Supplementary Material for "Constrained Multi-Objective Optimization"

A Some Useful Lemma

To simplify the notions, we use $\boldsymbol{\omega}^{(t)} = (x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})$ in the following proof.

Lemma A.1 *Yao et al. (2024)* ***Lemma A.1***. *Under Assumption 4.3 and 4.4, for the Moreau envelope-based Lagrange Multiplier function $L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}$ with $\gamma_1 \in (0, 1/\rho_f)$ and $\gamma_2 > 0$. That is,*

(1) The function $L_s(x, \boldsymbol{\lambda}, \boldsymbol{z})$ is continuously differentiable;

(2) The gradient of $L_s(x, \boldsymbol{\lambda}, \boldsymbol{z})$ has closed-form given by

$$\nabla_x L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}) = \arg\min_u \{H(u, \boldsymbol{\theta}^*) + \frac{1}{2}\|u - x\|^2\} + \sum_{i=1}^n \mu_i^* \nabla g_i(x),$$
$$\nabla_{\boldsymbol{\lambda}} L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}) = \frac{\boldsymbol{\lambda} - \boldsymbol{\theta}^*}{\gamma_1},$$
$$\nabla_{\boldsymbol{\lambda}} L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}) = \frac{\boldsymbol{\mu}^* - \boldsymbol{z}}{\gamma_2},$$

where $\boldsymbol{\theta}^* := \boldsymbol{\theta}^*(x, \boldsymbol{\lambda}, \boldsymbol{z})$ and $\boldsymbol{\mu}^* := \boldsymbol{\mu}^*(x, \boldsymbol{\lambda}, \boldsymbol{z})$ is the unique saddle point of the following min-max problem:

$$\min_\theta \max_{\boldsymbol{\mu}} \left\{ H(x, \boldsymbol{\theta}) + \sum_{i=1}^N \mu_i g_i(x) + \frac{1}{2\gamma_1} \sum_{i=1}^N \|\theta_i - \lambda_i\|^2 - \frac{1}{2\gamma_2} \sum_{i=1}^N \|z_i - \mu_i\|^2 \right\}.$$

(3) Furthermore, for any $\rho_v \geq \rho_f/(1 - \gamma_1\rho_f)$, $L_s(x, \boldsymbol{\lambda}, \boldsymbol{z})$ is ρ_v-weakly convex with respect to variables $(x, \boldsymbol{\lambda})$ on for any fixed $\boldsymbol{z}$.

Proof: The proof is similar to the proof of **Lemma A.1** in Yao et al. (2024).

Lemma A.2 *Yao et al. (2024)* ***Lemma A.2*** *and* ***Lemma A.4***.

Under Assumption 4.3 and 4.4, let $\gamma_1 \in (0, 1/\rho_f)$ and $\gamma_2 > 0$. Then, for any $\rho_v \geq \rho_f/(1 - \gamma_1\rho_f)$, the following inequality holds:

$$-L_s(x_1, \boldsymbol{\lambda}, \boldsymbol{z}) \leq -L_s(x_2, \boldsymbol{\lambda}, \boldsymbol{z}) - \langle \nabla_x L_s(x_2, \boldsymbol{\lambda}, \boldsymbol{z}), x_1 - x_2 \rangle + \frac{\rho_v}{2}\|x_1 - x_2\|^2,$$
$$-L_s(x, \boldsymbol{\lambda}_1, \boldsymbol{z}) \leq -L_s(x, \boldsymbol{\lambda}_2, \boldsymbol{z}) - \langle \nabla_{\boldsymbol{\lambda}} L_s(x, \boldsymbol{\lambda}_2, \boldsymbol{z}), \boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2 \rangle + \frac{\rho_v}{2}\|\boldsymbol{\lambda}_1 - \boldsymbol{\lambda}_2\|^2,$$
$$-L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}_1) \leq -L_s(x, \boldsymbol{\lambda}, \boldsymbol{z}_2) - \langle \nabla_{\boldsymbol{\lambda}} L_s(x, \boldsymbol{\lambda}x, \boldsymbol{z}_2), \boldsymbol{z}_1 - \boldsymbol{z}_2 \rangle + \frac{L_z}{2}\|\boldsymbol{z}_1 - \boldsymbol{z}_2\|^2,$$

where $L_z := (\gamma_2\rho_T + 1)/(\gamma_2^2\rho_T)$.

