 24w_{\mathbf{r}}^2 \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2.
\end{aligned} \tag{93}$$

Similar results can be derived for other variables as well,

$$\begin{aligned}
\|(\nabla \tilde{G}^k)_{\alpha_l}\|^2 &\leq \frac{3}{\eta_{\alpha}^2} \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 + 3((c_{\alpha}^{\bar{k}_l-1})^2 - (c_{\alpha}^{\bar{k}_l-1})^2)\|\boldsymbol{\alpha}_l^k\|^2 + 6\|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 + 24w_{\mathbf{p}}^2 \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2. \\
\|(\nabla \tilde{G}^k)_{\beta_l}\|^2 &\leq \frac{3}{\eta_{\beta}^2} \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 + 3((c_{\beta}^{\bar{k}_l-1})^2 - (c_{\beta}^{\bar{k}_l-1})^2)\|\boldsymbol{\beta}_l^k\|^2 + 6\|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 + 6\|\mathbf{v}^{\bar{k}_l} - \mathbf{v}^k\|^2. \\
\|(\nabla \tilde{G}^k)_{\gamma_s}\|^2 &\leq \frac{3}{\eta_{\gamma}^2} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + 3((c_{\gamma}^{k+1})^2 - (c_{\gamma}^k)^2)\|\gamma_s^k\|^2 + 6\|\mathbf{b}_s\|^2 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 + 24w_q^2 \|q_s^{k+1} - q_s^k\|^2.
\end{aligned} \tag{94}$$

According to Assumption 1, we have,

$$\|\mathbf{v}^{\bar{k}_l} - \mathbf{v}^k\|^2 \leq \tau k_1 \vartheta \leq \tau k_1 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2. \tag{95}$$

Combining it with (90)-(94), we can obtain that,

$$\begin{aligned}
\|\nabla \tilde{G}^k\|^2 &= \sum_{l=1}^L (\|(\nabla \tilde{G}^k)_{\boldsymbol{\mu}_l}\|^2 + \|(\nabla \tilde{G}^k)_{\mathbf{r}_l}\|^2 + \|(\nabla \tilde{G}^k)_{\mathbf{p}_l}\|^2 + \|(\nabla \tilde{G}^k)_{\lambda_l}\|^2 + \|(\nabla \tilde{G}^k)_{\alpha_l}\|^2 + \|(\nabla \tilde{G}^k)_{\beta_l}\|^2) \\
&\quad + \sum_{s=1}^{|\mathcal{P}^k|} \|(\nabla \tilde{G}^k)_{q_s}\|^2 + \|(\nabla \tilde{G}^k)_{\mathbf{v}}\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \|(\nabla \tilde{G}^k)_{\gamma_s}\|^2 \\
&\leq (\frac{1}{\eta_{\boldsymbol{\mu}}^2} + 18) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 + (\frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2) \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 + (\frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2) \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 \\
&\quad + (\frac{1}{\eta_q^2} + 24w_q^2) \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 + (\frac{1}{\eta_{\mathbf{v}}^2} + \sum_{s=1}^{|\mathcal{P}^k|} 6\|\mathbf{b}_s\|^2 + 6L\tau k_1) \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \\
&\quad + \frac{3}{\eta_{\lambda}^2} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\lambda}^{\bar{k}_l-1})^2 - (c_{\lambda}^{\bar{k}_l-1})^2) \|\boldsymbol{\lambda}_l^k\|^2 \\
&\quad + \frac{3}{\eta_{\alpha}^2} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\alpha}^{\bar{k}_l-1})^2 - (c_{\alpha}^{\bar{k}_l-1})^2) \|\boldsymbol{\alpha}_l^k\|^2 \\
&\quad + \frac{3}{\eta_{\beta}^2} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\beta}^{\bar{k}_l-1})^2 - (c_{\beta}^{\bar{k}_l-1})^2) \|\boldsymbol{\beta}_l^k\|^2 \\
&\quad + \frac{3}{\eta_{\gamma}^2} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + 3 \sum_{s=1}^{|\mathcal{P}^k|} ((c_{\gamma}^{k+1})^2 - (c_{\gamma}^k)^2) \|\gamma_s^k\|^2.
\end{aligned} \tag{96}$$

