Appendix E
Complete Proof of Theorem 2

In this section, we provide the comprehensive proof of Theorem 2. We will start by introducing several definitions. Following that, three lemmas are presented, which are crucial components in the proof of Theorem 2. Finally, we provide the complete proof of Theorem 2.

Definition 2. *Based on the definition of ∇G^k, we further define :*

$$\begin{aligned}
&(\nabla G^k)_{\boldsymbol{\mu}_l} = \nabla_{\boldsymbol{\mu}_l} 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\}),\\
&(\nabla G^k)_{\mathbf{r}_l} = \nabla_{\mathbf{r}_l} 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\}),\\
&(\nabla G^k)_{\mathbf{p}_l} = \nabla_{\mathbf{p}_l} 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\}),\\
&(\nabla G^k)_{q_s} = \nabla_{q_s} 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\}),\\
&(\nabla G^k)_{\mathbf{v}} = \nabla_{\mathbf{v}} 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\}),\\
&(\nabla G^k)_{\boldsymbol{\lambda}_l} = \nabla_{\lambda_l} 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\}),\\
&(\nabla G^k)_{\boldsymbol{\alpha}_l} = \nabla_{\alpha_l} 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\}),\\
&(\nabla G^k)_{\boldsymbol{\beta}_l} = \nabla_{\beta_l} 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\}),\\
&(\nabla G^k)_{\gamma_s} = \nabla_{\gamma_s} 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\}).
\end{aligned} \tag{E.1}$$

Definition 3. *Similar to the definition of ∇G^k, let $\nabla \widetilde{G}^k$ denote the gradient of $\widetilde{L}_p$ in the k-th iteration, i.e.,*

$$\nabla \widetilde{G}^k = \begin{bmatrix}
\{\nabla_{\boldsymbol{\mu}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{\mathbf{r}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{\mathbf{p}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{q_s} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\nabla_{\mathbf{v}} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\\
\{\nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{\boldsymbol{\alpha}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{\boldsymbol{\beta}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}\\
\{\nabla_{\gamma_s} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\mathbf{v}^k,\{q_s^k\},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\}
\end{bmatrix}. \tag{E.2}$$

And we further define :

$$\begin{aligned}
&(\nabla \widetilde{G}^k)_{\boldsymbol{\mu}_l} = \nabla_{\boldsymbol{\mu}_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\mathbf{r}_l} = \nabla_{\mathbf{r}_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\mathbf{p}_l} = \nabla_{\mathbf{p}_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{q_s} = \nabla_{q_s} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\mathbf{v}} = \nabla_{\mathbf{v}} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\boldsymbol{\lambda}_l} = \nabla_{\lambda_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\boldsymbol{\alpha}_l} = \nabla_{\alpha_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\boldsymbol{\beta}_l} = \nabla_{\beta_l} \widetilde{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\}),\\
&(\nabla \widetilde{G}^k)_{\gamma_s} = \nabla_{\gamma_s} \widetilde{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\}).
\end{aligned} \tag{E.3}$$

Definition 4. *In the k^{th} iteration of our algorithm, we define the last iteration in which the l^{th} worker was active as $\hat{k}_l$, and the next iteration in which the l^{th} worker will be active as $\overline{k}_l$. Furthermore, we represent the set of iteration indices in which the l^{th} worker is active during the $K_1 + K + \tau$ iteration as $\mathcal{V}_l(K)$. And the j^{th} element in $\mathcal{V}_l(K)$ is represented as $\hat{v}_l(j)$.*

Based on the above definitions, we next provide the proof of Lemma 1, Lemma 2 and Lemma 3

Lemma 1. *According to Eq. (21), function L_p has Lipschitz continuous Hessian and let L_w denote the Lipschitz constant. Based on definition of $\eta_{\boldsymbol{\mu}}$, $\eta_{\mathbf{v}}$, $\eta_{\mathbf{r}}$, $\eta_{\mathbf{p}}$ and η_q, we define $\eta_{\boldsymbol{\mu}}^k$, $\eta_{\mathbf{v}}^k$, $\eta_{\mathbf{r}}^k$, $\eta_{\mathbf{p}}^k$ and η_q^k as:*

$$\begin{aligned}
&\eta_{\boldsymbol{\mu}}^k = \frac{1}{\frac{8\xi}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2} + \frac{8\xi}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2} + \frac{8\xi}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2}}, \quad \eta_{\mathbf{v}}^k = \frac{1}{\frac{8\xi}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2} + \frac{8\xi}{\eta_{\gamma}(c_{\gamma}^k)^2}\sum_{s=1}^{L}\|b_s\|^2}, \\
&\eta_{\mathbf{r}}^k = \frac{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2}{32 w_{\mathbf{r}}^2 \xi}, \quad \eta_{\mathbf{p}}^k = \frac{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2}{32 w_{\mathbf{p}}^2 \xi}, \quad \eta_q^k = \frac{\eta_{\gamma}(c_{\gamma}^k)^2}{32 w_q^2 \xi}.
\end{aligned} \tag{38}$$

We can obtain that,

$$\begin{aligned}
&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) \\
&- L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) \\
&\le (\frac{L_w}{6} + 1 - \frac{1}{\eta_{\boldsymbol{\mu}}^k})\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + (-\frac{1}{\eta_{\mathbf{r}}^k} + \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} + w_{\boldsymbol{\lambda}})\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + (\frac{L_w}{6} - \frac{1}{\eta_{\mathbf{v}}^k})\|\mathbf{v}^{k+1} - \mathbf{v}^k\|^2
\end{aligned} \tag{39}$$

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

$$\begin{aligned}
&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) \\
&\le \sum_{l=1}^{L}\left(\left\langle \nabla_{\boldsymbol{\mu}_l} L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}), \boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k \right\rangle + \frac{L_w}{6}||\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k||^2\right) \\
&+ \frac{1}{2}\sum_{l=1}^{L}\left\langle \nabla_{\boldsymbol{\mu}_l}^2 L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})(\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k), \boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k \right\rangle.
\end{aligned} \tag{40}$$

