Nonconvex Multiview Subspace Clustering Framework With Efficient Method Designs and Theoretical Analysis
Supplementary File

Zhi Wang, **Zhuo Liu**, **Dong Hu** and **Tao Jia***

College of Computer and Information Science, Southwest University, Chongqing, China

chiw@swu.edu.cn, zhuoliu.swu@gmail.com, donghu@email.swu.edu.cn, tjia@swu.edu.cn

This is the supplementary file for paper entitled "Nonconvex Multiview Subspace Clustering Framework With Efficient Method Designs and Theoretical Analysis", which gives detailed proofs of Theorem 1 presented in the paper.

Before starting the proof of Theorem 1, we first introduce some preliminary knowledge, which is usable in subsequent proofs.

Definition 1 (ℓ_q $(0 < q < 1)$ norm [Marjanovic and Solo, 2012])**.** *Given a matrix* $\mathbf{X} \in \mathbb{R}^{d\times n}$*, then its* ℓ_q *norm with* $0 < q < 1$ *can be defined as*

$$\|\mathbf{X}\|_q^q = \sum_{i=1}^{\min(d,n)} \sigma_i(\mathbf{X})^q, \tag{1}$$

where $\sigma_i(\mathbf{X})$ *denotes the* i*-th singular value of* $\mathbf{X}$.

Lemma 1 ([Lewis and Sendov, 2005])**.** *Suppose that* $F(\cdot) : \mathbb{R}^{d\times n} \to \mathbb{R}$ *be denoted by* $F(\mathbf{X}) = f \circ \boldsymbol{\sigma}(\mathbf{X})$ *and the function* $f(\cdot) : \mathbb{R}^r \to \mathbb{R}$ *be absolutely symmetric and differentiable at* $\boldsymbol{\sigma}(\mathbf{X})$*. Let the SVD of matrix* $\mathbf{X}$ *be* $\mathbf{U}\,\mathrm{Diag}(\boldsymbol{\sigma}(\mathbf{X}))\mathbf{V}^T$*, then the subdifferential of* $F(\mathbf{X})$ *(i.e.,* $f \circ \boldsymbol{\sigma}$*) at a matrix* $\mathbf{X}$ *can be given by*

$$\frac{\partial F(\mathbf{X})}{\partial \mathbf{X}} = \partial(f \circ \boldsymbol{\sigma})(\mathbf{X}) = \mathbf{U}\,\mathrm{Diag}(\partial f(\boldsymbol{\sigma}(\mathbf{X})))\mathbf{V}^T, \tag{2}$$

where $\partial f(\boldsymbol{\sigma}(\mathbf{X})) = (\partial f(\sigma_1(\mathbf{X}))/\partial \mathbf{X}, \ldots, f(\sigma_r(\mathbf{X}))/\partial \mathbf{X})$*, and each* $f(\sigma_i(\mathbf{X}))/\partial \mathbf{X} = \delta_i f'(\sigma_i(\mathbf{X}))$ *with* $\delta_i \in \partial|\sigma_i(\mathbf{X})|$.

Lemma 2. *Let the SVD of matrix* $\mathbf{X} \in \mathbb{R}^{d\times n}$ *be* $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$*, then the subdifferential of* ℓ_q $(0 < q < 1)$ *norm can be given by*

$$\partial\|\mathbf{X}\|_q^q = \mathbf{U}\,\mathrm{Diag}(\{\delta_i f'(\boldsymbol{\Sigma}_{i,i})\}_{0\le i\le r})\mathbf{V}^T, \tag{3}$$

where $r = \min(d, n)$*,* $f'(\boldsymbol{\Sigma}_{i,i}) = q(\boldsymbol{\Sigma}_{i,i} + \epsilon)^{q-1}$ *and* $\delta_i \in \partial|\boldsymbol{\Sigma}_{i,i} + \epsilon|$ *with a small enough constant* $\epsilon > 0$.