Proof: The first 2 conclusions follow directly from *Lemma A.2* that $L_s(x, \boldsymbol{\lambda}, \boldsymbol{z})$ is ρ_v-weakly convex with respect to variables $(x, \boldsymbol{\lambda})$ on for any fixed $\boldsymbol{z}$, and the third conclusion is similar to the proof of **Lemma A.4** in Yao et al. (2024).

Lemma A.3 *Yao et al. (2024)* ***Lemma A.3***. *Under Assumption 4.3 and 4.4, let $\gamma_1 \in (0, 1/\rho_f)$ and $\gamma_2 > 0$. Then, for any $(x_1, \boldsymbol{\lambda}_1, \boldsymbol{z}_1)$ and $(x_2, \boldsymbol{\lambda}_2, \boldsymbol{z}_2)$, the following*

Lipschitz property holds:

$$\begin{aligned}&\|(\boldsymbol{\theta}^*(x_1,\boldsymbol{\lambda}_1,\boldsymbol{z}_1),\boldsymbol{\mu}^*(x_1,\boldsymbol{\lambda}_1,\boldsymbol{z}_1))-(\boldsymbol{\theta}^*(x_2,\boldsymbol{\lambda}_2,\boldsymbol{z}_2),\boldsymbol{\mu}^*(x_2,\boldsymbol{\lambda}_2,\boldsymbol{z}_2))\|\\ \leq&\frac{L_f+L_g+C_ZL_g}{\rho_T}\|x_1-x_2\|+\frac{1}{\gamma_1\rho_T}\|\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2\|+\frac{1}{\gamma_2\rho_T}\|\boldsymbol{z}_1-\boldsymbol{z}_2\|\\ \leq&L_{\theta,\mu}\|(x_1,\boldsymbol{\lambda}_1,\boldsymbol{z}_1)-(x_2,\boldsymbol{\lambda}_2,\boldsymbol{z}_2)\|,\end{aligned}$$

where $\rho_T:=\min\{1/\gamma_1-\rho_f,1/\gamma_2\}$, $C_Z=\max_{z\in Z}\|z\|$, *and* $L_{\theta,\mu}:=\sqrt{3}\max\{L_f+L_g+C_ZL_g,1/\gamma_1,1/\gamma_2\}/\rho_T$.

Proof: The proof is similar to the proof of **Lemma A.3** in Yao et al. (2024).

B Proof of main theorem and Lemmas

B.1 Proof of Theorem 4.6

Theorem 4.6 If Assumptions of Assumptions 4.2, 4.3 and 4.4 hold, let $\gamma_1\in(0,1/\rho_\gamma)$, $\gamma_2>0$, $c_t=\underline{c}(t+1)^p$ with $p\in(0,1/2)$ and $\underline{c}>0$. Pick $\eta_t\in(0,\rho_\gamma/L_B^2)$, then there exists $c_\alpha,c_\beta>0$ such that when $\alpha\in(\underline{\alpha},c_\alpha)$ and$\beta\in(\underline{\beta},c_\beta)$, with $\underline{\alpha},\underline{\beta}>0$, the sequence of $(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)},\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)})$ generated by Algorithm 1: MLM-CMOO satisfies

$$\min_t\left\|(\boldsymbol{\theta}^t,\boldsymbol{\mu}^t)-(\boldsymbol{\theta}_r^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}_r^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}))\right\|=\mathcal{O}(\frac{1}{\sqrt{T}}),$$

and

$$\min_t R_t(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})=\mathcal{O}(\frac{1}{\sqrt{T^{1-2p}}}).$$

Proof: First, using the descent lemma in *Lemma 4.5* and its condition, telescoping the inequality for $t=0,1,...,T-1$, we get

$$\begin{aligned}V_T-V_0\leq&-\frac{1}{4\alpha}\sum_{t=0}^{T-1}\left(\left\|x^{(t+1)}-x^{(t)}\right\|^2+\left\|\boldsymbol{\lambda}^{(t+1)}-\boldsymbol{\lambda}^{(t)}\right\|^2\right)-\frac{1}{4\beta}\sum_{t=0}^{T-1}\left\|\boldsymbol{z}^{(t+1)}-\boldsymbol{z}^{(t)}\right\|^2\\&-\eta\rho_TC_{\theta,\mu}\sum_{t=0}^{T-1}\left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)})-(\boldsymbol{\theta}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})\right\|^2.\end{aligned}$$