According to (23) and the definition of $\eta_{\boldsymbol{\mu}}$, we have that,

$$\begin{aligned}
&\left\langle \nabla_{\boldsymbol{\mu}_l} L_p(\{\boldsymbol{\mu}_l^{\hat{k}_l}\},\{\mathbf{r}_l^{\hat{k}_l}\},\{\mathbf{p}_l^{\hat{k}_l}\},\{q_s^{\hat{k}_l}\},\mathbf{v}^{\hat{k}_l},\{\boldsymbol{\lambda}_l^{\hat{k}_l}\},\{\boldsymbol{\alpha}_l^{\hat{k}_l}\},\{\boldsymbol{\beta}_l^{\hat{k}_l}\},\{\gamma_s^{\hat{k}_l}\}), \boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k \right\rangle = -\frac{1}{\eta_{\boldsymbol{\mu}}}||\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k||^2 \\
&\le -\frac{1}{\eta_{\boldsymbol{\mu}}^k}||\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k||^2.
\end{aligned} \tag{41}$$

Given that $\boldsymbol{\mu}_l^k = \boldsymbol{\mu}_l^{\hat{k}_l}$, $\boldsymbol{\lambda}_l^k = \boldsymbol{\lambda}_l^{\hat{k}_l}$, $\boldsymbol{\alpha}_l^k = \boldsymbol{\alpha}_l^{\hat{k}_l}$, and $\boldsymbol{\beta}_l^k = \boldsymbol{\beta}_l^{\hat{k}_l}$, combining (41) with the definition of $(\nabla \widetilde{G}^k)_{\boldsymbol{\mu}_l}$, we have,

$$\left\langle \nabla_{\boldsymbol{\mu}_l} L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}), \boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k \right\rangle \le -\frac{1}{\eta_{\boldsymbol{\mu}}^k}||\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k||^2. \tag{42}$$

According to the definition of L_p, we have,

$$\frac{1}{2}\sum_{l=1}^{L}\left\langle \nabla_{\boldsymbol{\mu}_l}^2 L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})(\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k), \boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k \right\rangle = \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2. \tag{43}$$

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

$$\begin{aligned}
&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) - L_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}) \\
&\le (-\frac{1}{\eta_{\boldsymbol{\mu}}^k} + \frac{L_w}{6} + 1)\sum_{l=1}^{L}\left(||\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k||^2\right).
\end{aligned} \tag{44}$$

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

$$\begin{aligned}
&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\} - L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\} \\
&\le \sum_{l=1}^{L}\left\langle \nabla_{\mathbf{r}_l} L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\}), \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \right\rangle + \frac{L_w}{6}\sum_{l=1}^{L}||\mathbf{r}_l^{k+1} - \mathbf{r}_l^k||^3 \\
&+ \frac{1}{2}\sum_{l=1}^{L}\left\langle \nabla_{\mathbf{r}_l}^2 L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})(\mathbf{r}_l^{k+1} - \mathbf{r}_l^k), \mathbf{r}_l^{k+1} - \mathbf{r}_l^k \right\rangle,
\end{aligned} \tag{45}$$

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

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

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

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

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

$$\frac{L_w}{6}\sum_{l=1}^{L}||\mathbf{r}_l^{k+1}-\mathbf{r}_l^k||^3 \le \frac{w_{\mathbf{r}}L_w M^{\frac{1}{2}}}{3}\sum_{l=1}^{L}||\mathbf{r}_l^{k+1}-\mathbf{r}_l^k||^2. \tag{48}$$

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

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

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

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

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

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

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

Lemma 2. *$\forall k \ge K_1$, 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\})\\&\le(\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\frac{L_w}{6}+1-\frac{1}{\eta_{\boldsymbol{\mu}}^k})\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2+(\frac{4w_{\mathbf{r}}^2}{a_1}+\frac{w_{\mathbf{r}}L_wM^{\frac{1}{2}}}{3}+w_{\boldsymbol{\lambda}}-\frac{1}{\eta_{\mathbf{r}}^k})\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2\\&+(\frac{4w_{\mathbf{p}}^2}{a_2}+\frac{w_{\mathbf{p}}L_wM^{\frac{1}{2}}}{3}+w_{\boldsymbol{\alpha}}-\frac{1}{\eta_{\mathbf{p}}^k})\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2+(\frac{4w_q^2}{a_4}+\frac{w_qL_w}{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_{\mathbf{v}}^k}+\frac{L_w}{6})||\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\\&+(\frac{a_1}{2}-\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}+\frac{1}{2\eta_{\boldsymbol{\lambda}}})\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\\&+(\frac{a_2}{2}-\frac{c_{\boldsymbol{\alpha}}^{k-1}-c_{\boldsymbol{\alpha}}^k}{2}+\frac{1}{2\eta_{\boldsymbol{\alpha}}})\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\\&+(\frac{a_3}{2}-\frac{c_{\boldsymbol{\beta}}^{k-1}-c_{\boldsymbol{\beta}}^k}{2}+\frac{1}{2\eta_{\boldsymbol{\beta}}})\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\\&+(\frac{a_4}{2}-\frac{c_{\gamma}^{k-1}-c_{\gamma}^k}{2}+\frac{1}{2\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{53}$$

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

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

$$\left\langle \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k - \eta_{\boldsymbol{\lambda}} \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}), \boldsymbol{\lambda} - \boldsymbol{\lambda}_l^{k+1} \right\rangle \geq 0. \tag{54}$$

Let $\boldsymbol{\lambda} = \boldsymbol{\lambda}_l^k$, we can obtain:

$$\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) - \frac{1}{\eta_{\boldsymbol{\lambda}}}(\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k), \boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k+1} \right\rangle \leq 0. \tag{55}$$

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

$$\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}) - \frac{1}{\eta_{\boldsymbol{\lambda}}}(\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}), \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \leq 0. \tag{56}$$

Since $\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k = 0, l \notin \mathcal{Q}^{k+1}$, inequality (56) holds for l. According to (22), $\widetilde{L}_p$ is concave with respect to $\boldsymbol{\lambda}_l$. Based on (56), we can obtain that,

$$\begin{aligned}
&\widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k+1}\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\
&- \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}) \\
&\leq \sum_{l=1}^{L} \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}), \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&\leq \sum_{l=1}^{L} \Big(\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle + \frac{1}{\eta_{\boldsymbol{\lambda}}} \langle \boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \rangle \Big) \\
&- \sum_{l=1}^{L} \Big(\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle .
\end{aligned} \tag{57}$$

And we have that,

$$\begin{aligned}
&\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&= \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&+ \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k - (\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}) \right\rangle \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k - (\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}) \right\rangle \\
&+ \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1} \right\rangle \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}, \boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1} \right\rangle .
\end{aligned} \tag{58}$$