Proof. According to the Definition 1, $\|\mathbf{X}\|_q^q = \sum_{i=1}^{\min(d,n)}(\boldsymbol{\Sigma}_{i,i})^q$ can be regarded as a function with independent variable $\{\boldsymbol{\Sigma}_{i,i} | 1 \le i \le \min(d, n)\}$. It follows Lemma 1 that the subdifferential of $\|\mathbf{X}\|_q^q$ at $\mathbf{X}$ can be derived as

$$\begin{aligned}\partial\|\mathbf{X}\|_q^q &= \mathbf{U}\,\mathrm{Diag}(\partial f(\boldsymbol{\Sigma}))\mathbf{V}^T \\ &= \mathbf{U}\,\mathrm{Diag}(\{\delta_i f'(\boldsymbol{\Sigma}_{i,i})\}_{0\le i\le r})\mathbf{V}^T,\end{aligned} \tag{4}$$

where $r = \min(d, n)$, $f'(\boldsymbol{\Sigma}_{i,i}) = q(\boldsymbol{\Sigma}_{i,i})^{q-1}$ with $\delta_i \in \partial|\boldsymbol{\Sigma}_{i,i}|$ when $\boldsymbol{\Sigma}_{i,i} \neq 0$. In order to handle the case of $\boldsymbol{\Sigma}_{i,i} = 0$ (denominator is zero), $q(\boldsymbol{\Sigma}_{i,i})^{q-1}$ can be approximated by $q(\boldsymbol{\Sigma}_{i,i}+\epsilon)^{q-1}$, where $\epsilon > 0$ denotes a small enough constant, and naturally $\delta_i \in \partial|\boldsymbol{\Sigma}_{i,i} + \epsilon|$. □

Next, we will give a detailed proof of Theorem 1.

Proof. **1)** First, we prove that the generated sequence $\{\Theta^k = (\mathbf{C}_v^k,\ \mathbf{P}_v^k,\ \mathbf{S}_v^k,\ \mathbf{E}_v^k;\ \mathbf{M}_{v,1}^k,\ \mathbf{M}_{v,2}^k,\ \mathbf{M}_{v,3}^k) | v = 1, \ldots, V\}_{k=1}^{\infty}$ is bounded when $\lim_{k\to\infty}(\mathcal{G}^{k+1} - \mathcal{G}^k) = \mathbf{0}$ with $\mathcal{G} = \{\mathbf{C}_v, \mathbf{E}_v\}$ in any view holds. At the $(k + 1)$th iteration in each view, the optimality condition of $\mathbf{P}_v^{k+1} \in \mathbb{R}^{n\times n}$ holds by following its update rule, that is

$$\mathbf{0} \in \partial\|\mathbf{P}_v^{k+1}\|_q^q + \mu^k(\mathbf{P}_v^{k+1} - (\mathbf{C}_v^k + \mathbf{M}_{v,1}^k/\mu^k)). \tag{5}$$

Using the rule $\mathbf{M}_{v,1}^{k+1} = \mathbf{M}_{v,1}^k + \mu^k(\mathbf{C}_v^{k+1} - \mathbf{P}_v^{k+1})$, equation (5) can be derived as

$$\mathbf{M}_{v,1}^{k+1} \in \partial\|\mathbf{P}_v^{k+1}\|_q^q + \mu^k(\mathbf{C}_v^{k+1} - \mathbf{C}_v^k). \tag{6}$$

Let $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ be the SVD of matrix $\mathbf{P}_v^{k+1}$, it follows Lemma 2 that

$$\begin{aligned}\|\partial\|\mathbf{P}_v^{k+1}\|_q^q\|_F^2 &= \|\mathbf{U} \times \partial f(\boldsymbol{\Sigma}) \times \mathbf{V}^T\|_F^2 \\ &= \|\partial f(\boldsymbol{\Sigma})\|_F^2 \\ &= \sum_{i=1}^{n} [\partial f(\boldsymbol{\Sigma}_{i,i})]^2 \\ &= \sum_{i=1}^{n} \delta_i^2 q^2(\boldsymbol{\Sigma}_{i,i} + \epsilon)^{2(q-1)} \\ &\le nq^2\epsilon^{2(q-1)},\end{aligned} \tag{7}$$

where the inequality holds by $0 \le \delta_i^2 \le 1$ and $\boldsymbol{\Sigma}_{i,i} \ge 0$. Thus $\partial\|\mathbf{P}_v^{k+1}\|_q^q$ is bounded. Combining it with $\lim_{k\to\infty}(\mathbf{C}_v^{k+1} - \mathbf{C}_v^k) = \mathbf{0}$, the boundedness of $\{\mathbf{M}_{v,1}^{k+1}\}_{v=1}^V$ can be easily verified.