Let constant  $\underline{a}_6$  denote the lower bound of  $a_6^k$  ( $\underline{a}_6 > 0$ ), and we set constants  $d_1, d_2, d_3, d_4, d_5$  that,

$$\begin{aligned}
d_1 &= \frac{k_\tau \tau + 18k_\tau \tau \eta_\mu^2}{\eta_\mu^2 (a_6)^2} \geq \frac{k_\tau \tau + 18k_\tau \tau \eta_\mu^2}{\eta_\mu^2 (a_6^k)^2}, \\
d_2 &= \frac{k_\tau \tau + 24w_{\mathbf{r}}^2 k_\tau \tau \eta_{\mathbf{r}}^2}{\eta_{\mathbf{r}}^2 (a_6)^2} \geq \frac{k_\tau \tau + 24w_{\mathbf{r}}^2 k_\tau \tau \eta_{\mathbf{r}}^2}{\eta_{\mathbf{r}}^2 (a_6^k)^2}, \\
d_3 &= \frac{k_\tau \tau + 24w_{\mathbf{p}}^2 k_\tau \tau \eta_{\mathbf{p}}^2}{\eta_{\mathbf{p}}^2 (a_6)^2} \geq \frac{k_\tau \tau + 24w_{\mathbf{p}}^2 k_\tau \tau \eta_{\mathbf{p}}^2}{\eta_{\mathbf{p}}^2 (a_6^k)^2}, \\
d_4 &= \frac{1 + 24w_{\mathbf{q}}^2 \eta_{\mathbf{q}}^2}{\eta_{\mathbf{q}}^2 (a_6)^2} \geq \frac{1 + 24w_{\mathbf{q}}^2 \eta_{\mathbf{q}}^2}{\eta_{\mathbf{q}}^2 (a_6^k)^2}, \\
d_5 &= \frac{1 + (\sum_{s=1}^{|\mathcal{P}^k|} 6\|\mathbf{b}_s\|^2 + 6L\tau k_1) \eta_{\mathbf{v}}^2}{\eta_{\mathbf{v}}^2 (a_6)^2} \geq \frac{1 + (\sum_{s=1}^{|\mathcal{P}^k|} 6\|\mathbf{b}_s\|^2 + 6L\tau k_1) \eta_{\mathbf{v}}^2}{\eta_{\mathbf{z}}^2 (a_6^k)^2}.
\end{aligned} \tag{97}$$

where  $k_\tau$  is a positive constant and it satisfies that,

$$k_\tau \geq \max \left\{ \frac{\overline{d_5}(\frac{1}{\eta_\mu^2} + 18)}{\underline{d_5}(\frac{1}{\eta_\mu^2} + 18)}, \frac{\overline{d_5}(\frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2)}{\underline{d_5}(\frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2)}, \frac{\overline{d_5}(\frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2)}{\underline{d_5}(\frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2)} \right\}, \tag{98}$$

where  $\overline{\eta_\mu}$ ,  $\overline{\eta_{\mathbf{r}}}$ , and  $\overline{\eta_{\mathbf{p}}}$  are the upper bounds of  $\eta_\mu^k$ ,  $\eta_{\mathbf{r}}^k$ , and  $\eta_{\mathbf{p}}^k$ , respectively. By employing (96) and (97), we can obtain,