From assumptions, we have $\sum_{t=0}^{T-1}\left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)})-(\boldsymbol{\theta}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})\right\|^2$ is upper bounded, which is

$$\sum_{t=0}^{T-1}\left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)})-(\boldsymbol{\theta}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})\right\|^2\leq+\infty.$$

Thus, we have

$$\min_t\left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)})-(\boldsymbol{\theta}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})\right\|=\mathcal{O}(\frac{1}{\sqrt{T}}).$$

Secondly, according to the update rule of variables (x, y, z), we have

$$0 \in c_t(d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}) + \mathcal{M}_C(x^{(t)}, \boldsymbol{\lambda}^{(t)}) + \frac{c_t}{\alpha}((x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})),$$
$$0 \in c_t d_{\boldsymbol{z}}^{(t)} + \mathcal{M}_Z(\boldsymbol{z}^{(t+1)}) + \frac{c_t}{\beta}(\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}).$$

where $d_x^{(t)} = \frac{1}{c^{(t)}}\nabla_x F(x^{(t)}, \boldsymbol{\lambda}^{(t)}) + U + \sum_{i=1}^n \mu_i^{(t+1)} \nabla_x g_i(x^{(t)})$, $d_{\boldsymbol{\lambda}}^{(t)} = \frac{1}{c^{(t)}}(\boldsymbol{\lambda}^{(t)} - \boldsymbol{\lambda}') + \boldsymbol{V} - \frac{1}{\gamma_1}(\boldsymbol{\lambda})^{(t)} - \boldsymbol{\theta})^{(t+1)})$, and $d_{\boldsymbol{z}}^{(t)} = \boldsymbol{\mu}^{(t+1)} - \boldsymbol{z}^{(t)}$. Note, $U = \arg\min_u\{H(u, \boldsymbol{\lambda}^{(t)}) + \frac{1}{2}\left\|u - x^{(t)}\right\|^2\} - \arg\min_u\{H(u, \boldsymbol{\theta}^{(t+1)}) + \frac{1}{2}\left\|u - x^{(t)}\right\|^2\}$, $\boldsymbol{\lambda}' = \arg\min_{\boldsymbol{\lambda}} \left\|\sum_{i=1}^m \lambda_i \nabla f_i(x^{(t)})\right\|$, and $\boldsymbol{V} = \arg\min_v\{H(x^{(t)}, \boldsymbol{v}) + \frac{1}{2}\left\|\boldsymbol{v} - \boldsymbol{\lambda}^{(t)}\right\|^2\} - \arg\min_v\{H(u, \boldsymbol{\theta}^{(t+1)}) + \frac{1}{2}\left\|u - x^{(t)}\right\|^2\}$.

By the meanings of $d_x^{(t)}$, $d_{\boldsymbol{\lambda}}^{(t)}$, and $d_{\boldsymbol{z}}^{(t)}$, we obtain

$$\begin{aligned}(e_{x,\boldsymbol{\lambda}}^{(t)}, e_{\boldsymbol{z}}^{(t)}) \in& (\nabla F(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}), 0) + c_t(\sum_{i=1}^n \mu_i \nabla g_i(x^{(t+1)}), 0) \\ &- c_t(\nabla L_{i,s,r}(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t+1)}) + \mathcal{M}_{C\times Z}(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t+1)}),\end{aligned}$$

where

$$\begin{aligned}e_{x,\boldsymbol{\lambda}}^{(t)} :=& \nabla_{x,\boldsymbol{\lambda}}\phi_{c_t}(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}) - c_t(d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}) - \frac{c_t}{\alpha} - ((x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - ((x^{(t)}, \boldsymbol{\lambda}^{(t)})), \\ e_{\boldsymbol{z}}^{(t)} :=& \nabla_{\boldsymbol{z}}\phi_{c_t}(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}) - c_t(d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}) - c_t d_{\boldsymbol{z}}^{(t)} - \frac{c_t}{\beta} - (\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)})).\end{aligned}$$

