

---

# Supplementary Material

---

**Xiaoqin Zhang, Di Wang**

Institute of Intelligent System and Decision  
Wenzhou University

zhangxiaoqinnan@gmail.com, wangdi@wzu.edu.cn

**Zhengyuan Zhou**

Department of Electrical Engineering  
Stanford University  
zyzhou@stanford.edu

**Yi Ma**

Visual computing group  
Microsoft Research Asia  
mayi@microsoft.com

## 1 Proof of convergence

In this section, we study the global convergence of the proposed algorithm for solving (1).

$$\begin{aligned} \min_{\mathcal{L}, \mathcal{E}, M_i, \Delta \tilde{\Gamma}} \quad & \sum_{i=1}^3 \alpha_i \|M_i\|_* + \gamma \|\mathcal{E}\|_1 \\ \text{s.t.} \quad & \mathcal{A} \circ \Gamma + \Delta \tilde{\Gamma} = \mathcal{L} + \mathcal{E} \\ & L_{(i)} = M_i, \quad i = 1, 2, 3 \end{aligned} \quad (1)$$

If all matrices and tensors are column-vectorized by taking their columns and stacking them on one another to form column vectors, and let  $\mathbf{m}_i$ ,  $\mathbf{l}$ ,  $\mathbf{e}$ ,  $\Delta \tilde{\boldsymbol{\tau}}$ ,  $\mathbf{q}_i$ ,  $\mathbf{y}$  and  $\mathbf{a}_0$  be the column vectors

of  $M_i$ ,  $\mathcal{L}$ ,  $\mathcal{E}$ ,  $\Delta \tilde{\Gamma}$ ,  $Q_i$ ,  $\mathcal{Y}$  and  $\mathcal{A} \circ \Gamma$  respectively. Let  $\mathbf{x} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \\ \mathbf{l} \\ \mathbf{e} \end{pmatrix}$ ,  $\boldsymbol{\eta} = \begin{pmatrix} \mathbf{y} \\ \mathbf{q}_1 \\ \mathbf{q}_2 \\ \mathbf{q}_3 \end{pmatrix}$ ,  $A =$

$\begin{pmatrix} 0 & 0 & 0 & I & I \\ -I & 0 & 0 & P_1 & 0 \\ 0 & -I & 0 & P_2 & 0 \\ 0 & 0 & -I & P_3 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} -I \\ 0 \\ 0 \\ 0 \end{pmatrix}$ , and  $\mathbf{d} = \begin{pmatrix} \mathbf{a}_0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ , where  $P_1$ ,  $P_2$  and  $P_3$  are permutation matrices. Then the optimization problem (1) becomes

$$\begin{aligned} \min_{\mathbf{x}, \Delta \tilde{\boldsymbol{\tau}}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & A\mathbf{x} + B\Delta \tilde{\boldsymbol{\tau}} = \mathbf{d}, \end{aligned} \quad (2)$$

where  $f(\mathbf{x})$  is vector form of  $\sum_{i=1}^3 \alpha_i \|M_i\|_* + \gamma \|\mathcal{E}\|_1$ .

Accordingly, the proximal gradient descent scheme in Section 3.2 can be written as

$$\begin{cases} \mathbf{x}^{k+1} := \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{1}{2\mu\tau_1} \|\mathbf{x} - (\mathbf{x}^k - \tau_1 A^\top (A\mathbf{x}^k + B\Delta \tilde{\boldsymbol{\tau}}^k - \mathbf{d} - \mu\boldsymbol{\eta}^k))\|^2 \end{cases} \quad (3)$$

$$\begin{cases} \Delta \tilde{\boldsymbol{\tau}}^{k+1} := \arg \min_{\Delta \tilde{\boldsymbol{\tau}}} \frac{1}{2\mu\tau_2} \|\Delta \tilde{\boldsymbol{\tau}} - (\Delta \tilde{\boldsymbol{\tau}}^k - \tau_2 B^\top (A\mathbf{x}^{k+1} + B\Delta \tilde{\boldsymbol{\tau}}^k - \mathbf{d} - \mu\boldsymbol{\eta}^k))\|^2 \end{cases} \quad (4)$$

$$\begin{cases} \boldsymbol{\eta}^{k+1} := \boldsymbol{\eta}^k - (A\mathbf{x}^{k+1} + B\Delta \tilde{\boldsymbol{\tau}}^{k+1} - \mathbf{d})/\mu \end{cases} \quad (5)$$

To prove the global convergence, we need to prove the following lemma.

