A. Supplementary materials

A.1. Notation and Lemmas

We provide some additional notations for the proof. First, for $X \in \mathcal{X}$, we denote an overlapped Schatten sup-norm as

$$|||X|||_{s,\infty} := \max_k \max_r \sigma_r(X_{(k)}).$$

With the norm, we introduce a following Lemma.

Lemma 7. *(Lemma 1 in Tomioka et al. (2011)) For $X \in \mathcal{X}$, consider the infimum of the maximum mode-k spectral norm for $|||\cdot|||_{s^*}$ as*

$$|||X|||_{s^*} = \inf_{\frac{1}{K}\sum_{k=1}^K Y^{(k)}=X} \max_k \|Y^{(k)}_{(k)}\|_{s,\infty},$$

where $Y^{(k)} \in \mathcal{X}$ and $Y^{(k)}_{(k)}$ is the mode-k unfolding of $Y^{(k)}$ for all k. Then, $|||\cdot|||_{s^*}$ is the dual norm of the overlapped Schatten 1-norm $|||\cdot|||_s$. Moreover, the following inequality is valid:

$$|||X|||_{s^*} \leq |||X|||_m.$$

Proof is provided in Tomioka et al. (2011). By Lemma 7, we obtain the following Holder-type inequality as

$$|\langle X, X' \rangle| \leq |||X|||_s |||X'|||_{s^*} \leq |||X|||_s |||X'|||_m. \tag{13}$$

We also discuss a rank restriction for tensor and provide Lemma for the restriction.

We introduce another result to bound the the effect of the noise tensor such as $|||\mathfrak{X}^*(\mathcal{E})|||_m$.

Lemma 8. *(Lemma 3 in Tomioka et al. (2011)) Let $\mathfrak{X}$ and $\mathcal{E}$ be as the defined above. Then, with high probability, we have*

$$|||\mathfrak{X}^*(\mathcal{E})|||_m \leq \frac{\sigma}{K}\sum_{k=1}^K \left(\sqrt{I_k} + \sqrt{I_{\backslash k}}\right).$$

A.2. Proof of Lemma 2

The notation $\lesssim$ denotes that the left-hand side is bounded by the right-hand side up to a constant. Also, by the setting of the basis functions, we define a finite positive constant C_P satisfying $C_P \geq |||W|||_F / |||X|||_F$.

By the definition of $\hat{X}$, we obtain the following basis inequality:

$$\begin{aligned}&\frac{1}{2n}\|Y - \mathfrak{X}(\hat{X})\|^2 + \lambda_n |||W_{\hat{X}}|||_s + \mu_n |||W_{\hat{X}}|||_F \\ &\leq \frac{1}{2n}\|Y - \mathfrak{X}(X^*)\|^2 + \lambda_n |||W_{X^*}|||_s + \mu_n |||W_{X^*}|||_F.\end{aligned}$$

Let $\Delta_X := X^* - \hat{X}$ and $\Delta_W := W_{X^*} - W_{\hat{X}}$, and some calculation yields

$$\begin{aligned}\frac{1}{2n}\|\mathfrak{X}(\Delta_X)\|^2 \leq& \frac{1}{n}\langle \mathcal{E}, \mathfrak{X}(\Delta_X)\rangle \\ &+ \lambda_n(|||W_{\hat{X}} + \Delta_W|||_s - |||W_{\hat{X}}|||_s) \\ &+ \mu_n(|||W_{\hat{X}} + \Delta_W|||_F^2 - |||W_{\hat{X}}|||_F^2).\end{aligned} \tag{14}$$

Remind that $\mathfrak{X}$ satisfies the linearity properties.