Next, we estimate $\left\|e_{x,\boldsymbol{\lambda}}^{(t)}\right\|$. Using the estimates in Yao et al. (2024), we have

$$\begin{aligned}\left\|e_{x,\boldsymbol{\lambda}}^{(t)}\right\| \leq& c_t L_{\phi_1}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\| \\ &+ \frac{c_t}{\alpha}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\| + c_t C_{\phi_1} \\ &+ \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}), \boldsymbol{\mu}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\|,\end{aligned}$$

where $C_{\phi_1} := \sqrt{\max\{2(L_g + C_z L_g)^2, 2L_g^2\}}$.

For $\left\|e_{\boldsymbol{z}}^{(t)}\right\|$, we have

$$\left\|e_{\boldsymbol{z}}^{(t)}\right\| \leq (\frac{c_t}{\beta} + \frac{c_t}{\gamma_2})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\| + \frac{c_t}{\gamma_2}\left\|\boldsymbol{\mu}^{(t)} - \boldsymbol{\mu}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\|.$$

Thus,

$$\begin{aligned}R_t(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}) \leq& (\frac{c_t}{\beta} + \frac{c_t}{\gamma_2})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\| + \frac{c_t}{\alpha}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\| \\ &+ c_t(C_{\phi_1} + \frac{1}{\gamma_2})\left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}), \boldsymbol{\mu}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\| \\ &+ c_t L_{\phi_1}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\|.\end{aligned}$$

Let $\alpha_t \geq \underline{\alpha}$ and $\beta_t \geq \underline{\beta}$ for some positive constants $\underline{\alpha}$ and $\underline{\beta}$, we can show that there exists $C_R > 0$ such that

$$\frac{1}{c_t^2}R_t^2(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}) \leq C_R\left(\frac{1}{4\underline{\alpha}}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2 + \frac{1}{4\underline{\beta}}\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2\right.$$

$$+\eta\rho_T C_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)}),\boldsymbol{\mu}^*(x^{(t)},\boldsymbol{\lambda}^{(t)},\boldsymbol{z}^{(t)})\right\|^2\Big).$$

This completes the proof.

B.2 Proof of Lemma 4.7

Lemma 4.7. Under Assumption 4.3 and 4.4, let $\gamma_1 \in (0, 1/\rho_f)$, $\gamma_2 > 0$ and pick $\eta \in (0, \rho_T/L_b^2$, where $L_b := \max\{L_f + L_g + C_Z L_g + 1/\gamma_1, L_g + 1/\gamma_2\}$. Then, the sequence generated by Algorithm 1 satisfies

$$\begin{aligned}&\left\|(\boldsymbol{\theta}^{(t+1)},\boldsymbol{\mu}^{(t+1)}) - \big(\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}),\boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\big)\right\|\\ \le&(1-\eta\rho_T)\left\|(\boldsymbol{\theta}^{(t)},\boldsymbol{\mu}^{(t)}) - \big(\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}),\boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\big)\right\|.\end{aligned}$$

Proof: The proof is similar to the proof of **Lemma A.5** in Yao et al. (2024).

B.3 Proof of Lemma 4.8

Lemma 4.8. Suppose the assumption of 4.2, 4.3 and 4.4 hold, and let $\gamma_1 \in (0, 1/\rho_g)$, $\gamma_2 > 0$. Pick $\eta \in (0,\ \rho_\gamma/L_B^2)$ with $L_B := \max\{2L_g + C_z L_g + 1/\gamma_1,\ L_g + 1/\gamma_2\}$ then the sequence of $(\boldsymbol{\omega}^{(t)})$ generated by Algorithm 1: MLM-CMOO satisfies

$$\begin{aligned}\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \le \phi_{c_t}(\boldsymbol{\omega}^{(t)}) &- (\frac{1}{2\beta} - \frac{L_{v_z}}{2})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2\\ &- \left(\frac{1}{2\alpha} - \frac{L_{\phi_k}}{2} - \frac{\beta L_{\theta,\mu}^2}{\gamma_2^2}\right)\left(\left\|x^{(t+1)} - x^{(t)}\right\|^2 + \left\|\boldsymbol{\lambda}^{(t+1)} - \boldsymbol{\lambda}^{(t)}\right\|^2\right)\\ &+ \frac{\alpha}{2}\left(2(L_g + C_z L_g)^2 + \frac{1}{\gamma_1^2}\right)\left\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2\\ &+ (\alpha L_g^2 + \frac{\beta}{\gamma_2^2})\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2,\end{aligned}$$

where $L_{\phi_t} := L_f/c_t + L_g + \rho_v$.