$$\begin{aligned}
\|\nabla \tilde{G}^k\|^2 &\leq (a_6^k)^2 \left( d_1 \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + d_2 \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + d_3 \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \right) \\
&+ (a_6^k)^2 \left( d_4 \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 + d_5 \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \right) \\
&+ \left( \frac{1}{\eta_\mu^2} + 18 \right) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 - \left( \frac{1}{\eta_\mu^2} + 18 \right) k_\tau \tau \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 \\
&+ \left( \frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2 \right) \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 - \left( \frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
&+ \left( \frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2 \right) \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 - \left( \frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
&+ \frac{3}{\eta_\lambda^2} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\boldsymbol{\lambda}}^{\bar{k}_l-1})^2 - (c_{\boldsymbol{\lambda}}^{k-1})^2) \|\boldsymbol{\lambda}_l^k\|^2 \\
&+ \frac{3}{\eta_\alpha^2} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\boldsymbol{\alpha}}^{\bar{k}_l-1})^2 - (c_{\boldsymbol{\alpha}}^{k-1})^2) \|\boldsymbol{\alpha}_l^k\|^2 \\
&+ \frac{3}{\eta_\beta^2} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 + 3 \sum_{l=1}^L ((c_{\boldsymbol{\beta}}^{\bar{k}_l-1})^2 - (c_{\boldsymbol{\beta}}^{k-1})^2) \|\boldsymbol{\beta}_l^k\|^2 \\
&+ \frac{3}{\eta_\gamma^2} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + 3 \sum_{s=1}^{|\mathcal{P}^k|} ((c_\gamma^{\bar{k}_l-1})^2 - (c_\gamma^{k-1})^2) \|\gamma_s^k\|^2.
\end{aligned} \tag{99}$$

Let  $d_6^k$  denote a nonnegative sequence, i.e.,

$$d_6^k = \left( \max \left\{ d_1 a_6^k, d_2 a_6^k, d_3 a_6^k, d_4 a_6^k, d_5 a_6^k, \frac{30\tau}{\eta_\lambda}, \frac{30\tau}{\eta_\alpha}, \frac{30\tau}{\eta_\beta}, \frac{30}{\eta_\gamma} \right\} \right)^{-1}. \tag{100}$$

We denote the upper and lower bound of  $d_6$  as  $\bar{d}_6$  and  $\underline{d}_6$ , respectively. Combining (99) with the definition of  $d_6^k$ , we can

obtain that,

$$\begin{aligned}
d_6^k \|\nabla \tilde{G}^k\|^2 \leq & d_6^k \left( \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \|q_s^{k+1} - q_s^k\|^2 + \|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2 \right) \\
& + d_6^k \left( \frac{1}{\eta_\mu^2} + 18 \right) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 - d_6^k \left( \frac{1}{\eta_\mu^2} + 18 \right) k_\tau \tau \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 \\
& + d_6^k \left( \frac{1}{\eta_r^2} + 24w_r^2 \right) \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 - d_6^k \left( \frac{1}{\eta_r^2} + 24w_r^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + d_6^k \left( \frac{1}{\eta_p^2} + 24w_p^2 \right) \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 - d_6^k \left( \frac{1}{\eta_p^2} + 24w_p^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\lambda} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 + 3d_6^k \sum_{l=1}^L ((c_\lambda^{\bar{k}_l-1})^2 - (c_\lambda^{\bar{k}_l-1})^2) \|\boldsymbol{\lambda}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\alpha} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 + 3d_6^k \sum_{l=1}^L ((c_\alpha^{\bar{k}_l-1})^2 - (c_\alpha^{\bar{k}_l-1})^2) \|\boldsymbol{\alpha}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\beta} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 + 3d_6^k \sum_{l=1}^L ((c_\beta^{\bar{k}_l-1})^2 - (c_\beta^{\bar{k}_l-1})^2) \|\boldsymbol{\beta}_l^k\|^2 \\
& + \frac{1}{10\eta_\gamma} \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^{k+1} - \gamma_s^k\|^2 + 3d_6^k \sum_{s=1}^{|\mathcal{P}^k|} ((c_\gamma^{k-1})^2 - (c_\gamma^k)^2) \|\gamma_s^k\|^2.
\end{aligned} \tag{101}$$

