

## SUPPLEMENTARY MATERIAL

### Appendix: Convergence analysis

In this section, we prove all the theorems in this paper.

#### 1. proof of Theorem 1

**Proof** From Eq. (9), we know that

$$\text{vec}(\nabla_{\mathcal{G}} H(\bar{\mathcal{G}}, A_{(N)}) - \nabla_{\mathcal{G}} H(\mathcal{G}, A_{(N)})) = \text{vec}((\bar{\mathcal{G}} - \mathcal{G})(\otimes_{i=N}^1 ((A_i)^T A_i))).$$

Since  $\|\mathcal{G}\|_F = \|\text{vec}(\mathcal{G})\|_F$ , thus we infer

$$\begin{aligned} & \|\nabla_{\mathcal{G}} H(\bar{\mathcal{G}}, A_{(N)}) - \nabla_{\mathcal{G}} H(\mathcal{G}, A_{(N)})\|_F \\ &= \|\text{vec}((\bar{\mathcal{G}} - \mathcal{G})(\otimes_{i=N}^1 ((A_i)^T A_i)))\|_F \\ &\leq \|\otimes_{i=N}^1 ((A_i)^T A_i)\|_F \|\text{vec}(\bar{\mathcal{G}} - \mathcal{G})\|_F \\ &= L_{\nabla_{\mathcal{G}} H} \|\bar{\mathcal{G}} - \mathcal{G}\|_F, \end{aligned}$$

and from Eq. (9), we also have

$$\begin{aligned} & \|\nabla_{A_n} H(\mathcal{G}, \{A_j\}_{j=1}^{n-1}, \bar{A}, \{A_j\}_{j=n+1}^N) - \nabla_{A_n} H(\mathcal{G}, A_{(N)})\|_F \\ &\leq L_{\nabla_{A_n} H} \|\bar{A} - A_n\|_F, \end{aligned}$$

where  $B_n = G^{(n)}(\otimes_{i=N, i \neq n}^1 A_i)^T$  and

$$L_{\nabla_{A_n} H} = \begin{cases} \|B_n(B_n)^T\|_F + \alpha \|L\|_F, & \text{if } n = N, \\ \|B_n(B_n)^T\|_F, & \text{else.} \end{cases}$$

Thus, Theorem 1 is true.

#### 2. proof of Theorem 2

**Proof** (i) If step 12 of Algorithm 2 is not true, then the Theorem 2(i) is true, if step 12 of Algorithm 2 is true, since Theorem 1 and Proposition 4, we know that

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_{(N)}^k) &\leq \langle \nabla_{\mathcal{G}} H(\mathcal{G}^k, A_{(N)}^k), \mathcal{G}^{k+1} - \mathcal{G}^k \rangle \\ &\quad + H(\mathcal{G}^k, A_{(N)}^k) \\ &\quad + \frac{L_{\nabla_{\mathcal{G}} H}}{2} \|\mathcal{G}^{k+1} - \mathcal{G}^k\|_F^2. \end{aligned} \quad (15)$$

From Eq. (10), we obtain

$$\begin{aligned} F_G(\mathcal{G}^k) &\geq F_G(\mathcal{G}^{k+1}) + \frac{1}{2\sigma_G^k} \|\mathcal{G}^{k+1} - \mathcal{G}^k\|_F^2 \\ &\quad + \langle \nabla_{\mathcal{G}} H(\mathcal{G}^k, A_{(N)}^k), \mathcal{G}^{k+1} - \mathcal{G}^k \rangle. \end{aligned} \quad (16)$$

Then sum of the Eq. (15) and Eq. (16), we have

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_{(N)}^k) &\leq H(\mathcal{G}^k, A_{(N)}^k) + F_G(\mathcal{G}^k) - F_G(\mathcal{G}^{k+1}) \\ &\quad - \left( \frac{1}{2\sigma_G^k} - \frac{L_{\nabla_{\mathcal{G}} H}}{2} \right) \|\mathcal{G}^{k+1} - \mathcal{G}^k\|_F^2. \end{aligned} \quad (17)$$