**Lemma 1.** Assume that  $(\mathbf{x}^*, \Delta\tilde{\boldsymbol{\tau}}^*)$  is an optimal solution of (2) and  $\boldsymbol{\eta}^*$  is the corresponding optimal Lagrange multipliers. If the step sizes  $\tau_1 < 1/\lambda_{\max}(A^\top A)$  and  $\tau_2 < 1/\lambda_{\max}(B^\top B)$ , where  $\lambda_{\max}(C)$  denotes the largest eigenvalue of matrix  $C$ . Then the sequence  $(\mathbf{x}^k, \Delta\tilde{\boldsymbol{\tau}}^k, \boldsymbol{\eta}^k)$  generated by (3)-(5) satisfies:

(i) let  $z^k = \frac{1}{\mu\tau_1}\|\mathbf{x}^k - \mathbf{x}^*\|^2 - \frac{1}{\mu}\|A(\mathbf{x}^k - \mathbf{x}^*)\|^2 + \frac{1}{\mu\tau_2}\|\Delta\tilde{\boldsymbol{\tau}}^k - \Delta\tilde{\boldsymbol{\tau}}^*\|^2 + \mu\|\boldsymbol{\eta}^k - \boldsymbol{\eta}^*\|^2$ , the sequence  $\{z^k\}_{k=1}^\infty$  is monotonically non-increasing;

(ii)  $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \rightarrow 0$ ,  $\|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\| \rightarrow 0$ , and  $\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\| \rightarrow 0$ .

*Proof.* The optimality conditions for the subproblems (3) and (4) satisfy

$$-\frac{1}{\mu\tau_1}(\mathbf{x}^{k+1} - \mathbf{x}^k) - A^\top \left[ \frac{1}{\mu}(A\mathbf{x}^k + B\Delta\tilde{\boldsymbol{\tau}}^k - \mathbf{d}) - \boldsymbol{\eta}^k \right] \in \partial f(\mathbf{x}^{k+1}) \quad (6)$$

$$(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) + \tau_2 B^\top (A\mathbf{x}^{k+1} + B\Delta\tilde{\boldsymbol{\tau}}^k - \mathbf{d} - \mu\boldsymbol{\eta}^k) = \mathbf{0} \quad (7)$$

$(\mathbf{x}^*, \Delta\tilde{\boldsymbol{\tau}}^*, \boldsymbol{\eta}^*)$  also satisfy the following equations from the KKT conditions

$$A^\top \boldsymbol{\eta}^* \in \partial f(\mathbf{x}^*) \quad (8)$$

$$B^\top \boldsymbol{\eta}^* = \mathbf{0} \quad (9)$$

$$A\mathbf{x}^* + B\Delta\tilde{\boldsymbol{\tau}}^* = \mathbf{d} \quad (10)$$

Note that the fact that  $\partial f(\cdot)$  is a monotone operator, then

$$\langle \mathbf{x}^{k+1} - \mathbf{x}^*, \partial f(\mathbf{x}^{k+1}) - \partial f(\mathbf{x}^*) \rangle \geq 0 \quad (11)$$