Proof: Given Assumptions 4.2, 4.3, and4.4 that ∇F and ∇g are L_F- and L_g-Lipschitz continuous on their domain, respectively, and applying **Lemma 5.7** in Beck (2017)] and previous Lemmas, we obtain

$$\begin{aligned}\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \le&\phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \left\langle \nabla_{x,\boldsymbol{\lambda}}\phi_{c_t}(\boldsymbol{\omega}^{(t)}), (x^{(t+1)},\boldsymbol{\lambda}^{(t+1)}) - (x^{(t)},\boldsymbol{\lambda}^{(t)})\right\rangle\\ &+ \frac{L_{\phi_t}}{2}\left\|(x^{(t+1)},\boldsymbol{\lambda}^{(t+1)}) - (x^{(t)},\boldsymbol{\lambda}^{(t)})\right\|^2,\end{aligned}$$

with $L_{\phi_t} := L_F/c_t + L_g + \rho_v$. Based on the update rule of variable $x^{(t)}, \boldsymbol{\lambda}^{(t)}$, the convexity and the property of the proximal operator, we have

$$\left\langle (x^{(t)},\boldsymbol{\lambda}^{(t)}) - \alpha(d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}) - (x^{(t+1)},\boldsymbol{\lambda}^{(t+1)}), (x^{(t)},\boldsymbol{\lambda}^{(t)}) - (x^{(t+1)},\boldsymbol{\lambda}^{(t+1)})\right\rangle \le 0,$$

thus, we have

$$\left\langle (d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}), (x^{(t+1)},\boldsymbol{\lambda}^{(t+1)}) - (x^{(t)},\boldsymbol{\lambda}^{(t)})\right\rangle \le -\frac{1}{\alpha}\left\|(x^{(t+1)},\boldsymbol{\lambda}^{(t+1)}) - (x^{(t)},\boldsymbol{\lambda}^{(t)})\right\|^2.$$

Considering the formula of $\nabla_{x,\boldsymbol{\lambda}} L_{s,r}$ derived in **Lemma A.2** and the meanings of $d_x^{(t)}$, $d_{\boldsymbol{\lambda}}^{(t)}$ provided in the previous proof, we can obtain that

$$\begin{aligned}
&\left\|\nabla_{x,\boldsymbol{\lambda}} L_{s,r}(\boldsymbol{\omega}^{(t)}) - (d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)})\right\|^2 \\
=&\left\|\nabla_x H(x^{(t)}, \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})) + \sum_{i=1}^n \mu_i^*(\boldsymbol{\omega}^{(t)}) \nabla_x g(x^{(t)}) - \nabla_x H(x^{(t)}, \boldsymbol{\theta}^{(t+1)}) - \sum_{i=1}^n \mu_i^{(t+1)} \nabla_x g(x^{(t)})\right\|^2 \\
&+ \frac{1}{\gamma_1^2}\left\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\leq& 2\left\|\nabla_x H(x^{(t)}, \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})) + \sum_{i=1}^n \mu_i^*(\boldsymbol{\omega}^{(t)}) \nabla_x g(x^{(t)}) - \nabla_x H(x^{(t)}, \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})) - \sum_{i=1}^n \mu_i^{(t+1)} \nabla_x g(x^{(t)})\right\|^2 \\
&+ 2\left\|\nabla_x H(x^{(t)}, \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})) + \sum_{i=1}^n \mu_i^{(t+1)} \nabla_x g(x^{(t)}) - \nabla_x H(x^{(t)}, \boldsymbol{\theta}^{(t+1)}) - \sum_{i=1}^n \mu_i^{(t+1)} \nabla_x g(x^{(t)})\right\|^2 \\
&+ \frac{1}{\gamma_1^2}\left\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\leq& \left(2(L_f + C_Z L_g + \frac{1}{\gamma_1^2}\right)\left\|\boldsymbol{\theta}^{(t+1)}) - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 + 2L_g^2\left\|\boldsymbol{\mu}^{(t+1)}) - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2,
\end{aligned}$$