According to Assumption 1 and combining (88) with (101), we have,

$$\begin{aligned}
d_6^k \|\nabla \tilde{G}^k\|^2 \leq & F^k - F^{k+1} + \bar{d}_6 \left( \frac{1}{\eta_\mu^2} + 18 \right) \sum_{l=1}^L \|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2 - \underline{d}_6 \left( \frac{1}{\eta_\mu^2} + 18 \right) k_\tau \tau \sum_{l=1}^L \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 \\
& + \bar{d}_6 \left( \frac{1}{\eta_r^2} + 24w_r^2 \right) \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 - \underline{d}_6 \left( \frac{1}{\eta_r^2} + 24w_r^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + \bar{d}_6 \left( \frac{1}{\eta_p^2} + 24w_p^2 \right) \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 - \underline{d}_6 \left( \frac{1}{\eta_p^2} + 24w_p^2 \right) k_\tau \tau \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\lambda} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 - \frac{1}{10\eta_\lambda} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\alpha} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 - \frac{1}{10\eta_\alpha} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 \\
& + \frac{1}{10\tau\eta_\beta} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 - \frac{1}{10\eta_\beta} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 \\
& + 3\bar{d}_6 \sum_{l=1}^L ((c_\lambda^{\bar{k}_l-1})^2 - (c_\lambda^{\bar{k}_l-1})^2) \|\boldsymbol{\lambda}_l^k\|^2 + 3\bar{d}_6 \sum_{l=1}^L ((c_\alpha^{\bar{k}_l-1})^2 - (c_\alpha^{\bar{k}_l-1})^2) \|\boldsymbol{\alpha}_l^k\|^2 \\
& + 3\bar{d}_6 \sum_{l=1}^L ((c_\beta^{\bar{k}_l-1})^2 - (c_\beta^{\bar{k}_l-1})^2) \|\boldsymbol{\beta}_l^k\|^2 + 3\bar{d}_6 \sum_{s=1}^{|\mathcal{P}^k|} ((c_\gamma^{k-1})^2 - (c_\gamma^k)^2) \|\gamma_s^k\|^2 \\
& + \frac{c_\lambda^{k-1} - c_\lambda^k}{2} Mw_\lambda^2 + \frac{c_\alpha^{k-1} - c_\alpha^k}{2} Mw_\alpha^2 + \frac{c_\beta^{k-1} - c_\beta^k}{2} Mw_\beta^2 + \frac{c_\gamma^{k-1} - c_\gamma^k}{2} Pw_\gamma^2 \\
& + \frac{4}{\eta_\lambda} \left( \frac{c_\lambda^{k-2}}{c_\lambda^{k-1}} - \frac{c_\lambda^{k-1}}{c_\lambda^k} \right) \sum_{l=1}^L \|\boldsymbol{\lambda}_l^k\|^2 + \frac{4}{\eta_\alpha} \left( \frac{c_\alpha^{k-2}}{c_\alpha^{k-1}} - \frac{c_\alpha^{k-1}}{c_\alpha^k} \right) \sum_{l=1}^L \|\boldsymbol{\alpha}_l^k\|^2 \\
& + \frac{4}{\eta_\beta} \left( \frac{c_\beta^{k-2}}{c_\beta^{k-1}} - \frac{c_\beta^{k-1}}{c_\beta^k} \right) \sum_{l=1}^L \|\boldsymbol{\beta}_l^k\|^2 + \frac{4}{\eta_\gamma} \left( \frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k} \right) \sum_{s=1}^{|\mathcal{P}^k|} \|\gamma_s^k\|^2.
\end{aligned} \tag{102}$$