Similarly, since Theorem 1 and Proposition 4, we know that

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_1^{k+1}, \{A_i^k\}_{i=2}^N) &\leq \langle \nabla_{A_1} H(\mathcal{G}^{k+1}, A_{(N)}^k), A_1^{k+1} - A_1^k \rangle \\ &\quad + H(\mathcal{G}^{k+1}, A_{(N)}^k) \\ &\quad + \frac{L_{\nabla_{A_1} H}}{2} \|A_1^{k+1} - A_1^k\|_F^2. \end{aligned} \quad (18)$$

From Eq. (10), we can also obtain

$$\begin{aligned} F_1(A_1^k) &\geq F_1(A_1^{k+1}) + \frac{1}{2\sigma_1^k} \|A_1^{k+1} - A_1^k\|_F^2 \\ &\quad + \langle \nabla_{A_1} H(\mathcal{G}^{k+1}, A_{(N)}^k), A_1^{k+1} - A_1^k \rangle. \end{aligned} \quad (19)$$

Then sum of the Eq. (18), Eq. (19) and Eq. (17), we have

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_1^{k+1}, \{A_i^k\}_{i=2}^N) &\leq H(\mathcal{G}^k, A_{(N)}^k) + F_1(A_1^k) + F_G(\mathcal{G}^k) \\ &\quad - F_1(A_1^{k+1}) - F_G(\mathcal{G}^{k+1}) \\ &\quad - \rho \|A_1^{k+1} - A_1^k\|_F^2 \\ &\quad - \rho \|\mathcal{G}^{k+1} - \mathcal{G}^k\|_F^2, \end{aligned} \quad (20)$$

where  $\rho = \min(\frac{1}{2\sigma_G^k} - \frac{L_{\nabla_{\mathcal{G}} H}}{2}, \frac{1}{2\sigma_1^k} - \frac{L_{\nabla_{A_1} H}}{2})$ . Assuming Theorem 2(i) holds when  $n = N - 1$ , i.e.,

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_{(N-1)}^{k+1}, A_N^k) &\leq H(\mathcal{G}^k, A_{(N)}^k) + F_G(\mathcal{G}^k) + \sum_{i=1}^{N-1} F_i(A_i^k) \\ &\quad - \sum_{i=1}^{N-1} F_i(A_i^{k+1}) - F_G(\mathcal{G}^{k+1}) \\ &\quad - \rho \|A_{(N-1)}^{k+1} - A_{(N-1)}^k\|_F^2 \\ &\quad - \rho \|\mathcal{G}^{k+1} - \mathcal{G}^k\|_F^2, \end{aligned} \quad (21)$$

where  $\rho = \min(\frac{1}{2\sigma_G^k} - \frac{L_{\nabla_{\mathcal{G}} H}}{2}, \{\frac{1}{2\sigma_i^k} - \frac{L_{\nabla_{A_i} H}}{2}\}_{i=1}^{N-1})$ . Similarly, from Eq. (10), we infer

$$\begin{aligned} F_N(A_N^k) &\geq F_N(A_N^{k+1}) + \frac{1}{2\sigma_N^k} \|A_N^{k+1} - A_N^k\|_F^2 \\ &\quad + \langle \nabla_{A_N} H(\mathcal{G}^{k+1}, A_{(N-1)}^{k+1}, A_N^k), A_N^{k+1} - A_N^k \rangle. \end{aligned} \quad (22)$$

From Theorem 1 and Proposition 4, we also have

$$\begin{aligned} H(\mathcal{G}^{k+1}, A_{(N)}^{k+1}) &\leq H(\mathcal{G}^{k+1}, A_{(N-1)}^{k+1}, A_N^k) \\ &\quad + \langle \nabla_{A_N} H(\mathcal{G}^{k+1}, A_{(N-1)}^{k+1}, A_N^k), A_N^{k+1} - A_N^k \rangle \\ &\quad + \frac{L_{\nabla_{A_N} H}}{2} \|A_N^{k+1} - A_N^k\|_F^2. \end{aligned} \quad (23)$$