which yields

$$\begin{aligned}
&\left\langle \nabla_{x,\boldsymbol{\lambda}} L_{s,r}(\boldsymbol{\omega}^{(t)}) - (d_x^{(t)}, d_{\boldsymbol{\lambda}}^{(t)}), \boldsymbol{\lambda}^{(t+1)}), (x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)}) \right\rangle \\
\leq& \frac{\alpha}{2}\left(2(L_f + C_Z L_g + \frac{1}{\gamma_1^2}\right)\left\|\boldsymbol{\theta}^{(t+1)}) - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 + \alpha L_g^2\left\|\boldsymbol{\mu}^{(t+1)}) - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&+ \frac{1}{2\alpha}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2,
\end{aligned}$$

Combing with the above inequalities, we have

$$\begin{aligned}
\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \leq& \phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \left(\frac{1}{2\alpha} - \frac{L_{\phi_t}}{2}\right)\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2 \\
&+ \frac{\alpha}{2}\left(2(L_f + C_Z L_g + \frac{1}{\gamma_1^2}\right)\left\|\boldsymbol{\theta}^{(t+1)}) - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 + \alpha L_g^2\left\|\boldsymbol{\mu}^{(t+1)}) - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2
\end{aligned}$$

For variable $\boldsymbol{z}$, we have

$$\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \leq \phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \left\langle \nabla_{\boldsymbol{z}} \phi_{c_t}(\boldsymbol{\omega}^{(t)}), \boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)} \right\rangle + \frac{L_z}{2}\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2.$$

According to the property of the proximal gradient, we have

$$\left\langle d_{\boldsymbol{z}}^{(t)}, \boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)} \right\rangle \leq -\frac{1}{\beta}\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2$$

Thus, we have

$$\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \leq \phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \left\langle \nabla_{\boldsymbol{z}} \phi_{c_t}(\boldsymbol{\omega}^{(t)}) - d_{\boldsymbol{z}}^{(t)}, \boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)} \right\rangle + (\frac{L_z}{2} - \frac{1}{\beta})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2.$$

Based on the definition of $d_{\boldsymbol{z}}^{(t)}$ provided in the previous section, we have

$$\left\|\boldsymbol{\omega}^{(t)} - d_{\boldsymbol{z}}^{(t)}\right\|^2 \leq \frac{1}{\gamma_2^2}\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t)})\right\|^2,$$

and

$$\left\langle \nabla_{\boldsymbol{z}}\phi_{c_t}(\boldsymbol{\omega}^{(t)}) - d_{\boldsymbol{z}}^{(t)}, \boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\rangle \leq \frac{\beta}{2\gamma_2^2}\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t)})\right\|^2 + \frac{1}{2\beta}\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2$$

The, for variable $\boldsymbol{z}$, we can get

$$\begin{aligned}
\phi_{c_t}(\boldsymbol{\omega}^{(t+1)}) \leq& \phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \frac{\beta}{2\gamma_2^2}\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}, \boldsymbol{z}^{(t)})\right\|^2 + (\frac{L_z}{2} - \frac{1}{2\beta})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
\leq& \phi_{c_t}(\boldsymbol{\omega}^{(t)}) + \frac{\beta}{2\gamma_2^2}\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\|^2 + (\frac{L_z}{2} - \frac{1}{2\beta})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
&+ \frac{\beta L_{\boldsymbol{\theta},\boldsymbol{\mu}}^2}{2\gamma_2^2}\left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2.
\end{aligned}$$

Combining the inequities for variable $(x, \boldsymbol{\lambda})$ and $\boldsymbol{z}$, we can get Lemma 4.8.