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

$$\begin{aligned}
&\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
&\leq \left\langle \nabla_{\boldsymbol{\lambda}_l} 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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle + \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2}(||\lambda_l^{k+1}||^2 - ||\boldsymbol{\lambda}_l^k||^2) \\
&- \left\langle \nabla_{\boldsymbol{\lambda}_l} 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\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle - \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2}||\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k||^2 \\
&\leq \frac{a_1}{2}\|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\| + \frac{1}{a_1}\|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\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_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2}(||\lambda_l^{k+1}||^2 - ||\boldsymbol{\lambda}_l^k||^2) \\
&- \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2}||\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k||^2,
\end{aligned} \tag{59}$$

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

$$\begin{aligned}&\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\},\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k-(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\right\rangle\\&-\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^{k-1}\},\{\boldsymbol{\alpha}_l^{k-1}\},\{\boldsymbol{\beta}_l^{k-1}\},\{\gamma_s^{k-1}\},\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k-(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\right\rangle\\&\leq \frac{1}{2\eta_{\boldsymbol{\lambda}}}\|\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k-(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\|^2+\frac{\eta_{\boldsymbol{\lambda}}}{2}\|c_{\boldsymbol{\lambda}}^{k-1}(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\|^2.\end{aligned} \tag{60}$$

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

$$\begin{aligned}&\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\},\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1}\right\rangle\\&-\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^{k-1}\},\{\boldsymbol{\alpha}_l^{k-1}\},\{\boldsymbol{\beta}_l^{k-1}\},\{\gamma_s^{k-1}\},\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1}\right\rangle\\&\leq -\frac{1}{2c_{\boldsymbol{\lambda}}^{k-1}}\|c_{\boldsymbol{\lambda}}^{k-1}(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\|^2-\frac{c_{\boldsymbol{\lambda}}^{k-1}}{2}\|\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1}\|^2.\end{aligned} \tag{61}$$

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

$$\begin{aligned}&\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\},\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k\right\rangle\\&-\left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^k\},\{\mathbf{r}_l^k\},\{\mathbf{p}_l^k\},\{q_s^k\},\mathbf{v}^k,\{\boldsymbol{\lambda}_l^{k-1}\},\{\boldsymbol{\alpha}_l^{k-1}\},\{\boldsymbol{\beta}_l^{k-1}\},\{\gamma_s^{k-1}\},\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k\right\rangle\\&\leq \frac{a_1}{2}\|\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k\|+\frac{1}{a_1}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2+\frac{1}{a_1}\|\mathbf{r}_l^{k+1}\circ\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\circ\mathbf{r}_l^k\|^2+\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}(||\lambda_l^{k+1}||^2-||\boldsymbol{\lambda}_l^k||^2)\\&-\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}||\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k||^2+\frac{1}{2\eta_{\boldsymbol{\lambda}}}\|\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k-(\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1})\|^2-\frac{c_{\boldsymbol{\lambda}}^{k-1}}{2}\|\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1}\|^2.\end{aligned} \tag{62}$$

According to Assumption 1, we have,

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

In addition, the following equality can be obtained,

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

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

$$\begin{aligned}&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^{k+1}\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\\&-L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^k\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\\&\leq \frac{1}{a_1}\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2+\frac{4w_r^2}{a_1}\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2+(\frac{a_1}{2}-\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}+\frac{1}{2\eta_{\boldsymbol{\lambda}}})\sum_{l=1}^{L}\|\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k\|^2\\&+\frac{c_{\boldsymbol{\lambda}}^{k-1}}{2}\sum_{l=1}^{L}(\|\boldsymbol{\lambda}_l^{k+1}\|^2-\|\boldsymbol{\lambda}_l^k\|^2)+\frac{1}{2\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}\|\boldsymbol{\lambda}_l^k-\boldsymbol{\lambda}_l^{k-1}\|^2.\end{aligned} \tag{65}$$

Likewise, similar results can be obtained for other variables :

$$\begin{aligned}&L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^{k+1}\},\{\boldsymbol{\alpha}_l^{k+1}\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\\&-L_p(\{\boldsymbol{\mu}_l^{k+1}\},\{\mathbf{r}_l^{k+1}\},\{\mathbf{p}_l^{k+1}\},\{q_s^{k+1}\},\mathbf{v}^{k+1},\{\boldsymbol{\lambda}_l^{k+1}\},\{\boldsymbol{\alpha}_l^k\},\{\boldsymbol{\beta}_l^k\},\{\gamma_s^k\})\\&\leq \frac{1}{a_2}\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2+\frac{4w_{\mathbf{p}}^2}{a_2}\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2+(\frac{a_2}{2}-\frac{c_{\boldsymbol{\alpha}}^{k-1}-c_{\boldsymbol{\alpha}}^k}{2}+\frac{1}{2\eta_{\boldsymbol{\alpha}}})\sum_{l=1}^{L}\|\boldsymbol{\alpha}_l^{k+1}-\boldsymbol{\alpha}_l^k\|^2\\&+\frac{c_{\boldsymbol{\alpha}}^{k-1}}{2}\sum_{l=1}^{L}(\|\boldsymbol{\alpha}_l^{k+1}\|^2-\|\boldsymbol{\alpha}_l^k\|^2)+\frac{1}{2\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}\|\boldsymbol{\alpha}_l^k-\boldsymbol{\alpha}_l^{k-1}\|^2,\end{aligned} \tag{66}$$

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

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

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 (\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\frac{L_w}{6}+1-\frac{1}{\eta_{\boldsymbol{\mu}}^k})\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^{k}\|^2+(\frac{4w_{\mathbf{r}}^2}{a_1}+\frac{w_{\mathbf{r}}L_wM^{\frac{1}{2}}}{3}+w_{\boldsymbol{\lambda}}-\frac{1}{\eta_{\mathbf{r}}^k})\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^{k}\|^2\\
&+(\frac{4w_{\mathbf{p}}^2}{a_2}+\frac{w_{\mathbf{p}}L_wM^{\frac{1}{2}}}{3}+w_{\boldsymbol{\alpha}}-\frac{1}{\eta_{\mathbf{p}}^k})\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^{k}\|^2+(\frac{4w_q^2}{a_4}+\frac{w_qL_w}{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_{\mathbf{v}}^k}+\frac{L_w}{6})\|\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\\
&+(\frac{a_1}{2}-\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^{k}}{2}+\frac{1}{2\eta_{\boldsymbol{\lambda}}})\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\\
&+(\frac{a_2}{2}-\frac{c_{\boldsymbol{\alpha}}^{k-1}-c_{\boldsymbol{\alpha}}^{k}}{2}+\frac{1}{2\eta_{\boldsymbol{\alpha}}})\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\\
&+(\frac{a_3}{2}-\frac{c_{\boldsymbol{\beta}}^{k-1}-c_{\boldsymbol{\beta}}^{k}}{2}+\frac{1}{2\eta_{\boldsymbol{\beta}}})\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\\
&+(\frac{a_4}{2}-\frac{c_{\gamma}^{k-1}-c_{\gamma}^{k}}{2}+\frac{1}{2\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{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 :*