Substitute (6) and (8) in (11) and use the updating formula (5), we have

$$\begin{aligned} 0 &\leq \langle \mathbf{x}^{k+1} - \mathbf{x}^*, -\frac{1}{\mu\tau_1}(\mathbf{x}^{k+1} - \mathbf{x}^k) - A^\top \left[ \frac{1}{\mu}(A\mathbf{x}^k + B\Delta\tilde{\boldsymbol{\tau}}^k - \mathbf{d}) - \boldsymbol{\eta}^k \right] - A^\top \boldsymbol{\eta}^* \rangle \\ &= \langle \mathbf{x}^{k+1} - \mathbf{x}^*, -\frac{1}{\mu\tau_1}(\mathbf{x}^{k+1} - \mathbf{x}^k) \rangle - \langle A(\mathbf{x}^{k+1} - \mathbf{x}^*), \frac{1}{\mu}(A\mathbf{x}^k + B\Delta\tilde{\boldsymbol{\tau}}^k - \mathbf{d}) - \boldsymbol{\eta}^k + \boldsymbol{\eta}^* \rangle \\ &= -\frac{1}{\mu\tau_1} \langle \mathbf{x}^{k+1} - \mathbf{x}^*, \mathbf{x}^{k+1} - \mathbf{x}^k \rangle + \langle A(\mathbf{x}^{k+1} - \mathbf{x}^*), \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^* \rangle \\ &\quad + \frac{1}{\mu} \langle A(\mathbf{x}^{k+1} - \mathbf{x}^*), A(\mathbf{x}^{k+1} - \mathbf{x}^k) \rangle + \frac{1}{\mu} \langle A(\mathbf{x}^{k+1} - \mathbf{x}^*), B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \end{aligned} \quad (12)$$

Multiply (7) by  $\frac{1}{\mu\tau_2}$  and sum (9), we have

$$\frac{1}{\mu\tau_2}(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) + B^\top \left[ \frac{1}{\mu}(A\mathbf{x}^{k+1} + B\Delta\tilde{\boldsymbol{\tau}}^k - \mathbf{d}) - \boldsymbol{\eta}^k + \boldsymbol{\eta}^* \right] = \mathbf{0} \quad (13)$$

Taking the inner product with  $\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^*$  on both sides of above equation and use the updating formula (5), we obtain

$$\begin{aligned} 0 &= \frac{1}{\mu\tau_2} \langle \Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^*, \Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k \rangle + \langle B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^*), \boldsymbol{\eta}^* - \boldsymbol{\eta}^{k+1} \rangle \\ &\quad - \frac{1}{\mu} \langle B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^*), B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \end{aligned} \quad (14)$$

Using the KKT condition (10) and updating formula (5), (12) minus (14) is

$$\begin{aligned} 0 &\leq -\frac{1}{\mu\tau_1} \langle \mathbf{x}^{k+1} - \mathbf{x}^*, \mathbf{x}^{k+1} - \mathbf{x}^k \rangle - \frac{1}{\mu\tau_2} \langle \Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^*, \Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k \rangle \\ &\quad - \mu \langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^* \rangle - \langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \\ &\quad + \frac{1}{\mu} \langle A(\mathbf{x}^{k+1} - \mathbf{x}^*), A(\mathbf{x}^{k+1} - \mathbf{x}^k) \rangle \end{aligned}$$

By using the fact

$$2\langle \mathbf{a}^{k+1} - \mathbf{a}^*, \mathbf{a}^{k+1} - \mathbf{a}^k \rangle = \|\mathbf{a}^{k+1} - \mathbf{a}^*\|^2 - \|\mathbf{a}^k - \mathbf{a}^*\|^2 + \|\mathbf{a}^{k+1} - \mathbf{a}^k\|^2$$

we obtain

$$z^{k+1} \leq z^k - \left[ \frac{1}{\mu\tau_1} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \frac{1}{\mu} \|A(\mathbf{x}^{k+1} - \mathbf{x}^k)\|^2 \right] - \left[ \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + 2\langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \right] \quad (15)$$

