Appendix

A Proof of Theorem 3.1

Proof. Let $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} *_{\mathcal{L}} \boldsymbol{\mathcal{S}} *_{\mathcal{L}} \boldsymbol{\mathcal{V}}^{\mathrm{T}}$ be the t-SVD of $\boldsymbol{\mathcal{X}}$. By the orthogonality of $\boldsymbol{L}$ and the definition of tensor average rank, the problem (20) is equivalent to

$$\begin{aligned}
\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}}) \quad &\in \quad \arg\min_{\boldsymbol{\mathcal{Y}}} \frac{1}{2\lambda}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}}\|_{\mathrm{F}}^2 + \mathrm{rank}_{\mathrm{a}}(\boldsymbol{\mathcal{Y}}) && \text{(A.1)}\\
&\in \quad \arg\min_{\boldsymbol{\mathcal{Y}}} \frac{1}{2\lambda}\|\bar{\boldsymbol{\mathcal{Y}}} - \bar{\boldsymbol{\mathcal{X}}}\|_{\mathrm{F}}^2 + \mathrm{rank}_{\mathrm{a}}(\bar{\boldsymbol{\mathcal{Y}}}) && \text{(A.2)}\\
&= \quad \arg\min_{\boldsymbol{\mathcal{Y}}} \sum_{i=1}^{n_3}\Big(\frac{1}{2}\|\bar{\boldsymbol{X}}^{(i)} - \bar{\boldsymbol{Y}}^{(i)}\|_{\mathrm{F}}^2 + \lambda \mathrm{rank}(\bar{\boldsymbol{Y}}^{(i)})\Big), && \text{(A.3)}
\end{aligned}$$

which is separable to each frontal slice in the transform domain. By the Corollary 2.2 in [1], we get that the i-th frontal slice of $\overline{\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}})}$ solves the i-th subproblem of (A.3) . Hence, $\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}})$ solves problem (20). $\square$

B Proof of Proposition 3.1

Proof. Let $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} *_{\mathcal{L}} \boldsymbol{\mathcal{S}} *_{\mathcal{L}} \boldsymbol{\mathcal{V}}^{\mathrm{T}}$ be the t-SVD of $\boldsymbol{\mathcal{X}}$. By Theorem 3.1, the right side of (21) equals to

$$\frac{1}{2\lambda}\|\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}}) - \boldsymbol{\mathcal{X}}\|_{\mathrm{F}}^2 + \mathrm{rank}_{\mathrm{a}}(\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}})). \tag{B.1}$$

Substituting the explicit form of $\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}})$ in (18) into (B.1) and by the orthogonality of $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$, (B.1) can be simplified as

$$\frac{1}{2\lambda}\sum_{(i,k)\notin\Lambda} \bar{\boldsymbol{\mathcal{S}}}_{iik}^2 + \sharp\Lambda, \tag{B.2}$$

where Λ is the set defined by

$$\Lambda = \{(i,k)|\bar{\boldsymbol{\mathcal{S}}}_{iik} > \sqrt{2\lambda}\},$$

and $\sharp\Lambda$ denotes its cardinality. We can prove (21) by directly verifying the definition of $\Phi_{\lambda}(\boldsymbol{\mathcal{X}})$ in (15). Moreover, if $\lambda \leqslant \frac{\min_{i,k} \bar{\boldsymbol{\mathcal{S}}}_{iik}^2}{2}$, we get that $\boldsymbol{\mathcal{D}}_{\lambda}(\boldsymbol{\mathcal{X}}) = \boldsymbol{\mathcal{X}}$. Hence, $\Phi_{\lambda}(\boldsymbol{\mathcal{X}}) = \mathrm{rank}_{\mathrm{a}}(\boldsymbol{\mathcal{X}})$. $\square$

The following notation is needed for presenting the convergence analysis. For a tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$, let $\sigma(\bar{\boldsymbol{X}}^{(i)})$ denote the singular value vector of the i-th frontal slice of $\bar{\boldsymbol{\mathcal{X}}} = L(\boldsymbol{\mathcal{X}})$. We denote by $\sigma_j(\bar{\boldsymbol{X}}^{(i)})$ the j-th component of $\sigma(\bar{\boldsymbol{X}}^{(i)})$ with $1 \leqslant j \leqslant \min(n_1, n_2)$.

Lemma B.1. *Let $\{\boldsymbol{\mathcal{Y}}_k\}$ be generated by (23), then*

(1) Either $\sigma_j(\bar{\boldsymbol{Y}}_k^{(i)}) \geqslant \sqrt{2\lambda\mu}$ or $\sigma_j(\bar{\boldsymbol{Y}}_k^{(i)}) = 0$, for all $1 \leqslant i \leqslant n_3, 1 \leqslant j \leqslant \min(n_1, n_2)$,

(2) $\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}} \geqslant \sqrt{2\lambda\mu}$, if $\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}) \neq \mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_k)$.