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

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

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

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

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

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

$$\begin{aligned}
& L_p(\{\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 (\eta_{\boldsymbol{\lambda}} + \eta_{\boldsymbol{\alpha}} + \eta_{\boldsymbol{\beta}} + \frac{L_w}{6} + 1 - \frac{1}{\eta_{\boldsymbol{\mu}}^k}) \sum_{l=1}^{L} \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + (4w_{\mathbf{r}}^2 \eta_{\boldsymbol{\lambda}} + \frac{w_{\mathbf{r}} L_w M^{\frac{1}{2}}}{3} + w_{\boldsymbol{\lambda}} - \frac{1}{\eta_{\mathbf{r}}^k}) \sum_{l=1}^{L} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + (4w_{\mathbf{p}}^2 \eta_{\boldsymbol{\alpha}} + \frac{w_{\mathbf{p}} L_w M^{\frac{1}{2}}}{3} + w_{\boldsymbol{\alpha}} - \frac{1}{\eta_{\mathbf{p}}^k}) \sum_{l=1}^{L} \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 + (4w_q^2 \eta_\gamma + \frac{w_q L_w}{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_{\mathbf{v}}^k} + \frac{L_w}{6}) ||\mathbf{v}^{k+1} - \mathbf{v}^k||^2 + \eta_{\boldsymbol{\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_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2} + \frac{1}{\eta_{\boldsymbol{\lambda}}}) \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 \\
& + (-\frac{c_{\boldsymbol{\alpha}}^{k-1} - c_{\boldsymbol{\alpha}}^k}{2} + \frac{1}{\eta_{\boldsymbol{\alpha}}}) \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 \\
& + (-\frac{c_{\boldsymbol{\beta}}^{k-1} - c_{\boldsymbol{\beta}}^k}{2} + \frac{1}{\eta_{\boldsymbol{\beta}}}) \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\|^2) + \frac{1}{2\eta_{\boldsymbol{\beta}}} \sum_{l=1}^{L} \|\boldsymbol{\beta}_l^k - \boldsymbol{\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,

$$\begin{aligned}
& \frac{1}{\eta_{\boldsymbol{\lambda}}} \langle \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k - (\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}), \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \rangle \\
& \leq \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k+1}\}, \{\mathbf{r}_l^{k+1}\}, \{\mathbf{p}_l^{k+1}\}, \{q_s^{k+1}\}, \mathbf{v}^{k+1}, \{\boldsymbol{\lambda}_l^{k}\}, \{\boldsymbol{\alpha}_l^{k}\}, \{\boldsymbol{\beta}_l^{k}\}, \{\gamma_s^{k}\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
& - \left\langle \nabla_{\boldsymbol{\lambda}_l} \widetilde{L}_p(\{\boldsymbol{\mu}_l^{k}\}, \{\mathbf{r}_l^{k}\}, \{\mathbf{p}_l^{k}\}, \{q_s^{k}\}, \mathbf{v}^{k}, \{\boldsymbol{\lambda}_l^{k-1}\}, \{\boldsymbol{\alpha}_l^{k-1}\}, \{\boldsymbol{\beta}_l^{k-1}\}, \{\gamma_s^{k-1}\}, \boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k \right\rangle \\
& \leq \frac{1}{b_1^k} \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 + \frac{4w_{\mathbf{r}}^2}{b_1^k} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 + \frac{b_1^k}{2} \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k\|^2 + \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2} (||\boldsymbol{\lambda}_l^{k+1}||^2 - ||\boldsymbol{\lambda}_l^k||^2) \\
& - \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2} ||\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k||^2 + \frac{1}{2\eta_{\boldsymbol{\lambda}}} \|\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k - (\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1})\|^2 - \frac{c_{\boldsymbol{\lambda}}^{k-1}}{2} \|\boldsymbol{\lambda}_l^k - \boldsymbol{\lambda}_l^{k-1}\|^2.
\end{aligned} \tag{75}$$