On the other hand, since  $\tau_1 < 1/\lambda_{\max}(A^\top A)$  and  $\tau_2 < 1/\lambda_{\max}(B^\top B)$ , the following equations hold:

$$\begin{aligned} & \frac{1}{\tau_1} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|A(\mathbf{x}^{k+1} - \mathbf{x}^k)\|^2 \geq \left( \frac{1}{\tau_1} - \lambda_{\max}(A^\top A) \right) \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \geq 0 \quad (16) \\ & \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + 2\langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \\ & \geq \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 - \left[ \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + \frac{1}{\mu} \|B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k)\|^2 \right] \\ & = \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 - \frac{1}{\mu} \|B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k)\|^2 \\ & \geq \frac{1}{\mu} \left( \frac{1}{\tau_2} - \lambda_{\max}(B^\top B) \right) \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 \geq 0 \quad (17) \end{aligned}$$

Combining (15), (16) and (17), the proof of (i) is complete.

As the sequence  $\{z_k\}_{k=1}^\infty$  is monotonically non-increasing and non-negative, it has a limit. Then from (15), we can see that

$$\begin{aligned} & \frac{1}{\tau_1} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|A(\mathbf{x}^{k+1} - \mathbf{x}^k)\|^2 \rightarrow 0 \\ & \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + 2\langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \rightarrow 0 \end{aligned}$$

Combining (16) and (17), we have

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \rightarrow 0, \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\| \rightarrow 0$$

Note that

$$\begin{aligned} & \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + 2\langle \boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k, B(\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k) \rangle \\ & = \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 + 2\langle B^\top(\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k), \Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k \rangle \\ & \geq \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 - \left[ \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k+1} - \Delta\tilde{\boldsymbol{\tau}}^k\|^2 + \mu\tau_2 \|B^\top(\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k)\|^2 \right] \\ & \geq \mu\tau_2 \left( \frac{1}{\tau_2} - \lambda_{\max}(B^\top B) \right) \|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\|^2 \end{aligned}$$

Thus  $\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\eta}^k\| \rightarrow 0$ . The proof of (ii) is complete.  $\square$

In the following, we will give the global convergence result of our proposed algorithm.

**Theorem 1.** *The sequence  $\{(\mathbf{x}^k, \Delta\tilde{\boldsymbol{\tau}}^k, \boldsymbol{\eta}^k)\}_{k=1}^\infty$  generated by Section 3.2 with  $\tau_1 < 1/5$  and  $\tau_2 < 1$  converges to an optimal solution to problem (1).*

*Proof.* It is easy to check that  $\lambda_{\max}(A^\top A) = 1$  and  $\lambda_{\max}(B^\top B) = 5$ .

From Lemma 1(i), we know the sequence  $\{(\mathbf{x}^k, \Delta\tilde{\boldsymbol{\tau}}^k, \boldsymbol{\eta}^k)\}_{k=1}^\infty$  is bounded. Hence, there exists a convergent subsequence such that  $\lim_{j \rightarrow \infty} (\mathbf{x}^{k_j}, \Delta\tilde{\boldsymbol{\tau}}^{k_j}, \boldsymbol{\eta}^{k_j}) = (\mathbf{x}_0, \Delta\tilde{\boldsymbol{\tau}}_0, \boldsymbol{\eta}_0)$ . Then from the updating formula (5) and Lemma 1(ii), we have

$$A\mathbf{x}_0 + B\Delta\tilde{\boldsymbol{\tau}}_0 - \mathbf{d} = 0 \quad (18)$$