Proof. Item (1) follows immediately from the closed-form solution of $\{\boldsymbol{\mathcal{Y}}_k\}$ in (23). Suppose that $\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}) \neq \mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_k)$ for some k, by Item (1), there exist at least one i and one j such that either $\sigma_j(\bar{\boldsymbol{Y}}_{k+1}^{(i)}) \geqslant \sqrt{2\lambda\mu}$, $\sigma_j(\bar{\boldsymbol{Y}}_k^{(i)}) = 0$ or $\sigma_j(\bar{\boldsymbol{Y}}_{k+1}^{(i)}) = 0$, $\sigma_j(\bar{\boldsymbol{Y}}_k^{(i)}) \geqslant \sqrt{2\lambda\mu}$. For both cases, we have that

$$\|\sigma(\bar{\boldsymbol{Y}}_{k+1}^{(i)}) - \sigma(\bar{\boldsymbol{Y}}_k^{(i)})\|_2 \geqslant \sqrt{2\lambda\mu}.$$

Moreover, for any $1 \leqslant i \leqslant n_3$, we know from Corollary 7.1.4.3 of [2] that

$$\|\bar{\boldsymbol{Y}}_{k+1}^{(i)} - \bar{\boldsymbol{Y}}_k^{(i)}\|_{\mathrm{F}} \geqslant \|\sigma(\bar{\boldsymbol{Y}}_{k+1}^{(i)}) - \sigma(\bar{\boldsymbol{Y}}_k^{(i)})\|_2.$$

The above two inequalities and the orthogonality of $\boldsymbol{L}$ yield Item (2). ☐

C Proof of Theorem 3.2

Proof. Denote

$$E(\boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{X}}) = \frac{1}{2\lambda}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}}\|_{\mathrm{F}}^2 + \mathrm{rank}_a(\boldsymbol{\mathcal{Y}}).$$

By the second subproblem in (23), we conclude that

$$\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{X}}_{k+1}\|_{\mathrm{F}}^2 \leqslant \|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2.$$

Then we have that

$$\begin{aligned} \lambda E(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{X}}_{k+1}) &= \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{X}}_{k+1}\|_{\mathrm{F}}^2 + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}) \\ &\leqslant \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2 + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}). \end{aligned} \tag{C.1}$$

Expanding the term $\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2$ yields

$$\lambda E(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{X}}_{k+1}) \leqslant \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}}^2 + \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2 + \langle \boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\rangle + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}). \tag{C.2}$$

Hence, for any $\mu \in (0, 1)$, we obtain that

$$\lambda E(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{X}}_{k+1}) \leqslant \frac{1}{2\mu}\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}}^2 + \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2 + \langle \boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\rangle + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_{k+1}). \tag{C.3}$$

The first line in (23) can be rewritten as

$$\boldsymbol{\mathcal{Y}}_{k+1} \in \arg\min_{\boldsymbol{\mathcal{Y}}}\{\frac{1}{2\mu}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{Y}}_k + \mu(\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k)\|_{\mathrm{F}}^2 + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}})\}. \tag{C.4}$$

Expanding the quadratic term in (C.4) and replacing $\frac{\mu}{2}\|\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2$ as $\frac{1}{2}\|\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2$ (will not alter the minimizer) yield that

$$\boldsymbol{\mathcal{Y}}_{k+1} \in \arg\min_{\boldsymbol{\mathcal{Y}}} \frac{1}{2\mu}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}}^2 + \frac{1}{2}\|\boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}}^2 + \langle \boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{Y}}_k - \boldsymbol{\mathcal{X}}_k\rangle + \lambda\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}). \tag{C.5}$$

Substituting $\boldsymbol{\mathcal{Y}}_{k+1}$ and $\boldsymbol{\mathcal{Y}}_k$ into the objective function of (C.5), and by (C.3), we conclude that

$$E(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{X}}_{k+1}) \leqslant E(\boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{X}}_k). \tag{C.6}$$

Therefore, the sequence $\{E(\boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{X}}_k)\}$ is monotonically decreasing and hence converges. Moreover, by (C.2), (C.3) and (C.6), we get that

$$\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}}^2 \leqslant \frac{2\mu\lambda}{1-\mu}(E(\boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{X}}_k) - E(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{X}}_{k+1})). \tag{C.7}$$

This implies that $\lim_{k\to\infty} \|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}} = 0$. By the second subproblem of (23), $\boldsymbol{\mathcal{X}}_{k+1}$ is essentially the projection of $\boldsymbol{\mathcal{Y}}_{k+1}$ on the feasible set

$$\mathcal{C} = \{\boldsymbol{\mathcal{X}} \mid \boldsymbol{P}_{\boldsymbol{\Omega}}(\boldsymbol{\mathcal{X}}) = \boldsymbol{P}_{\boldsymbol{\Omega}}(\boldsymbol{\mathcal{M}})\}.$$

The nonexpansiveness of the projection operator implies $\lim_{k\to\infty} \|\boldsymbol{\mathcal{X}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_{\mathrm{F}} = 0$.