Similarly, in line with the update rule of $\mathbf{S}_v^{k+1} \in \mathbb{R}^{n\times n}$, its optimality condition can be satisfied, such that

$$\mathbf{0} \in \lambda\partial\|\mathbf{W} \odot \mathbf{S}_v^{k+1}\|_q^q + \mu^k(\mathbf{S}_v^{k+1} - (\mathbf{C}_v^k + \mathbf{M}_{v,2}^k/\mu^k)). \tag{8}$$

*Corresponding author.

Then, we can deduce

$$\mathbf{M}_{v,2}^{k+1} \in \lambda\partial\|\mathbf{W}\odot\mathbf{S}_v^{k+1}\|_q^q + \mu^k(\mathbf{C}_v^{k+1}-\mathbf{C}_v^k) \tag{9}$$

by the update rule $\mathbf{M}_{v,2}^{k+1} = \mathbf{M}_{v,2}^{k} + \mu^k(\mathbf{C}_v^{k+1}-\mathbf{S}_v^{k+1})$. Then we set $\mathbf{\Gamma}_v^{k+1} = \mathbf{W}\odot\mathbf{S}_v^{k+1}$ due to $\mathbf{W}$ is a predefined and fixed matrix. Different from the purpose of learning the lowest rank coefficient representation by the first term of our model, the sparse structural constraint (the second term) in our model plays a role of exploring the intrinsic and local subspace structure across views. Hence $\|\mathbf{\Gamma}_v^{k+1}\|_q^q$ is defined as $\sum_{i,j}\left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right|^q$. Then, we can obtain

$$\begin{aligned}\|\lambda\partial\|\mathbf{\Gamma}_v^{k+1}\|_q^q\|_F^2 &= \lambda^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left[q\left(\left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right|+\epsilon\right)^{q-1}\right]^2 \\ &\le \lambda^2 n^2 q^2 \epsilon^{2(q-1)}.\end{aligned} \tag{10}$$

Therefore, we can conclude that $\lambda\partial\|\mathbf{\Gamma}_v^{k+1}\|_q^q$ is bounded. Then, the boundedness of sequence $\{\mathbf{M}_{v,2}^{k+1}\}_{v=1}^V$ can be confirmed due to $\lim_{k\to\infty}(\mathbf{C}_v^{k+1}-\mathbf{C}_v^k)=\mathbf{0}$.

Next, we prove the boundedness of $\{\mathbf{M}_{v,3}^{k+1}\}_{v=1}^V$. According to the update rule of $\mathbf{C}_v^{k+1}$, its optimality condition meets, i.e.,

$$\begin{aligned}\mathbf{0} &= \mu^k(\mathbf{C}_v^{k+1}-\mathbf{P}_v^{k+1}) + \mu^k(\mathbf{C}_v^{k+1}-\mathbf{S}_v^{k+1}) + \mathbf{M}_{v,1}^k \\ &\quad + \mathbf{M}_{v,2}^k + \mu^k\mathbf{X}_v^T(\mathbf{X}_v\mathbf{C}_v^{k+1}+\mathbf{E}_v^k-\mathbf{X}_v) - \mathbf{X}_v^T\mathbf{M}_{v,3}^k.\end{aligned} \tag{11}$$

Under the update rules of $\mathbf{M}_{v,1}^{k+1}$, $\mathbf{M}_{v,2}^{k+1}$ and $\mathbf{M}_{v,3}^{k+1}$, we have

$$\mathbf{X}_v^T\mathbf{M}_{v,3}^{k+1} = \mathbf{M}_{v,1}^{k+1} + \mathbf{M}_{v,2}^{k+1} + \mu^k\mathbf{X}_v^T(\mathbf{E}_v^k-\mathbf{E}_v^{k+1}). \tag{12}$$

By combining the boundedness of $\{\mathbf{M}_{v,1}^{k+1}\}_{v=1}^V$ and $\{\mathbf{M}_{v,2}^{k+1}\}_{v=1}^V$ with $\lim_{k\to\infty}(\mathbf{E}_v^{k+1}-\mathbf{E}_v^k)=\mathbf{0}$, we can easily verify that sequence $\{\mathbf{M}_{v,3}^{k+1}\}_{v=1}^V$ is bounded.