Denoting  $\tilde{K}(\epsilon)$  as  $\tilde{K}(\epsilon) = \min\{k \mid \|\nabla \tilde{G}^{K_1+k}\|^2 \leq \frac{\epsilon}{4}, k \geq 2\}$ . Summing up (102) from  $K_1 + 2$  to  $K_1 + \tilde{K}(\epsilon)$ , we can obtain that,

$$\begin{aligned}
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} d_6^k \|\nabla \tilde{G}^k\|^2 &\leq F^{K_1+2} - F^{K_1+\tilde{K}(\epsilon)+1} + \frac{c_\lambda^1}{2} M w_\lambda^2 + \frac{c_\alpha^1}{2} M w_\alpha^2 + \frac{c_\beta^1}{2} M w_\beta^2 + \frac{c_\gamma^1}{2} P w_\gamma^2 \\
&+ \frac{4}{\eta_\lambda} \frac{c_\lambda^0}{c_\lambda^1} M w_\lambda^2 + \frac{4}{\eta_\alpha} \frac{c_\alpha^0}{c_\alpha^1} M w_\alpha^2 + \frac{4}{\eta_\beta} \frac{c_\beta^0}{c_\beta^1} M w_\beta^2 + \frac{4}{\eta_\gamma} \frac{c_\gamma^0}{c_\gamma^1} P w_\gamma^2 \\
&+ \overline{d_6} \left( \frac{1}{\eta_\mu^2} + 18 \right) \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mu_l^{\bar{k}_l} - \mu_l^k\|^2 - \underline{d_6} \left( \frac{1}{\eta_\mu^2} + 18 \right) k_\tau \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 \\
&+ \overline{d_6} \left( \frac{1}{\eta_r^2} + 24 w_r^2 \right) \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 - \underline{d_6} \left( \frac{1}{\eta_r^2} + 24 w_r^2 \right) k_\tau \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
&+ \overline{d_6} \left( \frac{1}{\eta_p^2} + 24 w_p^2 \right) \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 - \underline{d_6} \left( \frac{1}{\eta_p^2} + 24 w_p^2 \right) k_\tau \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
&+ \frac{1}{10\tau\eta_\lambda} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\lambda_l^{\bar{k}_l} - \lambda_l^k\|^2 - \frac{1}{10\eta_\lambda} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\lambda_l^{k+1} - \lambda_l^k\|^2 \\
&+ \frac{1}{10\tau\eta_\alpha} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\alpha_l^{\bar{k}_l} - \alpha_l^k\|^2 - \frac{1}{10\eta_\alpha} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\alpha_l^{k+1} - \alpha_l^k\|^2 \\
&+ \frac{1}{10\tau\eta_\beta} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\beta_l^{\bar{k}_l} - \beta_l^k\|^2 - \frac{1}{10\eta_\beta} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\beta_l^{k+1} - \beta_l^k\|^2 \\
&+ 3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\lambda^{\hat{k}_l-1})^2 - (c_\lambda^{\bar{k}_l-1})^2) \|\lambda_l^k\|^2 + 3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\alpha^{\hat{k}_l-1})^2 - (c_\alpha^{\bar{k}_l-1})^2) \|\alpha_l^k\|^2 \\
&+ 3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\beta^{\hat{k}_l-1})^2 - (c_\beta^{\bar{k}_l-1})^2) \|\beta_l^k\|^2 + 3\overline{d_6} (c_\gamma^1)^2 P w_\gamma^2.
\end{aligned} \tag{103}$$

For each worker  $l$ , we have that  $\bar{k}_l - \hat{k}_l \leq \tau$ , thus,

$$\begin{aligned}
3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\lambda^{\hat{k}_l-1})^2 - (c_\lambda^{\bar{k}_l-1})^2) \|\lambda_l^k\|^2 &\leq 3\tau\overline{d_6} (c_\lambda^1)^2 M w_\lambda^2, \\
3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\alpha^{\hat{k}_l-1})^2 - (c_\alpha^{\bar{k}_l-1})^2) \|\alpha_l^k\|^2 &\leq 3\tau\overline{d_6} (c_\alpha^1)^2 M w_\alpha^2, \\
3\overline{d_6} \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L ((c_\beta^{\hat{k}_l-1})^2 - (c_\beta^{\bar{k}_l-1})^2) \|\beta_l^k\|^2 &\leq 3\tau\overline{d_6} (c_\beta^1)^2 M w_\beta^2.
\end{aligned} \tag{104}$$