The result of $\lim_{k\to\infty} \|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}} = 0$ implies that there exists a number $K > 0$ such that $\|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}} < \sqrt{2\lambda\mu}$ for all $k > K$. By Item (2) of Lemma B.1, the $\mathrm{rank}_a(\boldsymbol{\mathcal{Y}}_k)$ keeps invariant for all $k > K$. □

C.1 Note

If we choose $(\boldsymbol{\mathcal{Y}}_0, \boldsymbol{\mathcal{X}}_0)$ be a feasible point of (22) and bounded, by (C.7) we have that

$$\sum_{k=1}^{+\infty} \|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}}^2 < +\infty,$$

which implies that $\lim_{k\to\infty} \|\boldsymbol{\mathcal{Y}}_{k+1} - \boldsymbol{\mathcal{Y}}_k\|_{\mathrm{F}} = 0$ and the sequence $\{\boldsymbol{\mathcal{Y}}_k\}$ is bounded. Moreover, the sequence $\{\boldsymbol{\mathcal{X}}_k\}$ is also bounded.

D Image Recovery

We show the detailed PSNR values of the standard images with different sampling rates in this appendix.

Figure 1: Standard images: airplane, baboon, barbara, boats, butterfly, house, lena, peppers (from left to right).

Table 1: Comparison of PSNR on the standard images at sampling rate $p = 0.2$.

Method	airplane	baboon	barbara	boats	butterfly	house	lena	peppers	average
HaLRTC	21.27	19.12	21.02	21.06	16.94	24.29	21.76	20.51	20.75
T-SVD	21.51	18.58	20.90	21.66	17.27	24.43	21.64	19.47	20.68
TNN	23.21	19.68	21.53	22.90	19.69	25.70	22.42	21.03	22.02
TNN-DCT	23.23	19.84	21.58	23.00	19.76	26.24	22.54	21.19	22.17
Ours	**23.99**	**20.13**	**22.33**	**23.72**	**20.95**	**27.36**	**23.51**	**22.06**	**23.01**

Table 2: Comparison of PSNR on the standard images at sampling rate $p = 0.4$.

Method	airplane	baboon	barbara	boats	butterfly	house	lena	peppers	average
HaLRTC	25.64	21.91	25.30	25.24	22.15	26.69	26.11	25.26	24.78
T-SVD	26.04	21.84	25.65	26.71	23.27	30.04	26.34	24.62	25.56
TNN	27.61	23.22	26.00	27.46	25.68	30.72	27.28	25.25	26.65
TNN-DCT	27.70	23.27	26.06	27.57	25.90	31.30	27.43	25.45	26.83
Ours	**28.89**	**23.82**	**27.20**	**28.95**	**28.28**	**32.68**	**29.08**	**26.72**	**28.20**

Table 3: Comparison of PSNR on the standard images at sampling rate $p = 0.6$.

Method	airplane	baboon	barbara	boats	butterfly	house	lena	peppers	average
HaLRTC	30.05	25.03	29.40	29.75	27.50	34.23	30.36	29.85	29.52
T-SVD	30.90	25.39	30.40	32.00	30.45	34.59	30.95	29.15	30.48
TNN	31.88	26.94	30.33	32.43	32.16	34.81	31.88	29.41	31.23
TNN-DCT	31.94	26.93	30.42	32.61	32.72	35.47	32.07	29.66	31.48
Ours	**32.72**	**27.87**	**31.83**	**34.84**	**35.91**	**36.71**	**34.17**	**31.01**	**33.13**

Table 4: Comparison of PSNR on the standard images at sampling rate $p = 0.8$.

Method	airplane	baboon	barbara	boats	butterfly	house	lena	peppers	average
HaLRTC	36.06	29.61	34.70	35.95	34.65	39.21	36.00	35.43	35.20
T-SVD	37.12	30.41	36.25	38.69	40.40	39.75	37.00	34.99	36.83
TNN	36.77	32.06	36.09	38.77	39.95	39.72	37.69	34.37	36.92
TNN-DCT	36.76	31.97	36.15	39.01	41.16	40.32	37.87	34.67	37.24
Ours	**37.21**	**33.30**	**38.51**	**41.50**	**44.33**	**41.24**	**40.23**	**36.38**	**39.09**

References

[1] Y. Zhang and Z. Lu. Penalty decomposition methods for rank minimization. In *Advances in Neural Information Processing Systems*, pages 46–54. 2011.

[2] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2012.