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

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

Combining (75) with (76), we have

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

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

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

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

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

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

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

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

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

Based on the setting of $c_{\boldsymbol{\lambda}}^k$, $c_{\boldsymbol{\alpha}}^k$, $c_{\boldsymbol{\beta}}^k$ and c_{γ}^k, we can obtain that, $\frac{\eta_{\boldsymbol{\lambda}}}{10} \geq \frac{1}{c_{\boldsymbol{\lambda}}^{k+1}}-\frac{1}{c_{\boldsymbol{\lambda}}^{k}}$, $\frac{\eta_{\boldsymbol{\alpha}}}{10} \geq \frac{1}{c_{\boldsymbol{\alpha}}^{k+1}}-\frac{1}{c_{\boldsymbol{\alpha}}^{k}}$, $\frac{\eta_{\boldsymbol{\beta}}}{10} \geq \frac{1}{c_{\boldsymbol{\beta}}^{k+1}}-\frac{1}{c_{\boldsymbol{\beta}}^{k}}$, $\frac{\eta_{\gamma}}{10} \geq \frac{1}{c_{\gamma}^{k+1}}-\frac{1}{c_{\gamma}^{k}}$, $\forall k \geq K_1$. 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 & (a_5-\frac{1}{\eta_{\boldsymbol{\mu}}^k}+\frac{16}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2}+\frac{16}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2}+\frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2})\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2 \\
& +(a_5-\frac{1}{\eta_{\mathbf{r}}^k}+\frac{64w_{\mathbf{r}}^2}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2})\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2 \\
& +(a_5-\frac{1}{\eta_{\mathbf{p}}^k}+\frac{64w_{\mathbf{p}}^2}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2})\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2 \\
& +(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 \\
& +(a_5-\frac{1}{\eta_{\mathbf{v}}^k}+\frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2}+\frac{16}{\eta_{\gamma}(c_{\gamma}^k)^2}\sum_{s=1}^{|\mathcal{P}^k|}\|b_s\|^2)\|\mathbf{v}^{k+1}-\mathbf{v}^k\|^2 \\
& -\frac{1}{10\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k||^2+\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1}||^2+\frac{4}{\eta_{\boldsymbol{\lambda}}}(\frac{c_{\boldsymbol{\lambda}}^{k-2}}{c_{\boldsymbol{\lambda}}^{k-1}}-\frac{c_{\boldsymbol{\lambda}}^{k-1}}{c_{\boldsymbol{\lambda}}^{k}})\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^k||^2 \\
& -\frac{1}{10\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}||\alpha_l^{k+1}-\alpha_l^k||^2+\frac{c_{\boldsymbol{\alpha}}^{k-1}-c_{\boldsymbol{\alpha}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{k+1}||^2+\frac{4}{\eta_{\boldsymbol{\alpha}}}(\frac{c_{\boldsymbol{\alpha}}^{k-2}}{c_{\boldsymbol{\alpha}}^{k-1}}-\frac{c_{\boldsymbol{\alpha}}^{k-1}}{c_{\boldsymbol{\alpha}}^{k}})\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^k||^2 \\
& -\frac{1}{10\eta_{\boldsymbol{\beta}}}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{k+1}-\boldsymbol{\beta}_l^k||^2+\frac{c_{\boldsymbol{\beta}}^{k-1}-c_{\boldsymbol{\beta}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{k+1}||^2+\frac{4}{\eta_{\boldsymbol{\beta}}}(\frac{c_{\boldsymbol{\beta}}^{k-2}}{c_{\boldsymbol{\beta}}^{k-1}}-\frac{c_{\boldsymbol{\beta}}^{k-1}}{c_{\boldsymbol{\beta}}^{k}})\sum_{l=1}^{L}||\boldsymbol{\beta}_l^k||^2 \\
& -\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}}(\frac{c_{\gamma}^{k-2}}{c_{\gamma}^{k-1}}-\frac{c_{\gamma}^{k-1}}{c_{\gamma}^{k}})\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

$$a_6^k = \min\{\frac{16}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2} + \frac{16}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2} + \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2}, \frac{64w_{\mathbf{r}}^2}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2}, \frac{64w_{\mathbf{p}}^2}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2}, \frac{64w_q^2}{\eta_\gamma(c_\gamma^k)^2}, \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2} + \frac{16}{\eta_\gamma(c_\gamma^k)^2}\sum_{s=1}^{L}\|b_s\|^2\}\frac{\xi-2}{2} - a_5. \tag{85}$$

where constant $\xi > 2$ and satisfies

$$\min\{\frac{16}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^0)^2} + \frac{16}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^0)^2} + \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^0)^2}, \frac{64w_{\mathbf{r}}^2}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^0)^2}, \frac{64w_{\mathbf{p}}^2}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^0)^2}, \frac{64w_q^2}{\eta_\gamma(c_\gamma^0)^2}, \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^0)^2} + \frac{16}{\eta_\gamma(c_\gamma^0)^2}\sum_{s=1}^{L}\|b_s\|^2\}\frac{\xi-2}{2} > a_5. \tag{86}$$

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

$$\begin{aligned}
a_5 - \frac{1}{\eta_{\boldsymbol{\mu}}^k} + \frac{16}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2} + \frac{16}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2} + \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2} &\le -a_6^k,\\
a_5 - \frac{1}{\eta_{\mathbf{r}}^k} + \frac{64w_{\mathbf{r}}^2}{\eta_{\boldsymbol{\lambda}}(c_{\boldsymbol{\lambda}}^k)^2} &\le -a_6^k,\\
a_5 - \frac{1}{\eta_{\mathbf{p}}^k} + \frac{64w_{\mathbf{p}}^2}{\eta_{\boldsymbol{\alpha}}(c_{\boldsymbol{\alpha}}^k)^2} &\le -a_6^k,\\
(a_5 - \frac{1}{\eta_q^k} + \frac{64w_q^2}{\eta_\gamma(c_\gamma^k)^2}) &\le -a_6^k,\\
a_5 - \frac{1}{\eta_{\mathbf{v}}^k} + \frac{16}{\eta_{\boldsymbol{\beta}}(c_{\boldsymbol{\beta}}^k)^2} + \frac{16}{\eta_\gamma(c_\gamma^k)^2}\sum_{s=1}^{L}\|b_s\|^2 &\le -a_6^k.
\end{aligned} \tag{87}$$

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

$$\begin{aligned}
& a_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)\\
& \le F^k - F^{k+1} - \frac{1}{10\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1} - \boldsymbol{\lambda}_l^k||^2 + \frac{c_{\boldsymbol{\lambda}}^{k-1} - c_{\boldsymbol{\lambda}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1}||^2 + \frac{4}{\eta_{\boldsymbol{\lambda}}}(\frac{c_{\boldsymbol{\lambda}}^{k-2}}{c_{\boldsymbol{\lambda}}^{k-1}} - \frac{c_{\boldsymbol{\lambda}}^{k-1}}{c_{\boldsymbol{\lambda}}^k})\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^k||^2\\
& - \frac{1}{10\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}||\alpha_l^{k+1} - \alpha_l^k||^2 + \frac{c_{\boldsymbol{\alpha}}^{k-1} - c_{\boldsymbol{\alpha}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{k+1}||^2 + \frac{4}{\eta_{\boldsymbol{\alpha}}}(\frac{c_{\boldsymbol{\alpha}}^{k-2}}{c_{\boldsymbol{\alpha}}^{k-1}} - \frac{c_{\boldsymbol{\alpha}}^{k-1}}{c_{\boldsymbol{\alpha}}^k})\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^k||^2\\
& - \frac{1}{10\eta_{\boldsymbol{\beta}}}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{k+1} - \boldsymbol{\beta}_l^k||^2 + \frac{c_{\boldsymbol{\beta}}^{k-1} - c_{\boldsymbol{\beta}}^k}{2}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{k+1}||^2 + \frac{4}{\eta_{\boldsymbol{\beta}}}(\frac{c_{\boldsymbol{\beta}}^{k-2}}{c_{\boldsymbol{\beta}}^{k-1}} - \frac{c_{\boldsymbol{\beta}}^{k-1}}{c_{\boldsymbol{\beta}}^k})\sum_{l=1}^{L}||\boldsymbol{\beta}_l^k||^2\\
& - \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}(\frac{c_\gamma^{k-2}}{c_\gamma^{k-1}} - \frac{c_\gamma^{k-1}}{c_\gamma^k})\sum_{s=1}^{|\mathcal{P}^k|}||\gamma_s^k||^2.
\end{aligned} \tag{88}$$