According to the definition of $\mathcal{L}_\mu(\cdot;\cdot)$ and the alternative scheme of each view in proposed Algorithm 1, we educe following two formulas

$$\begin{aligned}&\mathcal{L}_{\mu^k}(\mathbf{P}_v^{k+1},\mathbf{S}_v^{k+1},\mathbf{C}_v^{k+1},\mathbf{E}_v^{k+1},\mathbf{M}_{v,1}^k,\mathbf{M}_{v,2}^k,\mathbf{M}_{v,3}^k) \\ &\le \mathcal{L}_{\mu^k}(\mathbf{P}_v^{k},\mathbf{S}_v^{k},\mathbf{C}_v^{k},\mathbf{E}_v^{k},\mathbf{M}_{v,1}^k,\mathbf{M}_{v,2}^k,\mathbf{M}_{v,3}^k) \\ &= \mathcal{L}_{\mu^{k-1}}(\mathbf{P}_v^{k},\mathbf{S}_v^{k},\mathbf{C}_v^{k},\mathbf{E}_v^{k},\mathbf{M}_{v,1}^{k-1},\mathbf{M}_{v,2}^{k-1},\mathbf{M}_{v,3}^{k-1}) \\ &+ \left\langle\mathbf{M}_{v,1}^k-\mathbf{M}_{v,1}^{k-1},\mathbf{C}_v^k-\mathbf{P}_v^k\right\rangle \\ &+ \left\langle\mathbf{M}_{v,2}^k-\mathbf{M}_{v,2}^{k-1},\mathbf{C}_v^k-\mathbf{S}_v^k\right\rangle \\ &+ \left\langle\mathbf{M}_{v,3}^k-\mathbf{M}_{v,3}^{k-1},\mathbf{X}_v-\mathbf{X}_v\mathbf{C}_v^k-\mathbf{E}_v^k\right\rangle \\ &+ \frac{\mu^k-\mu^{k-1}}{2}(\|\mathbf{C}_v^k-\mathbf{P}_v^k\|_F^2+\|\mathbf{C}_v^k-\mathbf{S}_v^k\|_F^2 \\ &+ \|\mathbf{X}_v-\mathbf{X}_v\mathbf{C}_v^k-\mathbf{E}_v^k\|_F^2) \\ &= \mathcal{L}_{\mu^{k-1}}(\mathbf{P}_v^{k},\mathbf{S}_v^{k},\mathbf{C}_v^{k},\mathbf{E}_v^{k},\mathbf{M}_{v,1}^{k-1},\mathbf{M}_{v,2}^{k-1},\mathbf{M}_{v,3}^{k-1}) \\ &+ \frac{\mu^k+\mu^{k-1}}{2(\mu^{k-1})^2}(\|\mathbf{M}_{v,1}^k-\mathbf{M}_{v,1}^{k-1}\|_F^2+\|\mathbf{M}_{v,2}^k-\mathbf{M}_{v,2}^{k-1}\|_F^2 \\ &+ \|\mathbf{M}_{v,3}^k-\mathbf{M}_{v,3}^{k-1}\|_F^2)\end{aligned} \tag{13}$$