Then sum of the Eq. (21), Eq. (22) and Eq. (23), we have

$$J(\mathcal{G}^{k+1}, A_{(N)}^{k+1}) \leq J(\mathcal{G}^k, A_{(N)}^k) - \rho \|z^{k+1} - c^k\|_F^2, \quad (24)$$

where  $\rho = \min(\frac{1}{2\sigma_G^k} - \frac{L_{\nabla \mathcal{G}^k H}}{2}, \{\frac{1}{2\sigma_i^k} - \frac{L_{\nabla A_i^k H}}{2}\}_{i=1}^N)$ . This shows the Theorem 2(i) is true.

(ii) We know that  $\frac{\alpha}{2} \text{Tr}((A_N)^T L A_N)$  in Eq. (6) is non-negative. Hence, Eq. (6) is nonnegative. From Eq. (24), we have

$$\rho \|z^{k+1} - c^k\|_F^2 \leq J(\mathcal{G}^k, A_{(N)}^k) - J(\mathcal{G}^{k+1}, A_{(N)}^{k+1}).$$

Sum of both side, since Eq. (6) is nonnegative, thus we have

$$\begin{aligned} \rho \sum_{k=1}^{\infty} \|z^{k+1} - c^k\|_F^2 &\leq \sum_{k=1}^{\infty} (J(\mathcal{G}^k, A_{(N)}^k) - J(\mathcal{G}^{k+1}, A_{(N)}^{k+1})) \\ &= J(\mathcal{G}^1, A_{(N)}^1) - \inf J \\ &< \infty. \end{aligned}$$

It follows that

$$\lim_{k \rightarrow \infty} \|z^{k+1} - c^k\|_F^2 = 0.$$

### 3. proof of Lemma 1

**Proof** By Proposition 2 and Proposition 3, Eq. (10) follows that

$$0 \in \partial_{\mathcal{G}} F_G(\mathcal{G}^{k+1}) - \frac{1}{\sigma_G^k} (\mathcal{Y}^k - \mathcal{G}^{k+1}) + \nabla_{\mathcal{G}} H(\mathcal{Y}^k, A_{(N)}^k). \quad (25)$$

Thus we infer

$$\begin{aligned} f_G^{k+1} &= \nabla_{\mathcal{G}} H(\mathcal{G}^{k+1}, A_{(N)}^k) - \nabla_{\mathcal{G}} H(\mathcal{Y}^k, A_{(N)}^k) \\ &\quad + \frac{1}{\sigma_G^k} (\mathcal{Y}^k - \mathcal{G}^{k+1}) \\ &\in \nabla_{\mathcal{G}} H(\mathcal{G}^{k+1}, A_{(N)}^k) + \partial_{\mathcal{G}} (F_G(\mathcal{G}^{k+1}) + \sum_{i=1}^N F_i(A_i^k)) \\ &= \partial_{\mathcal{G}} J(\mathcal{G}^{k+1}, A_{(N)}^k). \end{aligned} \quad (26)$$

Similarly, by Proposition 2 and Proposition 3, Eq. (10) follows that

$$\begin{aligned} 0 &\in \partial_{A_i} F_i(A_i^{k+1}) - \frac{1}{\sigma_i^k} (Y_i^k - A_i^{k+1}) \\ &\quad + \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, Y_i^k, \{A_j^k\}_{j=i+1}^N). \end{aligned} \quad (27)$$

Thus we also have

$$\begin{aligned} f_i^{k+1} &= \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^i, \{A_j^k\}_{j=i+1}^N) \\ &\quad - \frac{1}{\sigma_i^k} (A_i^{k+1} - Y_i^k) \\ &\quad - \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, Y_i^k, \{A_j^k\}_{j=i+1}^N) \\ &\in \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, A_i^{k+1}, \{A_j^k\}_{j=i+1}^N) \\ &\quad + \partial_{A_i} (F_G(\mathcal{G}^{k+1}) + \sum_{j=1}^i F_j(A_j^{k+1}) + \sum_{j=i+1}^N F_j(A_j^k)) \\ &= \partial_{A_i} J(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, A_i^{k+1}, \{A_j^k\}_{j=i+1}^N). \end{aligned} \quad (28)$$