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

$$\begin{aligned}
(\nabla\widetilde{G}^k)_{\boldsymbol{\mu}_l} =& \nabla_{\boldsymbol{\mu}_l}\widetilde{L}_p(\{\boldsymbol{\mu}_l^{\hat{k}_l}\}, \{\mathbf{r}_l^{\hat{k}_l}\}, \{\mathbf{p}_l^{\hat{k}_l}\}, \{q_s^{\hat{k}_l}\}, \mathbf{v}^{\hat{k}_l}, \{\boldsymbol{\lambda}_l^{\hat{k}}\}, \{\boldsymbol{\alpha}_l^{\hat{k}}\}, \{\boldsymbol{\beta}_l^{\hat{k}}\}, \{\gamma_s^{\hat{k}}\})\\
&+ \nabla_{\boldsymbol{\mu}_l}\widetilde{L}_p(\{\boldsymbol{\mu}_l^k\}, \{\mathbf{r}_l^k\}, \{\mathbf{p}_l^k\}, \{q_s^k\}, \mathbf{v}^k, \{\boldsymbol{\lambda}_l^k\}, \{\boldsymbol{\alpha}_l^k\}, \{\boldsymbol{\beta}_l^k\}, \{\gamma_s^k\})\\
&- \nabla_{\boldsymbol{\mu}_l}\widetilde{L}_p(\{\boldsymbol{\mu}_l^{\hat{k}_l}\}, \{\mathbf{r}_l^{\hat{k}_l}\}, \{\mathbf{p}_l^{\hat{k}_l}\}, \{q_s^{\hat{k}_l}\}, \mathbf{v}^{\hat{k}_l}, \{\boldsymbol{\lambda}_l^{\hat{k}}\}, \{\boldsymbol{\alpha}_l^{\hat{k}}\}, \{\boldsymbol{\beta}_l^{\hat{k}}\}, \{\gamma_s^{\hat{k}}\}).
\end{aligned} \tag{89}$$

Combining it with (41), we have that,

$$\|(\nabla\widetilde{G}^k)_{\boldsymbol{\mu}_l}\|^2 \le \frac{1}{\eta_{\boldsymbol{\mu}}^2}\|\boldsymbol{\mu}_l^{\bar{k}_l} - \boldsymbol{\mu}_l^k\|^2. \tag{90}$$

Similar results can be derived for other variables,

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

According to the definition 3, we have,

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

Combining trigonometric inequality, (28) with Assumption 1, we can obtain

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

Similar results can be derived for other variables as well,

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

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\widetilde{G}^k||^2 =& \sum_{l=1}^{L}(||(\nabla\widetilde{G}^k)_{\boldsymbol{\mu}_l}||^2 + ||(\nabla\widetilde{G}^k)_{\mathbf{r}_l}||^2 + ||\nabla G^k)_{\mathbf{p}_l}||^2 + ||\nabla\widetilde{G}^k)_{\boldsymbol{\lambda}_l}||^2 + ||(\nabla G^k)_{\boldsymbol{\alpha}_l}||^2 + ||(\nabla\widetilde{G}^k)_{\boldsymbol{\beta}_l}||^2) \\
&\sum_{s=1}^{|\mathcal{P}^k|}||(\nabla\widetilde{G}^k)_{q_s}||^2 + ||(\nabla\widetilde{G}^k)_{\mathbf{v}}||^2 + \sum_{s=1}^{|\mathcal{P}^k|}||(\nabla\widetilde{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 \\
&+(\frac{1}{\eta_q^2}+24w_{\mathbf{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 \\
&+\frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\lambda}}^{\bar{k}_l-1})^2)||\boldsymbol{\lambda}_l^k||^2 \\
&+\frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\alpha}}^{\bar{k}_l-1})^2)||\boldsymbol{\alpha}_l^k||^2 \\
&+\frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\beta}}^{\bar{k}_l-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}^{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,

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

where k_τ is a positive constant and it satisfies that,

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

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

$$\begin{aligned}
||\nabla \widetilde{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) \\
& + (\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_{\boldsymbol{\mu}}^2} + 18) k_\tau \tau \sum_{l=1}^{L} \|\boldsymbol{\mu}_l^{k+1} - \boldsymbol{\mu}_l^k\|^2 \\
& + (\frac{1}{\eta_{\mathbf{r}}^2} + 24 w_{\mathbf{r}}^2) \sum_{l=1}^{L} \|\mathbf{r}_l^{\bar{k}_l} - \mathbf{r}_l^k\|^2 - (\frac{1}{\eta_{\mathbf{r}}^2} + 24 w_{\mathbf{r}}^2) k_\tau \tau \sum_{l=1}^{L} \|\mathbf{r}_l^{k+1} - \mathbf{r}_l^k\|^2 \\
& + (\frac{1}{\eta_{\mathbf{p}}^2} + 24 w_{\mathbf{p}}^2) \sum_{l=1}^{L} \|\mathbf{p}_l^{\bar{k}_l} - \mathbf{p}_l^k\|^2 - (\frac{1}{\eta_{\mathbf{p}}^2} + 24 w_{\mathbf{p}}^2) k_\tau \tau \sum_{l=1}^{L} \|\mathbf{p}_l^{k+1} - \mathbf{p}_l^k\|^2 \\
& + \frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l - 1})^2 - (c_{\boldsymbol{\lambda}}^{\bar{k}_l - 1})^2) ||\boldsymbol{\lambda}_l^k||^2 \\
& + \frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l - 1})^2 - (c_{\boldsymbol{\alpha}}^{\bar{k}_l - 1})^2) ||\boldsymbol{\alpha}_l^k||^2 \\
& + \frac{3}{\eta_{\boldsymbol{\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}}^{\hat{k}_l - 1})^2 - (c_{\boldsymbol{\beta}}^{\bar{k}_l - 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}^{k-1})^2 - (c_{\gamma}^k)^2) ||\gamma_s^k||^2.
\end{aligned} \tag{99}$$