and

$$\begin{aligned}&\mathcal{L}_{\mu^k}(\mathbf{P}_v^{k+1},\mathbf{S}_v^{k+1},\mathbf{C}_v^{k+1},\mathbf{E}_v^{k+1},\mathbf{M}_{v,1}^k,\mathbf{M}_{v,2}^k,\mathbf{M}_{v,3}^k) \\ &+ \frac{1}{2\mu^k}(\|\mathbf{M}_{v,1}^k\|_F^2+\|\mathbf{M}_{v,2}^k\|_F^2+\|\mathbf{M}_{v,3}^k\|_F^2) \\ &= \|\mathbf{P}_v^{k+1}\|_q^q + \lambda\|\mathbf{W}\odot\mathbf{S}_v^{k+1}\|_q^q + \beta\|\mathbf{E}_v^{k+1}\|_{2,q} \\ &+ \frac{\mu^k}{2}\|\mathbf{C}_v^{k+1}-\mathbf{P}_v^{k+1}+\frac{\mathbf{M}_{v,1}^k}{\mu^k}\|_F^2 \\ &+ \frac{\mu^k}{2}\|\mathbf{C}_v^{k+1}-\mathbf{S}_v^{k+1}+\frac{\mathbf{M}_{v,2}^k}{\mu^k}\|_F^2 \\ &+ \frac{\mu^k}{2}\|\mathbf{X}_v-\mathbf{X}_v\mathbf{C}_v^{k+1}-\mathbf{E}_v^{k+1}+\frac{\mathbf{M}_{v,3}^k}{\mu^k}\|_F^2\end{aligned} \tag{14}$$

Then, the summation of the right side in (13) can be derived from $k=1$ to $K(\ge 1)$, such that

$$\begin{aligned}&\mathcal{L}_{\mu^K}(\mathbf{P}_v^{K+1},\mathbf{S}_v^{K+1},\mathbf{C}_v^{K+1},\mathbf{E}_v^{K+1},\mathbf{M}_{v,1}^K,\mathbf{M}_{v,2}^K,\mathbf{M}_{v,3}^K) \\ &\le \mathcal{L}_{\mu^0}(\mathbf{P}_v^1,\mathbf{S}_v^1,\mathbf{C}_v^1,\mathbf{E}_v^1,\mathbf{M}_{v,1}^0,\mathbf{M}_{v,2}^0,\mathbf{M}_{v,3}^0) \\ &+ \sum_{k=1}^{K}\frac{\mu^k+\mu^{k-1}}{2(\mu^{k-1})^2}(\|\mathbf{M}_{v,1}^k-\mathbf{M}_{v,1}^{k-1}\|_F^2 \\ &+ \|\mathbf{M}_{v,2}^k-\mathbf{M}_{v,2}^{k-1}\|_F^2+\|\mathbf{M}_{v,3}^k-\mathbf{M}_{v,3}^{k-1}\|_F^2).\end{aligned} \tag{15}$$

By the update rule $\mu^k=\rho\mu^{k-1}$ and $\mu^k=\rho^k\mu^0$, we can get

$$\begin{aligned}\sum_{k=1}^{K}\frac{\mu^k+\mu^{k-1}}{2(\mu^{k-1})^2} &= \frac{\rho(\rho+1)}{2\mu_0(\rho-1)}(1-\frac{1}{\rho^K}) \\ &< \frac{\rho(\rho+1)}{2\mu_0(\rho-1)} < +\infty,\end{aligned} \tag{16}$$

where the first inequality holds due to $0<(1-1/\rho^K)<1$ by $\rho>1$. Thus $\sum_{k=1}^K(\mu^k+\mu^{k-1})/(2(\mu^{k-1})^2)$ is bounded. Taking it with the bounded sequences $\{\mathbf{M}_{v,1}^{k+1}\}_{v=1}^V$, $\{\mathbf{M}_{v,2}^{k+1}\}_{v=1}^V$ and $\{\mathbf{M}_{v,3}^{k+1}\}_{v=1}^V$, we can further obtain that $\mathcal{L}_{\mu^k}(\mathbf{P}_v^{k+1},\mathbf{S}_v^{k+1}$ $,\mathbf{C}_v^{k+1},\mathbf{E}_v^{k+1},\mathbf{M}_{v,1}^k,\mathbf{M}_{v,2}^k,\mathbf{M}_{v,3}^k)$ is bounded by the finite $\mathcal{L}_{\mu^0}(\mathbf{P}_v^1,\mathbf{S}_v^1,\mathbf{C}_v^1,\mathbf{E}_v^1,\mathbf{M}_{v,1}^0,\mathbf{M}_{v,2}^0,\mathbf{M}_{v,3}^0)$. On the basis of this, we can conclude that each term in both sides of (14) is bounded. Then, the boundedness of $\|\mathbf{P}_v^{k+1}\|_q^q$, $\|\mathbf{W}\odot\mathbf{S}_v^{k+1}\|_q^q$ and $\|\mathbf{E}_v^{k+1}\|_{2,q}$ are validated.