In our asynchronous algorithm, inactive workers do not update their variables in each master iteration. Thus, for any  $k$  which satisfies  $\hat{v}_l(j-1) \leq k < \hat{v}_l(j)$ , we have  $\mu_l^k = \mu_l^{\hat{v}_l(j)-1}$ . And for  $k \notin \mathcal{V}_l(K)$ , we have  $\|\mu_l^k - \mu_l^{k-1}\|^2 = 0$ . Since  $\hat{v}_l(j) - \hat{v}_l(j-1) \leq \tau$ , we can obtain

$$\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mu_l^{\bar{k}_l} - \mu_l^k\|^2 \leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mu_l^{k+1} - \mu_l^k\|^2 + 4\tau(\tau-1) M w_\mu^2. \tag{105}$$

Similarly, we can obtain

$$\begin{aligned}
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 &\leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + 4\tau(\tau-1)Mw_{\mathbf{r}}^2, \\
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 &\leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + 4\tau(\tau-1)Mw_{\mathbf{p}}^2, \\
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{\bar{k}_l} - \boldsymbol{\lambda}_l^k\|^2 &\leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\|^2 + 4\tau(\tau-1)Mw_{\boldsymbol{\lambda}}^2, \\
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{\bar{k}_l} - \boldsymbol{\alpha}_l^k\|^2 &\leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\alpha}_l^{k+1} - \boldsymbol{\alpha}_l^k\|^2 + 4\tau(\tau-1)Mw_{\boldsymbol{\alpha}}^2, \\
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{\bar{k}_l} - \boldsymbol{\beta}_l^k\|^2 &\leq \tau \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \sum_{l=1}^L \|\boldsymbol{\beta}_l^{k+1} - \boldsymbol{\beta}_l^k\|^2 + 4\tau(\tau-1)Mw_{\boldsymbol{\beta}}^2.
\end{aligned} \tag{106}$$

By employing (103), (104), (105) and (106), we can obtain

$$\begin{aligned}
\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} d_6^k \|\nabla \tilde{G}^k\|^2 &\leq F^{K_1+2} - F^{K_1+\tilde{K}(\epsilon)+1} \\
&+ \frac{c_{\boldsymbol{\lambda}}^1}{2} Mw_{\boldsymbol{\lambda}}^2 + \frac{c_{\boldsymbol{\alpha}}^1}{2} Mw_{\boldsymbol{\alpha}}^2 + \frac{c_{\boldsymbol{\beta}}^1}{2} Mw_{\boldsymbol{\beta}}^2 + \frac{c_{\gamma}^1}{2} Pw_{\gamma}^2 \\
&+ \frac{4}{\eta_{\boldsymbol{\lambda}}} \frac{c_{\boldsymbol{\lambda}}^0}{c_{\boldsymbol{\lambda}}^1} Mw_{\boldsymbol{\lambda}}^2 + \frac{4}{\eta_{\boldsymbol{\alpha}}} \frac{c_{\boldsymbol{\alpha}}^0}{c_{\boldsymbol{\alpha}}^1} Mw_{\boldsymbol{\alpha}}^2 + \frac{4}{\eta_{\boldsymbol{\beta}}} \frac{c_{\boldsymbol{\beta}}^0}{c_{\boldsymbol{\beta}}^1} Mw_{\boldsymbol{\beta}}^2 + \frac{4}{\eta_{\gamma}} \frac{c_{\gamma}^0}{c_{\gamma}^1} Pw_{\gamma}^2 \\
&+ 3\tau \bar{d}_6 (c_{\boldsymbol{\lambda}}^1)^2 Mw_{\boldsymbol{\lambda}}^2 + 3\tau \bar{d}_6 (c_{\boldsymbol{\alpha}}^1)^2 Mw_{\boldsymbol{\alpha}}^2 + 3\tau \bar{d}_6 (c_{\boldsymbol{\beta}}^1)^2 Mw_{\boldsymbol{\beta}}^2 + 3\tau \bar{d}_6 (c_{\gamma}^1)^2 Pw_{\gamma}^2 \\
&+ \left( \frac{2Mw_{\boldsymbol{\lambda}}^2}{5\eta_{\boldsymbol{\lambda}}} + \frac{2Mw_{\boldsymbol{\alpha}}^2}{5\eta_{\boldsymbol{\alpha}}} + \frac{2Mw_{\boldsymbol{\beta}}^2}{5\eta_{\boldsymbol{\beta}}} + 4\bar{d}_6 \left( \frac{1}{\eta_{\boldsymbol{\mu}}^2} + 18 \right) Mw_{\boldsymbol{\mu}}^2 \tau + 4\bar{d}_6 \left( \frac{1}{\eta_{\mathbf{r}}^2} + 24w_{\mathbf{r}}^2 \right) Mw_{\mathbf{r}}^2 \tau + 4\bar{d}_6 \left( \frac{1}{\eta_{\mathbf{p}}^2} + 24w_{\mathbf{p}}^2 \right) Mw_{\mathbf{p}}^2 \tau \right) (\tau-1) \\
&= (d_7 + k_d \tau)(\tau-1).
\end{aligned} \tag{107}$$