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

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

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 \widetilde{G}^k||^2 \leq & a_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(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{\bar{k}_l}-\boldsymbol{\mu}_l^k\|^2-d_6^k(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)k_\tau\tau\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2 \\
&+d_6^k(\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-d_6^k(\frac{1}{\eta_{\mathbf{r}}^2}+24w_{\mathbf{r}}^2)k_\tau\tau\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2 \\
&+d_6^k(\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-d_6^k(\frac{1}{\eta_{\mathbf{p}}^2}+24w_{\mathbf{p}}^2)k_\tau\tau\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{\bar{k}_l}-\boldsymbol{\lambda}_l^k||^2+3d_6^k\sum_{l=1}^{L}((c_{\boldsymbol{\lambda}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\lambda}}^{\bar{k}_l-1})^2)||\boldsymbol{\lambda}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{\bar{k}_l}-\boldsymbol{\alpha}_l^k||^2+3d_6^k\sum_{l=1}^{L}((c_{\boldsymbol{\alpha}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\alpha}}^{\bar{k}_l-1})^2)||\boldsymbol{\alpha}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\beta}}}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{\bar{k}_l}-\boldsymbol{\beta}_l^k||^2+3d_6^k\sum_{l=1}^{L}((c_{\boldsymbol{\beta}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\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 \widetilde{G}^k||^2 \leq & F^k-F^{k+1}+\overline{d_6}(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{\bar{k}_l}-\boldsymbol{\mu}_l^k\|^2-\underline{d_6}(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)k_\tau\tau\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2 \\
&+\overline{d_6}(\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-\underline{d_6}(\frac{1}{\eta_{\mathbf{r}}^2}+24w_{\mathbf{r}}^2)k_\tau\tau\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2 \\
&+\overline{d_6}(\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-\underline{d_6}(\frac{1}{\eta_{\mathbf{p}}^2}+24w_{\mathbf{p}}^2)k_\tau\tau\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{\bar{k}_l}-\boldsymbol{\lambda}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\lambda}}}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{\bar{k}_l}-\boldsymbol{\alpha}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\alpha}}}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{\bar{k}_l}-\boldsymbol{\alpha}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\beta}}}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{\bar{k}_l}-\boldsymbol{\beta}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\beta}}}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{\bar{k}_l}-\boldsymbol{\beta}_l^k||^2 \\
&+3\overline{d_6}\sum_{l=1}^{L}((c_{\boldsymbol{\lambda}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\lambda}}^{\bar{k}_l-1})^2)||\boldsymbol{\lambda}_l^k||^2+3\overline{d_6}\sum_{l=1}^{L}((c_{\boldsymbol{\alpha}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\alpha}}^{\bar{k}_l-1})^2)||\boldsymbol{\alpha}_l^k||^2 \\
&+3\overline{d_6}\sum_{l=1}^{L}((c_{\boldsymbol{\beta}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\beta}}^{\bar{k}_l-1})^2)||\boldsymbol{\beta}_l^k||^2+3\overline{d_6}\sum_{s=1}^{|\mathcal{P}^k|}((c_\gamma^{k-1})^2-(c_\gamma^k)^2)||\gamma_s^k||^2 \\
&+\frac{c_{\boldsymbol{\lambda}}^{k-1}-c_{\boldsymbol{\lambda}}^k}{2}Mw_{\boldsymbol{\lambda}}^2+\frac{c_{\boldsymbol{\alpha}}^{k-1}-c_{\boldsymbol{\alpha}}^k}{2}Mw_{\boldsymbol{\alpha}}^2+\frac{c_{\boldsymbol{\beta}}^{k-1}-c_{\boldsymbol{\beta}}^k}{2}Mw_{\boldsymbol{\beta}}^2+\frac{c_\gamma^{k-1}-c_\gamma^k}{2}Pw_\gamma^2 \\
&+\frac{4}{\eta_{\boldsymbol{\lambda}}}(\frac{c_{\boldsymbol{\lambda}}^{k-2}}{c_{\boldsymbol{\lambda}}^{k-1}}-\frac{c_{\boldsymbol{\lambda}}^{k-1}}{c_{\boldsymbol{\lambda}}^k})\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^k||^2+\frac{4}{\eta_{\boldsymbol{\alpha}}}(\frac{c_{\boldsymbol{\alpha}}^{k-2}}{c_{\boldsymbol{\alpha}}^{k-1}}-\frac{c_{\boldsymbol{\alpha}}^{k-1}}{c_{\boldsymbol{\alpha}}^k})\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^k||^2 \\
&+\frac{4}{\eta_{\boldsymbol{\beta}}}(\frac{c_{\boldsymbol{\beta}}^{k-2}}{c_{\boldsymbol{\beta}}^{k-1}}-\frac{c_{\boldsymbol{\beta}}^{k-1}}{c_{\boldsymbol{\beta}}^k})\sum_{l=1}^{L}||\boldsymbol{\beta}_l^k||^2+\frac{4}{\eta_\gamma}(\frac{c_\gamma^{k-2}}{c_\gamma^{k-1}}-\frac{c_\gamma^{k-1}}{c_\gamma^k})\sum_{s=1}^{|\mathcal{P}^k|}||\gamma_s^k||^2.
\end{aligned}
\tag{102}
$$