Let the SVD of $\mathbf{P}_v^{k+1}$ be $\mathbf{U\Sigma V}^T$, we can have $\|\mathbf{P}_v^{k+1}\|_q^q = \sum_{i=1}^n\mathbf{\Sigma}_{i,i}^q$. Due to its boundedness, there exists a positive constant γ, such that

$$0\le\mathbf{\Sigma}_{i,i}^q\le\gamma \Longrightarrow \mathbf{\Sigma}_{i,i}\le\gamma^{1/q}, \tag{17}$$

and further obtain

$$\|\mathbf{P}_v^{k+1}\|_F^2 = \sum_{i=1}^n\mathbf{\Sigma}_{i,i}^2 \le n\gamma^{2/q}. \tag{18}$$

Thus the sequence $\{\mathbf{P}_v^{k+1}\}_{v=1}^V$ is bounded. Taking the boundedness of $\{\mathbf{P}_v^{k+1}\}_{v=1}^V$ and $\{\mathbf{M}_{v,1}^{k+1}\}_{v=1}^V$ into account, the boundedness of $\{\mathbf{C}_v^{k+1}\}_{v=1}^V$ can be easily obtained by the forth term in right side of (14).

Similarly, we suppose $\mathbf{\Gamma}_v^{k+1} = \mathbf{W} \odot \mathbf{S}_v^{k+1}$, then for $\|\mathbf{\Gamma}_v^{k+1}\|_q^q = \sum_{i,j} \left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right|^q$, there also exists a positive scalar ζ meet

$$0 \leq \left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right|^q \leq \zeta \Longrightarrow \left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right| \leq \zeta^{1/q}, \tag{19}$$

then make

$$\|\mathbf{\Gamma}_v^{k+1}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left|\left[\mathbf{\Gamma}_v^{k+1}\right]_{i,j}\right|^2 \leq n^2 \zeta^{2/q}. \tag{20}$$

Hence the boundedness of $\{\mathbf{S}_v^{k+1}\}_{v=1}^V$ is verified due to matrix $\mathbf{W}$ with $\mathbf{W}_{i,j} \in [0,1]$ is predefined and delimited.

Given the definition of $\|\mathbf{E}_v^{k+1}\|_{2,q}$ as

$$\|\mathbf{E}_v^{k+1}\|_{2,q} = \left[\sum_{i=1}^{d_v}\left(\sum_{j=1}^{n}\left|\left[\mathbf{E}_v^{k+1}\right]_{i,j}\right|^2\right)^{q/2}\right]^{1/q}, \tag{21}$$

where d_v denotes the dimension of each feature view. Due to the boundedness of $\|\mathbf{E}_v^{k+1}\|_{2,q}$ with $0 < q < 1$, there satisfies

$$0 \leq \left|\left[\mathbf{E}_v^{k+1}\right]_{i,j}\right|^2 \leq \psi \Longrightarrow \left|\left[\mathbf{E}_v^{k+1}\right]_{i,j}\right| \leq \psi^{1/2}, \tag{22}$$

where ψ is a positive constant. Then we have

$$\|\mathbf{E}_v^{k+1}\|_F^2 = \sum_{i=1}^{d_v}\sum_{j=1}^{n}\left|\left[\mathbf{E}_v^{k+1}\right]_{i,j}\right|^2 \leq d_v n \psi. \tag{23}$$

Thus $\{\mathbf{E}_v^{k+1}\}_{v=1}^V$ is bounded.

Finally, it can be concluded that the sequence $\{\Theta^k = (\mathbf{C}_v^k,\ \mathbf{P}_v^k,\ \mathbf{S}_v^k,\ \mathbf{E}_v^k;\ \mathbf{M}_{v,1}^k,\ \mathbf{M}_{v,2}^k,\ \mathbf{M}_{v,3}^k) | v = 1, \ldots, V\}_{k=1}^{\infty}$ is bounded.