Since Theorem 1, we have

$$\begin{aligned} \|f_G^{k+1}\|_F &= \|\nabla_{\mathcal{G}} H(\mathcal{G}^{k+1}, A_{(N)}^k) + \frac{1}{\sigma_G^k} (\mathcal{Y}^k - \mathcal{G}^{k+1}) \\ &\quad - \nabla_{\mathcal{G}} H(\mathcal{Y}^k, A_{(N)}^k)\|_F, \\ &\leq \|\nabla_{\mathcal{G}} H(\mathcal{G}^{k+1}, A_{(N)}^k) - \nabla_{\mathcal{G}} H(\mathcal{Y}^k, A_{(N)}^k)\|_F \\ &\quad + \|\frac{1}{\sigma_G^k} (\mathcal{Y}^k - \mathcal{G}^{k+1})\|_F \\ &\leq (L_{\nabla_{\mathcal{G}} H} + \frac{1}{\sigma_G^k}) \|\mathcal{G}^{k+1} - \mathcal{Y}^k\|_F. \end{aligned} \quad (29)$$

Similarly, we also have

$$\begin{aligned} \|f_i^{k+1}\|_F &= \|\nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^i, \{A_j^k\}_{j=i+1}^N) \\ &\quad + \frac{1}{\sigma_i^k} (Y_i^k - A_i^{k+1}) \\ &\quad - \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, Y_i^k, \{A_j^k\}_{j=i+1}^N)\|_F, \\ &\leq \|\nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^i, \{A_j^k\}_{j=i+1}^N) \\ &\quad - \nabla_{A_i} H(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^{i-1}, Y_i^k, \{A_j^k\}_{j=i+1}^N)\|_F \\ &\quad + \|\frac{1}{\sigma_i^k} (A_i^{k+1} - Y_i^k)\|_F \\ &\leq (L_{\nabla_{A_i} H} + \frac{1}{\sigma_i^k}) \|A_i^{k+1} - Y_i^k\|_F. \end{aligned} \quad (30)$$

Thus we infer

$$\|f_G^{k+1}\|_F + \|\{f_i^{k+1}\}_{i=1}^N\|_F \leq \rho_b \|z^{k+1} - c^k\|_F, \quad (31)$$

where  $\rho_b = \max(\frac{1}{\sigma_G^k} + L_{\nabla_{\mathcal{G}} H}, \{\frac{1}{\sigma_i^k} + L_{\nabla_{A_i} H}\}_{i=1}^N)$ .

### 4. proof of Theorem 3

**Proof** (i) Observe that  $z'$  can be viewed as an intersection of compact sets

$$z' = \bigcap_{q \in \mathbb{N}} \overline{\bigcup_{k \geq q} \{z^k\}}.$$

From proposition 1,  $z'$  is compact.

(ii)  $\forall \bar{z} \in z'$ , there exists a subsequence  $z^{k_j}$  such that

$$\lim_{j \rightarrow \infty} z^{k_j} = \bar{z}.$$

Let

$$F(z^{k_j}) = \sum_{i=1}^N F_i(A_i^{k_j}) + F_G(\mathcal{G}^{k_j}). \quad (32)$$