where  $d_7$  and  $k_d$  are constants. Set constant  $d_8$  as

$$d_8 = \left( \max \left\{ d_1, d_2, d_3, d_4, d_5, \frac{30\tau}{\eta_{\boldsymbol{\lambda}} a_6}, \frac{30\tau}{\eta_{\boldsymbol{\alpha}} a_6}, \frac{30\tau}{\eta_{\boldsymbol{\beta}} a_6}, \frac{30}{\eta_{\gamma} a_6} \right\} \right) \geq \frac{1}{d_6^k a_6^k}. \tag{108}$$

Thus, we can obtain that

$$\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \frac{1}{d_8 a_6^k} \|\nabla \tilde{G}^{K_1+\tilde{K}(\epsilon)}\|^2 \leq \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \frac{1}{d_8 a_6^k} \|\nabla \tilde{G}^k\|^2 \leq \sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} d_6^k \|\nabla \tilde{G}^k\|^2 \leq (d_7 + k_d \tau)(\tau-1). \tag{109}$$

Thus, we can obtain

$$\|\nabla \tilde{G}^{K_1+\tilde{K}(\epsilon)}\|^2 \leq \frac{(d_7 + k_d \tau)(\tau-1)d_8}{\sum_{k=K_1+2}^{K_1+\tilde{K}} \frac{1}{a_6^k}}. \tag{110}$$

According to the setting of  $c_{\boldsymbol{\lambda}}^k$ ,  $c_{\boldsymbol{\alpha}}^k$ ,  $c_{\boldsymbol{\beta}}^k$  and  $c_{\gamma}^k$ , we have,

$$\frac{1}{a_6^k} \geq \frac{1}{8(\xi-2)(k+1)^{\frac{1}{2}} \min\{\eta_{\boldsymbol{\lambda}} + \eta_{\boldsymbol{\alpha}} + \eta_{\boldsymbol{\beta}}, 4w_{\mathbf{r}}^2 \eta_{\boldsymbol{\lambda}}, 4w_{\mathbf{p}}^2 \eta_{\boldsymbol{\alpha}}, 4w_{\mathbf{q}}^2 \eta_{\gamma}, \eta_{\gamma} + \eta_{\gamma} \sum_{s=1}^L \|b_s\|^2\}}. \tag{111}$$

Summing up  $a_6^k$  from  $k = K_1 + 2$  to  $k = K_1 + \tilde{K}$ , we have

$$\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \frac{1}{a_6^k} \geq \frac{(K_1 + \tilde{K}(\epsilon))^{\frac{1}{2}} - (K_1 + 2)^{\frac{1}{2}}}{8(\xi-2) \min\{\eta_{\boldsymbol{\lambda}} + \eta_{\boldsymbol{\alpha}} + \eta_{\boldsymbol{\beta}}, 4w_{\mathbf{r}}^2 \eta_{\boldsymbol{\lambda}}, 4w_{\mathbf{p}}^2 \eta_{\boldsymbol{\alpha}}, 4w_{\mathbf{q}}^2 \eta_{\gamma}, \eta_{\gamma} + \eta_{\gamma} \sum_{s=1}^L \|b_s\|^2\}}. \tag{112}$$