By letting  $k = k_j - 1$  in (6) and  $k = k_j$  in (7), we have

$$\begin{aligned} & -\frac{1}{\mu\tau_1}(\mathbf{x}^{k_j} - \mathbf{x}^{k_j-1}) - A^\top \left\{ \frac{1}{\mu}(A\mathbf{x}^{k_j} + B\Delta\tilde{\boldsymbol{\tau}}^{k_j} - \mathbf{d}) - \boldsymbol{\eta}^{k_j} + \right. \\ & \quad \left. \frac{1}{\mu} [A(\mathbf{x}^{k_j-1} - \mathbf{x}^{k_j}) + B(\Delta\tilde{\boldsymbol{\tau}}^{k_j-1} - \Delta\tilde{\boldsymbol{\tau}}^{k_j})] + \boldsymbol{\eta}^{k_j} - \boldsymbol{\eta}^{k_j-1} \right\} \in \partial f(\mathbf{x}^{k_j}) \\ & (\Delta\tilde{\boldsymbol{\tau}}^{k_j+1} - \Delta\tilde{\boldsymbol{\tau}}^{k_j}) + \tau_2 B^\top (A\mathbf{x}^{k_j} + B\Delta\tilde{\boldsymbol{\tau}}^{k_j} - \mathbf{d} - \mu\boldsymbol{\eta}^{k_j} + A(\mathbf{x}^{k_j+1} - \mathbf{x}^{k_j})) = \mathbf{0} \end{aligned}$$

Let  $j \rightarrow \infty$ , by Lemma 1(ii), we have

$$A^\top \boldsymbol{\eta}_0 \in \partial f(\mathbf{x}_0) \quad (19)$$

$$B^\top \boldsymbol{\eta}_0 = \mathbf{0} \quad (20)$$

(18), (19) and (20) show that  $(\mathbf{x}_0, \Delta\tilde{\boldsymbol{\tau}}_0, \boldsymbol{\eta}_0)$  satisfies the KKT conditions for (2) and thus is an optimal solution to (2).

To complete the proof, we next show that the whole sequence  $\{(\mathbf{x}^k, \Delta\tilde{\boldsymbol{\tau}}^k, \boldsymbol{\eta}^k)\}_{k=1}^\infty$  converges to  $(\mathbf{x}_0, \Delta\tilde{\boldsymbol{\tau}}_0, \boldsymbol{\eta}_0)$ . By choosing  $(\mathbf{x}^*, \Delta\tilde{\boldsymbol{\tau}}^*, \boldsymbol{\eta}^*) = (\mathbf{x}_0, \Delta\tilde{\boldsymbol{\tau}}_0, \boldsymbol{\eta}_0)$  in Lemma 1 (i), we have

$$\frac{1}{\mu\tau_1} \|\mathbf{x}^{k_j} - \mathbf{x}_0\|^2 - \frac{1}{\mu} \|A(\mathbf{x}^{k_j} - \mathbf{x}_0)\|^2 + \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^{k_j} - \Delta\tilde{\boldsymbol{\tau}}_0\|^2 + \mu \|\boldsymbol{\eta}^{k_j} - \boldsymbol{\eta}_0\|^2 \rightarrow 0$$

By Lemma 1 (i), we know that the sequence  $\{\frac{1}{\mu\tau_1} \|\mathbf{x}^k - \mathbf{x}_0\|^2 - \frac{1}{\mu} \|A(\mathbf{x}^k - \mathbf{x}_0)\|^2 + \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^k - \Delta\tilde{\boldsymbol{\tau}}_0\|^2 + \mu \|\boldsymbol{\eta}^k - \boldsymbol{\eta}_0\|^2\}_{k=1}^\infty$  has a limit. Hence,

$$\frac{1}{\mu\tau_1} \|\mathbf{x}^k - \mathbf{x}_0\|^2 - \frac{1}{\mu} \|A(\mathbf{x}^k - \mathbf{x}_0)\|^2 + \frac{1}{\mu\tau_2} \|\Delta\tilde{\boldsymbol{\tau}}^k - \Delta\tilde{\boldsymbol{\tau}}_0\|^2 + \mu \|\boldsymbol{\eta}^k - \boldsymbol{\eta}_0\|^2 \rightarrow 0$$

So  $\lim_{k \rightarrow \infty} (\mathbf{x}^k, \Delta\tilde{\boldsymbol{\tau}}^k, \boldsymbol{\eta}^k) = (\mathbf{x}_0, \Delta\tilde{\boldsymbol{\tau}}_0, \boldsymbol{\eta}_0)$ . The proof is complete.  $\square$