2) In the light of Bolzano-Weierstrass theorem [Bartle and Sherbert, 2000], there must exists at least one convergent subsequence in bounded sequence $\{\Theta^k\}_{k=1}^{\infty}$. Suppose the subsequence is represented by $\{\Theta^k\}_{k=1}^{\infty}$ itself. Without losing generality, it converges to an accumulation point Θ^*. Then we can get

$$\begin{aligned}&\lim_{k\to\infty}(\mathbf{C}_v^k,\ \mathbf{P}_v^k,\ \mathbf{S}_v^k,\ \mathbf{E}_v^k;\ \mathbf{M}_{v,1}^k,\ \mathbf{M}_{v,2}^k,\ \mathbf{M}_{v,3}^k)\\ &\quad = (\mathbf{C}_v^*,\ \mathbf{P}_v^*,\ \mathbf{S}_v^*,\ \mathbf{E}_v^*;\ \mathbf{M}_{v,1}^*,\ \mathbf{M}_{v,2}^*,\ \mathbf{M}_{v,3}^*)\end{aligned} \tag{24}$$

According to the update rules of dual multipliers, we obtain

$$\begin{cases}\dfrac{\mathbf{M}_{v,1}^{k+1} - \mathbf{M}_{v,1}^k}{\mu^k} = \mathbf{C}_v^{k+1} - \mathbf{P}_v^{k+1}, & (25)\\ \dfrac{\mathbf{M}_{v,2}^{k+1} - \mathbf{M}_{v,2}^k}{\mu^k} = \mathbf{C}_v^{k+1} - \mathbf{S}_v^{k+1}, & (26)\\ \dfrac{\mathbf{M}_{v,3}^{k+1} - \mathbf{M}_{v,3}^k}{\mu^k} = \mathbf{X}_v - \mathbf{X}_v\mathbf{C}_v^{k+1} - \mathbf{E}_v^{k+1}. & (27)\end{cases}$$

With the fact $\lim_{k\to\infty}\mu^k = +\infty$ and the boundedness of $\{\mathbf{M}_{v,1}^{k+1}\}_{v=1}^V$, $\{\mathbf{M}_{v,2}^{k+1}\}_{v=1}^V$ and $\{\mathbf{M}_{v,3}^{k+1}\}_{v=1}^V$, we have

$$\lim_{k\to\infty}\frac{\mathbf{M}_{v,1}^{k+1} - \mathbf{M}_{v,1}^k}{\mu^k} = \lim_{k\to\infty}\mathbf{C}_v^{k+1} - \mathbf{P}_v^{k+1} = \mathbf{0}, \tag{28}$$

$$\lim_{k\to\infty}\frac{\mathbf{M}_{v,2}^{k+1} - \mathbf{M}_{v,2}^k}{\mu^k} = \lim_{k\to\infty}\mathbf{C}_v^{k+1} - \mathbf{S}_v^{k+1} = \mathbf{0}, \tag{29}$$

$$\lim_{k\to\infty}\frac{\mathbf{M}_{v,3}^{k+1} - \mathbf{M}_{v,3}^k}{\mu^k} = \lim_{k\to\infty}\mathbf{X}_v - \mathbf{X}_v\mathbf{C}_v^{k+1} - \mathbf{E}_v^{k+1} = \mathbf{0}. \tag{30}$$

Therefore, the Primal Feasibility is satisfied by (24), i.e.,

$$\begin{cases}\mathbf{C}_v^* - \mathbf{P}_v^* = \mathbf{0}, & (31)\\ \mathbf{C}_v^* - \mathbf{S}_v^* = \mathbf{0}, & (32)\\ \mathbf{X}_v - \mathbf{X}_v\mathbf{C}_v^* - \mathbf{E}_v^* = \mathbf{0}. & (33)\end{cases}$$

Additionally, the optimality conditions of variables $\mathbf{P}_v^{k+1}$ and $\mathbf{S}_v^{k+1}$ hold, i.e., (6) and (9), which indicate

$$\mathbf{0} \in \partial\|\mathbf{P}_v^{k+1}\|_q^q - \mathbf{M}_{v,1}^{k+1} \overset{k\to\infty}{\Longrightarrow} \mathbf{M}_{v,1}^* \in \partial\|\mathbf{P}_v^*\|_q^q \tag{34}$$