Since  $F_i$  and  $F_G$  are lower semicontinuous [1], from Definition 3 and Eq. (32), we obtain that

$$\liminf_{j \rightarrow \infty} F(z^{k_j}) \geq F(\bar{z}). \quad (33)$$

843

844

845

846

847

848

849

850

851

852

853

854

855

856

Choosing  $k = k_j - 1$ , from Eq. (10), we infer

$$\begin{aligned} \lim_{j \rightarrow \infty} \sup F_G(\mathcal{G}^{k_j}) &\leq \lim_{j \rightarrow \infty} \sup (F_G(\bar{\mathcal{G}}) + \frac{1}{2\sigma_G^k} \|\bar{\mathcal{G}} - \mathcal{Y}^k\|_F^2 \\ &\quad + \langle \nabla_{\mathcal{G}} H(\mathcal{Y}^k, A_{(N)}^k), \bar{\mathcal{G}} - \mathcal{Y}^k \rangle), \\ \lim_{j \rightarrow \infty} \sup F_i(A_i^{k_j}) &\leq \lim_{j \rightarrow \infty} \sup (F_i(\bar{A}_i) + \frac{1}{2\sigma_i^k} \|\bar{A}_i - Y_i^k\|_F^2 \\ &\quad + \langle \nabla_{A_i} h_i(Y_i^k), \bar{A}_i - Y_i^k \rangle), \end{aligned} \quad (34)$$

where  $h_i(Y_i^k) = H(\mathcal{G}^{k+1}, \{A_n^{k+1}\}_{n=1}^{i-1}, Y_i^k, \{A_n^k\}_{n=i+1}^N)$ ,  $\lim_{j \rightarrow \infty} \mathcal{G}^{k_j} = \bar{\mathcal{G}}$ ,  $\lim_{j \rightarrow \infty} A_i^{k_j} = \bar{A}_i$ . Since  $\lim_{j \rightarrow \infty} z^{k_j} = \bar{z}$  and  $\lim_{j \rightarrow \infty} \|z^{k_j} - c^{k_j-1}\|_F = 0$  (Theorem 2(ii)), we can get that

$$\begin{aligned} \lim_{j \rightarrow \infty} \|\bar{A}_i - Y_i^k\|_F &\leq \lim_{j \rightarrow \infty} \|\bar{A}_i - A_i^{k_j}\|_F \\ &\quad + \lim_{j \rightarrow \infty} \|A_i^{k_j} - Y_i^{k_j-1}\|_F \\ &= 0, \\ \lim_{j \rightarrow \infty} \|\bar{\mathcal{G}} - \mathcal{Y}^k\|_F &\leq \lim_{j \rightarrow \infty} \|\bar{\mathcal{G}} - \mathcal{G}^{k_j}\|_F \\ &\quad + \lim_{j \rightarrow \infty} \|\mathcal{G}^{k_j} - \mathcal{Y}^{k_j-1}\|_F \\ &= 0. \end{aligned} \quad (35)$$

From Eq. (34) and Eq. (35), we infer

$$\begin{aligned} \lim_{j \rightarrow \infty} \sup F_i(A_i^{k_j}) &\leq \lim_{j \rightarrow \infty} \sup F_i(\bar{A}_i), \\ \lim_{j \rightarrow \infty} \sup F_G(\mathcal{G}^{k_j}) &\leq \lim_{j \rightarrow \infty} \sup F_G(\bar{\mathcal{G}}), \end{aligned} \quad (36)$$

where  $\lim_{j \rightarrow \infty} \mathcal{G}^{k_j} = \bar{\mathcal{G}}$ ,  $\lim_{j \rightarrow \infty} A_i^{k_j} = \bar{A}_i$ . Therefore, from Eq. (32) and Eq. (36), we have

$$\lim_{j \rightarrow \infty} \sup F(z^{k_j}) \leq \lim_{j \rightarrow \infty} \sup F(\bar{z}). \quad (37)$$

From Eq. (32), Eq. (37) and Eq. (33), we infer

$$\lim_{j \rightarrow \infty} F(z^{k_j}) = F(\bar{z}). \quad (38)$$

Since Theorem 2, we have

$$\lim_{j \rightarrow \infty} H(z^{k_j}) = H(\bar{z}). \quad (39)$$

Thus from Eq. (38) and Eq. (39), we infer

$$\begin{aligned} \lim_{j \rightarrow \infty} H(z^{k_j}) + \lim_{j \rightarrow \infty} F(z^{k_j}) &= \lim_{j \rightarrow \infty} (H(z^{k_j}) + F(z^{k_j})) \\ &= \lim_{j \rightarrow \infty} J(z^{k_j}) \\ &= J(\bar{z}). \end{aligned}$$

This means  $J$  is constant on  $\bar{z}$ .