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

$$
\begin{aligned}
\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)} d_6^k||\nabla \widetilde{G}^k||^2 \leq & F^{K_1+2}-F^{K_1+\widetilde{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 \\
&+\overline{d_6}(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{\bar{k}_l}-\boldsymbol{\mu}_l^k\|^2-\underline{d_6}(\frac{1}{\eta_{\boldsymbol{\mu}}^2}+18)k_\tau\tau\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k\|^2 \\
&+\overline{d_6}(\frac{1}{\eta_{\mathbf{r}}^2}+24w_{\mathbf{r}}^2)\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\mathbf{r}_l^{\bar{k}_l}-\mathbf{r}_l^k\|^2-\underline{d_6}(\frac{1}{\eta_{\mathbf{r}}^2}+24w_{\mathbf{r}}^2)k_\tau\tau\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\mathbf{r}_l^{k+1}-\mathbf{r}_l^k\|^2 \\
&+\overline{d_6}(\frac{1}{\eta_{\mathbf{p}}^2}+24w_{\mathbf{p}}^2)\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\mathbf{p}_l^{\bar{k}_l}-\mathbf{p}_l^k\|^2-\underline{d_6}(\frac{1}{\eta_{\mathbf{p}}^2}+24w_{\mathbf{p}}^2)k_\tau\tau\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\mathbf{p}_l^{k+1}-\mathbf{p}_l^k\|^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\lambda}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{\bar{k}_l}-\boldsymbol{\lambda}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\lambda}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\lambda}_l^{k+1}-\boldsymbol{\lambda}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\alpha}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{\bar{k}_l}-\boldsymbol{\alpha}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\alpha}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\alpha}_l^{k+1}-\boldsymbol{\alpha}_l^k||^2 \\
&+\frac{1}{10\tau\eta_{\boldsymbol{\beta}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{\bar{k}_l}-\boldsymbol{\beta}_l^k||^2-\frac{1}{10\eta_{\boldsymbol{\beta}}}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\beta}_l^{k+1}-\boldsymbol{\beta}_l^k||^2 \\
&+3\overline{d_6}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}((c_{\boldsymbol{\lambda}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\lambda}}^{\bar{k}_l-1})^2)||\boldsymbol{\lambda}_l^k||^2+3\overline{d_6}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}((c_{\boldsymbol{\alpha}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\alpha}}^{\bar{k}_l-1})^2)||\boldsymbol{\alpha}_l^k||^2 \\
&+3\overline{d_6}\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}((c_{\boldsymbol{\beta}}^{\hat{k}_l-1})^2-(c_{\boldsymbol{\beta}}^{\bar{k}_l-1})^2)||\boldsymbol{\beta}_l^k||^2+3\overline{d_6}(c_{\gamma}^1)^2Pw_{\gamma}^2.
\end{aligned} \tag{103}
$$

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

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

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 $\boldsymbol{\mu}_l^k=\boldsymbol{\mu}_l^{\hat{v}_l(j)-1}$. And for $k\notin\mathcal{V}_l(K)$, we have $\|\boldsymbol{\mu}_l^k-\boldsymbol{\mu}_l^{k-1}\|^2=0$. Since $\hat{v}_l(j)-\hat{v}_l(j-1)\leq\tau$, we can obtain

$$
\sum_{k=K_1+2}^{K_1+\widetilde{K}(\epsilon)}\sum_{l=1}^{L}\|\boldsymbol{\mu}_l^{\bar{k}_l}-\boldsymbol{\mu}_l^k\|^2\leq\tau\sum_{k=K_1+2}^{K_1+\widetilde{T}(\epsilon)}\sum_{l=1}^{L}||\boldsymbol{\mu}_l^{k+1}-\boldsymbol{\mu}_l^k||^2+4\tau(\tau-1)Mw_{\boldsymbol{\mu}}^2. \tag{105}
$$

Similarly, we can obtain

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

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

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

where d_7 and k_d are constants. Set constant d_8 as

$$d_8=\left(\max\left\{d_1,d_2,d_3,d_4,d_5,\frac{30\tau}{\eta_{\boldsymbol{\lambda}}\underline{a_6}},\frac{30\tau}{\eta_{\boldsymbol{\alpha}}\underline{a_6}},\frac{30\tau}{\eta_{\boldsymbol{\beta}}\underline{a_6}},\frac{30}{\eta_{\gamma}\underline{a_6}}\right\}\right)\geq\frac{1}{d_6^ka_6^k}.\tag{108}$$

Thus, we can obtain that

$$\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)}\frac{1}{d_8a_6^k}\|\nabla\widetilde{G}^{K_1+\tilde{K}(\epsilon)}\|^2\leq\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)}\frac{1}{d_8a_6^k}\|\nabla\widetilde{G}^k\|^2\leq\sum_{k=K_1+2}^{K_1+\tilde{K}(\epsilon)}d_6^k\|\nabla\widetilde{G}^k\|^2\leq(d_7+k_d\tau)(\tau-1).\tag{109}$$

Thus, we can obtain

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

According to the setting of $c_{\boldsymbol{\lambda}}^k$, $c_{\boldsymbol{\alpha}}^k$, $c_{\boldsymbol{\beta}}^k$ and c_{γ}^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 $k=K_1+\widetilde{K}$, we have

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

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

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

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

$$K_1+\widetilde{K}(\epsilon)\geq(\frac{d_9(d_7+k_d\tau)(\tau-1)d_8}{\epsilon}+(K_1+2)^{\frac{1}{2}})^2. \tag{114}$$

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

$$||\nabla G^k||-||\nabla\widetilde{G}^k||\leq||\nabla G^k-\nabla\widetilde{G}^k||\leq\sqrt{\sum_{l=1}^{L}||c_{\boldsymbol{\lambda}}^{k-1}\boldsymbol{\lambda}_l^k||^2+\sum_{l=1}^{L}||c_{\boldsymbol{\alpha}}^{k-1}\boldsymbol{\alpha}_l^k||^2++\sum_{l=1}^{L}||c_{\boldsymbol{\beta}}^{k-1}\boldsymbol{\beta}_l^k||^2++\sum_{s=1}^{|\mathcal{P}^k|}||c_{\gamma}^{k-1}\gamma_s^k||^2}. \tag{115}$$

If $k\geq16(\frac{Mw_{\boldsymbol{\lambda}}^2}{\eta_{\lambda}^2}+\frac{Mw_{\boldsymbol{\alpha}}^2}{\eta_{\boldsymbol{\alpha}}^2}+\frac{Mw_{\boldsymbol{\beta}}^2}{\eta_{\boldsymbol{\beta}}^2}+\frac{Pw_{\gamma}^2}{\eta_{\gamma}^2})^2\frac{1}{\epsilon^2}$, we have

$$\sqrt{\sum_{l=1}^{L}||c_{\boldsymbol{\lambda}}^{k-1}\boldsymbol{\lambda}_l^k||^2+\sum_{l=1}^{L}||c_{\boldsymbol{\alpha}}^{k-1}\boldsymbol{\alpha}_l^k||^2++\sum_{l=1}^{L}||c_{\boldsymbol{\beta}}^{k-1}\boldsymbol{\beta}_l^k||^2++\sum_{s=1}^{|\mathcal{P}^k|}|||c_{\gamma}^{k-1}\gamma_s^k||^2}\leq\frac{\sqrt{\epsilon}}{2}. \tag{116}$$

Combining it with (114), we can conclude that

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

which concludes our proof. $\square$