B.4 Proof of Lemma 4.5

Lemma 4.5 Under Assumptions 4.2, 4.3 and 4.4 hold, let $\gamma_1 \in (0, 1/\rho_T)$, $\gamma_2 > 0$, $c_t \leq c_{t+1}$ and $\eta_t \in (\eta, \rho_\gamma/L_B^2)$ with $\eta > 0$, then there exist constants $c_\alpha, c_\beta > 0$ such that when $0 < \alpha \leq c_\alpha$ and$0 < \beta \leq c_\beta$, the sequence of $(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})$ generated by Algorithm 1: MLM-CMOO satisfies

$$\begin{aligned}
V_{t+1} - V_t \leq& -\frac{1}{4\alpha}\left\|x^{(t+1)} - x^{(t)}\right\|^2 - \frac{1}{4\alpha}\left\|\boldsymbol{\lambda}^{(t+1)} - \boldsymbol{\lambda}^{(t)}\right\|^2 - \frac{1}{4\beta}\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
&- \eta\rho_T C_{\theta,\mu}\left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)}), \boldsymbol{\mu}^*(x^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{z}^{(t)})\right\|^2.
\end{aligned}$$

Proof: From Lemma 4.8 and server aggregation rule, we have

$$\begin{aligned}
\phi_{c_t}(\boldsymbol{\omega}^{(t)}) \leq& \phi_{c_t}(\boldsymbol{\omega}^{(t)}) - (\frac{1}{2\beta} - \frac{L_{v_z}}{2})\left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
&- \left(\frac{1}{2\alpha} - \frac{L_{\phi_k}}{2} - \frac{\beta L_{\theta,\mu}^2}{\gamma_2^2}\right)\left(\left\|x^{(t+1)} - x^{(t)}\right\|^2 + \left\|\boldsymbol{\lambda}^{(t+1)} - \boldsymbol{\lambda}^{(t)}\right\|^2\right) \\
&+ \frac{\alpha}{2}\left(2(L_g + C_z L_g)^2 + \frac{1}{\gamma_1^2}\right)\left\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&+ (\alpha L_g^2 + \frac{\beta}{\gamma_2^2})\left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2.
\end{aligned} \tag{8}$$

Since $c_{t+1} \geq c_t$, we can infer that $(F(x^{(t)}, \boldsymbol{\lambda}^{(t)}) - \underline{F})/c_{t+1} \leq (F(x^{(t)}, \boldsymbol{\lambda}^{(t)}) - \underline{F})/c_t$. Combining with inequality equation 8 leads to

$$V_{t+1} - V_t = \phi_{c_{t+1}}(\boldsymbol{\omega}^{(t+1)})) - \phi_{c_t}(\boldsymbol{\omega}^{(t)})) + C_{\boldsymbol{\theta},\boldsymbol{\mu}}\left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t+1)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t+1)})\right\|^2$$

$$
\begin{aligned}
&- C_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 \\
\le& - (\frac{1}{2\alpha} - \frac{L_{\phi_t}}{2} - \frac{\beta L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2}) \left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2 - (\frac{1}{2\beta} - \frac{L_{v_z}}{2}) \left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
&+ (\alpha L_g^2 + \frac{\beta}{\gamma_2^2}) \left\|\boldsymbol{\mu}^{(t+1)} - \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 + \frac{\alpha}{2}\left(2(L_g + C_z L_g)^2 + \frac{1}{\gamma_1^2}\right) \left\|\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&+ C_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t+1)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t+1)})\right\|^2 \\
&- C_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 \\
\le& - (\frac{1}{2\alpha} - \frac{L_{\phi_t}}{2} - \frac{\beta L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2}) \left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2 - (\frac{1}{2\beta} - \frac{L_{v_z}}{2}) \left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2 \\
&+ C_{\theta,\lambda}\Big\{ - \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 + \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&+2\max\{\alpha,\beta\} \left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \Big\},
\end{aligned}
$$

where the last inequality follows from the fact that $C_{\theta,\lambda} := \max\{(L_g + C_Z L_g)^2 + 1/(2\gamma_1^2) + L_g^2,\ 1/\gamma_2^2\}$.