(iii) From Eq. (26) and Eq. (28), we know that

$$\begin{aligned} f_G^{k+1} &\in \partial_{\mathcal{G}} J(\mathcal{G}^{k+1}, A_{(N)}^k), \\ f_i^{k+1} &\in \partial_{A_i} J(\mathcal{G}^{k+1}, \{A_j^{k+1}\}_{j=1}^i, \{A_j^k\}_{j=i+1}^N), \end{aligned} \quad (40)$$

where the definitions of  $f_i^{k+1}$  and  $f_G^{k+1}$  are the same as in Eq. (12). From Theorem 2(ii) and Lemma 1, we infer

$$\lim_{k \rightarrow \infty} (\|\{f_i^{k+1}\}_{i=1}^N\|_F + \|f_G^{k+1}\|_F) \leq \lim_{k \rightarrow \infty} \rho_b \|z^{k+1} - c^k\|_F = 0. \quad (41)$$

Let

$$z_1^k = \mathcal{G}^k, z_{i+1}^k = A_i^k. \quad (42)$$

From Theorem 2(i), we know that

$$\begin{aligned} \lim_{k \rightarrow \infty} J(\{z_i^{k+1}\}_{i=1}^{j-1}, z_j^{k+1}, \{z_i^k\}_{i=j+1}^{N+1}) &= \lim_{k \rightarrow \infty} J(z^k) \\ &= J(z^*). \end{aligned} \quad (43)$$

Therefore, from Eq. (40), Eq. (41), Eq. (42) and Eq. (43), we have

$$\begin{aligned} 0 &\in \partial_{z_j} \lim_{k \rightarrow \infty} J(\{z_i^{k+1}\}_{i=1}^{j-1}, z_j^{k+1}, \{z_i^k\}_{i=j+1}^{N+1}) \\ &= \partial_{z_j} J(z^*). \end{aligned}$$

This means  $0 \in \partial J(z^*)$ .

## 5. proof of Theorem 4

**Proof** From Theorem 1, we know that  $\frac{\alpha}{2} \text{Tr}((A_N)^T L A_N)$  is an infinitely differentiable function, and the norm of its derivatives of any order is also continuous, so  $\frac{\alpha}{2} \text{Tr}((A_N)^T L A_N)$  is a real analytic function (real analytic functions are all KŁ functions [2]). Since the Frobenius norm,  $\ell_p$ -norm and Eq. (7) are also all KŁ functions [1], it follows that Eq. (6) is a KŁ function. Therefore, from Definition 6, there exists a concave function  $\phi$  so that

$$\phi'(J(z^k) - J(\bar{z})) \text{dist}(0, \partial J(z^k)) \geq 1. \quad (44)$$

From  $\phi$  is the convex function, we have

$$\begin{aligned} \phi(J(z^{k+1}) - J(\bar{z})) &\leq \phi(J(z^k) - J(\bar{z})) \\ &\quad + \phi'(J(z^k) - J(\bar{z}))(J(z^{k+1}) - J(z^k)). \end{aligned} \quad (45)$$

From Lemma 1 and Theorem 3, we infer

$$\text{dist}(0, \partial J(z^k)) \leq \rho_b \|z^k - c^{k-1}\|_F. \quad (46)$$

Since the Eq. (44) and Eq. (46), we infer

$$\begin{aligned} \phi'(J(z^k) - J(\bar{z})) &\geq \frac{1}{\text{dist}(0, \partial J(z^k))} \\ &\geq \frac{1}{\rho_b \|z^k - c^{k-1}\|_F}. \end{aligned} \quad (47)$$