Thus, plugging (112) into (110), we can obtain:

$$\|\nabla \tilde{G}^{K_1+\tilde{K}(\epsilon)}\|^2 \leq \frac{(d_7 + k_d \tau)(\tau-1)d_8}{\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)} \frac{1}{a_6^k}} \leq \frac{d_9(d_7 + k_d \tau)(\tau-1)d_8}{(K_1 + \tilde{K}(\epsilon))^{\frac{1}{2}} - (K_1 + 2)^{\frac{1}{2}}}, \tag{113}$$

where  $d_9 = 32(\xi - 2) \min\{\eta_\lambda + \eta_\alpha + \eta_\beta, 4w_{\mathbf{r}}^2\eta_\lambda, 4w_{\mathbf{p}}^2\eta_\alpha, 4w_{\mathbf{q}}^2\eta_\gamma, \eta_\gamma + \eta_\gamma \sum_{s=1}^L \|b_s\|^2\}$ . According to the definition of  $\tilde{K}(\epsilon)$ , we have:

$$K_1 + \tilde{K}(\epsilon) \geq \left( \frac{d_9(d_7 + k_d\tau)(\tau - 1)d_8}{\epsilon} + (K_1 + 2)^{\frac{1}{2}} \right)^2. \quad (114)$$

Combining the definition of  $\nabla G^k$  and  $\nabla \tilde{G}^k$  with trigonometric inequality, we then get:

$$\|\nabla G^k\| - \|\nabla \tilde{G}^k\| \leq \|\nabla G^k - \nabla \tilde{G}^k\| \leq \sqrt{\sum_{l=1}^L \|c_\lambda^{k-1} \lambda_l^k\|^2 + \sum_{l=1}^L \|c_\alpha^{k-1} \alpha_l^k\|^2 + \sum_{l=1}^L \|c_\beta^{k-1} \beta_l^k\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \|c_\gamma^{k-1} \gamma_s^k\|^2}. \quad (115)$$

If  $k \geq 16(\frac{Mw_\lambda^2}{\eta_\lambda^2} + \frac{Mw_\alpha^2}{\eta_\alpha^2} + \frac{Mw_\beta^2}{\eta_\beta^2} + \frac{Pw_\gamma^2}{\eta_\gamma^2})^2 \frac{1}{\epsilon^2}$ , we have

$$\sqrt{\sum_{l=1}^L \|c_\lambda^{k-1} \lambda_l^k\|^2 + \sum_{l=1}^L \|c_\alpha^{k-1} \alpha_l^k\|^2 + \sum_{l=1}^L \|c_\beta^{k-1} \beta_l^k\|^2 + \sum_{s=1}^{|\mathcal{P}^k|} \|c_\gamma^{k-1} \gamma_s^k\|^2} \leq \frac{\sqrt{\epsilon}}{2}. \quad (116)$$

Combining it with (114), we can conclude that

$$K(\epsilon) \sim \mathcal{O} \left( \max \left\{ \left( 16 \left( \frac{Mw_\lambda^2}{\eta_\lambda^2} + \frac{Mw_\alpha^2}{\eta_\alpha^2} + \frac{Mw_\beta^2}{\eta_\beta^2} + \frac{Pw_\gamma^2}{\eta_\gamma^2} \right)^2 \frac{1}{\epsilon^2}, \left( \frac{d_9(d_7 + k_d\tau)(\tau - 1)d_8}{\epsilon} + (K_1 + 2)^{\frac{1}{2}} \right)^2 \right\} \right). \quad (117)$$

which concludes our proof.  $\square$