Then, for the last 3 terms in the previous equation, we have

$$
\begin{aligned}
&- \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 + \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&+ 2\alpha \left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\overset{a}{\le}&(1 + \frac{1}{\epsilon_t}) \left\|(\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t+1)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
&- \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 \\
&+ (1 + \epsilon_t + 2\alpha) \left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\overset{b}{\le}&(1 + \frac{1}{\epsilon_t}) L_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|\boldsymbol{\omega}^{(t+1)} - \boldsymbol{\omega}^{(t)}\right\|^2 - \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 \\
&+ (1 + \epsilon_t + 2\alpha)(1 - \eta\rho_T)^2 \left\|(\boldsymbol{\theta}^{(t+1)}, \boldsymbol{\mu}^{(t+1)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\le&(1 + \frac{2}{\eta\rho_T}) L^2_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|\boldsymbol{\omega}^{(t+1)} - \boldsymbol{\omega}^{(t)}\right\|^2 - \eta\rho_T \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2,
\end{aligned}
$$

where a from Lemma A.5 and A.7 for $\epsilon > 0$, and b from setting $\epsilon = \eta\rho_T/2$ and picking $\alpha \le \eta\rho_T/4$ where holds that $(1 + \epsilon + 2\alpha)(1 - \eta\rho_T) \le 1$.

Similarly, we can show that when $\beta \le \eta\rho_T/4$, it holds that

$$
\begin{aligned}
&- \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2 + \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \boldsymbol{\mu}^*(\boldsymbol{\omega}^{(t)})\right\|^2 \\
\le&(1 + \frac{2}{\eta\rho_T}) L^2_{\boldsymbol{\theta},\boldsymbol{\mu}} \left\|\boldsymbol{\omega}^{(t+1)} - \boldsymbol{\omega}^{(t)}\right\|^2 - \eta\rho_T \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2.
\end{aligned}
$$

Combining the above inequities, we have

$$
\begin{aligned}
V_{t+1} - V_t \le& - \left(\frac{1}{2\alpha} - \frac{L_{\phi_t}}{2} - \frac{\beta L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2} - (1 + \frac{2}{\eta\rho_T}) L^2_{\boldsymbol{\theta},\boldsymbol{\mu}} C_{\theta,\lambda}\right) \left\|(x^{(t+1)}, \boldsymbol{\lambda}^{(t+1)}) - (x^{(t)}, \boldsymbol{\lambda}^{(t)})\right\|^2 \\
&- \left(\frac{1}{2\beta} - \frac{L_{v_z}}{2} - (1 + \frac{2}{\eta\rho_T}) L^2_{\boldsymbol{\theta},\boldsymbol{\mu}} C_{\theta,\lambda}\right) \left\|\boldsymbol{z}^{(t+1)} - \boldsymbol{z}^{(t)}\right\|^2
\end{aligned}
$$

$$+ \eta\rho_T C_{\theta,\lambda} \left\|(\boldsymbol{\theta}^{(t)}, \boldsymbol{\mu}^{(t)}) - (\boldsymbol{\theta}^*(\boldsymbol{\omega}^{(t)}), \mu^*(\boldsymbol{\omega}^{(t)}))\right\|^2.$$

When $c_{t+1} \geq c_t$, $\eta \geq \underline{\eta} > 0$, $\alpha \leq \underline{\eta}\rho_T/4$ and $\beta \leq \underline{\eta}\rho_T/4$, then $\frac{L_{\phi_t}}{2} + \frac{\beta L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2} + (1 + \frac{2}{\eta\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} \leq \frac{L_{\phi_0}}{2} - \frac{\underline{\eta}\rho_T L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2} - (1 + \frac{2}{\underline{\eta}\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} =: C_\alpha$ and $\frac{L_{v_z}}{2} + (1 + \frac{2}{\eta\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} \leq \frac{L_{v_z}}{2} + (1 + \frac{2}{\underline{\eta}\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} =: C_\beta$

Consequently, if $C_\alpha, C_\beta > 0$ satisfies $C_\alpha \leq \min\left\{\frac{\underline{\eta}\rho_T}{4}, \frac{1}{4C_\alpha}\right\}$and $C_\beta \leq \min\left\{\frac{\underline{\eta}\rho_T}{4}, \frac{1}{4C_\beta}\right\}$, it holds that $\frac{L_{\phi_t}}{2} + \frac{\beta L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}}{\gamma_2^2} + (1 + \frac{2}{\eta\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} \geq \frac{1}{4\alpha}$ and $\frac{L_{v_z}}{2} + (1 + \frac{2}{\eta\rho_T})L^2_{\boldsymbol{\theta},\boldsymbol{\mu}}C_{\boldsymbol{\theta},\boldsymbol{\mu}} \geq \frac{1}{4\beta}$

This completes the proof.