Let  $J(k) = J(z^k) - J(\bar{z})$ , from Eq. (45), Eq. (46), Eq. (47) and Theorem 2(i), we have

$$\begin{aligned} \phi(J(k)) - \phi(J(k+1)) &\geq \phi'(J(k))(J(k) - J(k+1)) \\ &\geq \frac{J(k) - J(k+1)}{\rho_b \|z^k - c^{k-1}\|_F} \\ &\geq \frac{\rho \|z^{k+1} - c^k\|_F^2}{\rho_b \|z^k - c^{k-1}\|_F}. \end{aligned}$$

Define  $C = \frac{\rho}{\rho_b}$ ,  $C$  is a constant, so we infer

$$\|z^{k+1} - c^k\|_F^2 \leq C(\phi(J(k)) - \phi(J(k+1)))\|z^k - c^{k-1}\|_F.$$

Using the fact that  $2ab \leq a^2 + b^2$

$$2\|z^{k+1} - c^k\|_F \leq C(\phi(J(k)) - \phi(J(k+1))) + \|z^k - c^{k-1}\|_F.$$

Sum both sides

$$\begin{aligned} 2 \sum_{k=l+1}^K \|z^{k+1} - c^k\|_F &\leq \sum_{k=l+1}^K \|z^k - c^{k-1}\|_F \\ &\quad + C(\phi(J(l+1)) - \phi(J(K+1))) \\ &= C(\phi(J(l+1)) - \phi(J(K+1))) \\ &\quad + \|z^{l+1} - c^l\|_F + \sum_{k=l+1}^K \|z^{k+1} - c^k\|_F. \end{aligned} \quad (48)$$

From Eq. (48), we can get that

$$\begin{aligned} \lim_{K \rightarrow \infty} \sum_{k=l+1}^K \|z^{k+1} - c^k\|_F &\leq \|z^{l+1} - c^l\|_F + C\phi(J(l+1)) \\ &\quad - \lim_{K \rightarrow \infty} C\phi(J(K+1)) \\ &< \infty. \end{aligned}$$

Thus we have

$$\sum_{k=l+1}^{\infty} \|z^{k+1} - c^k\|_F < \infty. \quad (49)$$

From Algorithm 2, no matter  $c^k = z^k + \beta_k(z^k - z^{k-1})$  or  $c^k = z^k$ , we always have

$$\|z^{k+1} - z^k\|_F - \beta_k \|z^k - z^{k-1}\|_F \leq \|z^{k+1} - c^k\|_F. \quad (50)$$

From Eq. (49) and Eq. (50), we know that

$$\sum_{k=l+1}^{\infty} (\|z^{k+1} - z^k\|_F - \beta_k \|z^k - z^{k-1}\|_F) \leq \sum_{k=l+1}^{\infty} \|z^{k+1} - c^k\|_F < \infty.$$

Since

we also have

$$\begin{aligned} \|z^{k+1} - z^k\|_F - \beta_{max} \|z^k - z^{k-1}\|_F &\leq \\ \|z^{k+1} - z^k\|_F - \beta_k \|z^k - z^{k-1}\|_F, \end{aligned}$$

$$\begin{aligned} \sum_{k=l+1}^{\infty} (\|z^{k+1} - z^k\|_F - \beta_{max} \|z^k - z^{k-1}\|_F) \\ \leq \sum_{k=l+1}^{\infty} (\|z^{k+1} - z^k\|_F - \beta_k \|z^k - z^{k-1}\|_F) \\ < \infty, \end{aligned} \quad (51)$$

and

$$\begin{aligned} \sum_{k=l+1}^{\infty} \|z^{k+1} - z^k\|_F - \sum_{k=l+1}^{\infty} \beta_{max} \|z^k - z^{k-1}\|_F \\ = \sum_{k=l+1}^{\infty} (\|z^{k+1} - z^k\|_F - \beta_{max} \|z^{k+1} - z^k\|_F) \\ + \beta_{max} \|z^{l+1} - z^l\|_F \\ = \sum_{k=l+1}^{\infty} (1 - \beta_{max}) \|z^{k+1} - z^k\|_F + \beta_{max} \|z^{l+1} - z^l\|_F. \end{aligned} \quad (52)$$