and

$$\begin{aligned}&\mathbf{0} \in \lambda\partial\|\mathbf{W}\odot\mathbf{S}_v^{k+1}\|_q^q - \mathbf{M}_{v,2}^{k+1}\\ &\overset{k\to\infty}{\Longrightarrow} \mathbf{M}_{v,2}^* \in \lambda\partial\|\mathbf{W}\odot\mathbf{S}_v^*\|_q^q\end{aligned} \tag{35}$$

by (24) and $\lim_{k\to\infty}(\mathbf{C}_v^{k+1} - \mathbf{C}_v^k) = \mathbf{0}$. For variable $\mathbf{E}_v^{k+1}$, its optimality condition is satisfied by the update rules of $\mathbf{E}_v^{k+1}$ and $\mathbf{M}_{v,3}^{k+1}$. With (24), it implies that

$$\mathbf{0} \in \beta\partial\|\mathbf{E}_v^{k+1}\|_{2,q} - \mathbf{M}_{v,3}^{k+1} \overset{k\to\infty}{\Longrightarrow} \mathbf{M}_{v,3}^* \in \beta\partial\|\mathbf{E}_v^*\|_{2,q}. \tag{36}$$

Therefore, the Stationary Conditions are satisfied by (34), (35) and (36). In summary, it can be concluded that any converged point Θ^* of $\{\Theta^k\}_{k=1}^{\infty}$ produced by Algorithm 1 satisfies KKT conditions with stationary point $\{(\mathbf{C}_v^*,\ \mathbf{P}_v^*,\ \mathbf{S}_v^*,\ \mathbf{E}_v^*) | v = 1, \ldots, V\}$. □

In what follows, we will consider the convergence rate of our proposed method.

Theorem 2 (Rate of Convergence.)**.** *Let* $\{\Theta^k = (\mathbf{C}_v^k,\ \mathbf{P}_v^k,\ \mathbf{S}_v^k,\ \mathbf{E}_v^k;\ \mathbf{M}_{v,1}^k,\ \mathbf{M}_{v,2}^k,\ \mathbf{M}_{v,3}^k) |\ v = 1, \ldots, V\}_{k\in N}$ *be the sequence generated by Algorithm 1, which converges to* $\Theta^* = (\mathbf{C}_v^*,\ \mathbf{P}_v^*,\ \mathbf{S}_v^*,\ \mathbf{E}_v^*;\ \mathbf{M}_{v,1}^*,\ \mathbf{M}_{v,2}^*,\ \mathbf{M}_{v,3}^*)$. *Suppose that* $\mathcal{L}_\mu(\cdot)$ *has the KŁ property at* Θ^* *with* $\psi(\gamma) = c\gamma^{1-\beta}$, *where* $\beta \in [0, 1)$ *and the constant* $c > 0$. *Then we have the following estimations:*

1) If $\beta = 0$, *the sequence* $\{\Theta^k\}_{k\in N}$ *converges in a finite number of iterations.*

2) If $\beta \in (0, 1/2]$, *there exist* $\upsilon \in [0, 1)$ *and* $c > 0$ *such that* $\|\Theta^k - \Theta^*\|_F \leq c\upsilon^k$.

3) If $\beta \in (1/2, 1)$, *there exists* $c > 0$ *such that* $\|\Theta^k - \Theta^*\|_F \leq ck^{(\beta-1)/(2\beta-1)}$.

Proof. These derivations are quite standard [Liu *et al.*, 2020], hence we omit details of this proof in our Supplemental Materials. □

References

[Bartle and Sherbert, 2000] Robert G Bartle and Donald R Sherbert. *Introduction to real analysis*, volume 2. Wiley New York, 2000.

[Lewis and Sendov, 2005] Adrian S. Lewis and Hristo S. Sendov. Nonsmooth analysis of singular values. part i: Theory. *Set-Valued Analysis*, 13(3):213–241, 2005.

[Liu *et al.*, 2020] Risheng Liu, Shichao Cheng, Yi He, Xin Fan, Zhouchen Lin, and Zhongxuan Luo. On the convergence of learing-based iterative methods for nonconvex inverse problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12):3027–3039, 2020.

[Marjanovic and Solo, 2012] Goran Marjanovic and Victor Solo. On l_q optimization and matrix completion. *IEEE Transactions on Signal Processing*, 60(11):5714–5724, 2012.