Let  $s^k = \|z^{k+1} - z^k\|_F$ , from Eq. (51) and Eq. (52), we have

$$\begin{aligned} \sum_{k=l+1}^{\infty} (1 - \beta_{max}) s^k - \beta_{max} s^l &= \sum_{k=l}^{\infty} (s^{k+1} - \beta_{max} s^k) \\ &\leq \sum_{k=l+1}^{\infty} \|z^{k+1} - c^k\|_F \\ &< \infty. \end{aligned}$$

Thus we infer

$$\sum_{k=l+1}^{\infty} (1 - \beta_{max}) s^k < \infty. \quad (53)$$

From  $(1 - \beta_{max})$  is a positive constant and Eq. (53), we infer

$$\sum_{k=l+1}^{\infty} \|z^{k+1} - z^k\|_F = \sum_{k=l+1}^{\infty} s^k < \infty.$$

Thus we have

$$\sum_{k=0}^{\infty} \|z^{k+1} - z^k\|_F = \sum_{k=0}^{\infty} s^k < \infty.$$

This shows that

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \|z^{K+p} - z^K\|_F &= \lim_{K \rightarrow \infty} \left\| \sum_{k=K+1}^{K+p} (z^{k+1} - z^k) \right\|_F \\
 &\leq \lim_{K \rightarrow \infty} \sum_{k=K+1}^{K+p} \|z^{k+1} - z^k\|_F \\
 &= \lim_{K \rightarrow \infty} \sum_{k=K+1}^{\infty} s^k \\
 &= 0.
 \end{aligned} \tag{54}$$

This means the Theorem 4 is true.

## 6. proof of Theorem 5

**Proof** From Appendix B of [3] and Section 3.5 of [4], we know that the computational cost of  $\nabla_{A_i} H(\mathcal{G}, A_{(N)})$  is

$$\mathcal{O}\left(\sum_{j=1}^N \left(\sum_{i=1}^j I_i\right) \left(\sum_{i=j}^N r_i\right)\right). \tag{55}$$

Similarly, the computational cost of  $\nabla_{\mathcal{G}} H(\mathcal{G}, A_{(N)})$  is

$$\mathcal{O}\left(\sum_{j=1}^N \left(\sum_{i=1}^j r_i\right) \left(\sum_{i=j}^N I_i\right)\right). \tag{56}$$

The cost of computing the Lipschitz constant, projection to nonnegative, and tensor unfolding is negligible compared to the cost of computing partial gradients  $\nabla_{\mathcal{G}} H(\mathcal{G}, A_{(N)})$  and  $\nabla_{A_j} H(\mathcal{G}, A_{(N)})$ . Similarly, the cost of the sparsity projection is also negligible compared to the cost of computing these partial gradients, as the time complexity of finding the largest  $s$  elements in an array of  $p$  elements is  $\mathcal{O}(p + s \log_2(s))$ . Therefore, from Eq. (55) and Eq. (56), the time complexity of the Algorithm 1 and Algorithm 2 in each iteration is approximately estimated as follows.

$$\mathcal{O}\left(\sum_{j=1}^N \left(\sum_{i=1}^j r_i\right) \left(\sum_{i=j}^N I_i\right) + N \left(\sum_{j=1}^N \left(\sum_{i=1}^j I_i\right) \left(\sum_{i=j}^N r_i\right)\right)\right).$$

## References

- [1] J. Bolte, S. Sabach, and M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program.*, vol.146, no.1, pp.459–494, 2014.
- [2] Y. Xu and W. Yin, A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion, *Siam. J. Imaging. Sci.*, vol.6, no.3, pp.1758–1789, 2013.
- [3] Y. Xu, Alternating proximal gradient method for sparse nonnegative Tucker decomposition, *Mathematical Programming Computation*, vol.7, pp.39–70, 2015.
- [4] W. Gong, Z. Huang, and L. Yang, Accurate regularized Tucker decomposition for image restoration, *Appl. Math. Model.*, vol.123, pp.75–86, 2023.