Here, we evaluate each of the terms on the right-hand side of (14). About the first term, we obtain

$$\frac{1}{n}\langle \mathcal{E}, \mathfrak{X}(\Delta_X)\rangle = \frac{1}{n}\langle \mathfrak{X}^*(\mathcal{E}), \Delta\rangle$$

$$\leq \frac{1}{n}|||\mathfrak{X}^*(\mathcal{E})|||_{s^*}|||\Delta_X|||_s \leq \frac{1}{n}|||\mathfrak{X}^*(\mathcal{E})|||_m|||\Delta_X|||_s.$$

by the definition of adjoint operators, the Holder's inequality, and Lemma 7. Let $\lambda^* = \frac{1}{n}|||\mathfrak{X}^*(\mathcal{E})|||_m$ for brevity. Here, we discuss the relation between $\Delta_{X,(k)}$ and $\Delta_{W,(k)}$.

Also, using the setting on Φ and the Holder's inequality, we have

$$\begin{aligned}|||\Delta_X|||_s &= \frac{1}{K}\sum_{k=1}^K \|\Delta_{X,(k)}\|_s && (15)\\ &\leq \frac{1}{K}\sum_{k=1}^K \|\Gamma_{(k)}\|\|\Delta_{W,(k)}\| \\ &\leq |||\Delta_W|||_F \frac{1}{K}\sum_{k=1}^K \|\Gamma_{(k)}\| \\ &=: |||\Delta_W|||_F C_\Gamma. && (16)\end{aligned}$$

About the third term in (14), since we have we have

$$\lambda_n(|||W_{\hat{X}} + \Delta_X|||_s - |||W_{\hat{X}}|||_s) \leq \lambda_n|||\Delta_W|||_s,$$

by the triangle inequality. The third term in (14) is bounded as

$$\begin{aligned}&\mu_n(|||W_{X^*}|||_F^2 - |||W_{\hat{X}}|||_F^2)\\ &= \frac{\mu_n}{K}\left(\sum_{k=1}^K\sum_{m_k=1}^\infty \left(w^*_{m_1\dots m_K}\right)^2 - \sum_{k=1}^K\sum_{m_k=1}^{M^{(k)}} \hat{w}^2_{m_1\dots m_K}\right)\\ &= \frac{\mu_n}{K}\left(\sum_{k=1}^K\sum_{m=1}^{M^{(k)}}\left\{\left(w^*_{m_1\dots m_K}\right)^2 - \hat{w}^2_{m_1\dots m_K}\right)\right\} + \sum_{k=1}^K\sum_{m_k>M^{(k)}}\left(w^*_{m_1\dots m_K}\right)^2\right).\end{aligned}$$

Here, we let $A := \frac{1}{K}\sum_{k=1}^K\sum_{m_k>M^{(k)}}\left(w^*_{m_1\dots m_K}\right)^2$ for brevity. To bound the term $\sum_{k=1}^K\sum_{m=1}^{M^{(k)}}\left\{\left(w^*_{m_1\dots m_K}\right)^2 - \hat{w}^2_{m_1\dots m_K}\right)\right\}$, we define a projection $\Pi : \mathbb{R}^{M^{(1)}\times\cdots\times M^{(K)}} \to \Theta$ is smooth and low-rank, and let $\overline{\Delta}_W := \Pi(W_{X^*}) - W_{\hat{X}}$. Then, we have

$$\begin{aligned}&\frac{1}{K}\sum_{k=1}^K\sum_{m=1}^{M^{(k)}}\left\{\left(w^*_{m_1\dots m_K}\right)^2 - \hat{w}^2_{m_1\dots m_K}\right)\right\}\\ &= |||\Pi(W_{X^*})|||^2 - |||W_{\hat{X}}|||^2\\ &= |||W^2_{\hat{X}} + \overline{\Delta}_W|||^2 - |||W_{\hat{X}}|||_F^2 \leq |||\overline{\Delta}_W|||_F^2.\end{aligned}$$

Then we have

$$\mu_n(|||W_{X^*}|||_F^2 - |||W_{\hat{X}}|||_F^2) \leq \mu_n(|||\overline{\Delta}_W|||_F^2 + A).$$

About the second term in (14), we obtain

$$\begin{aligned}&|||\Delta_W|||_s\\ &\leq \frac{1}{K}\sum_{k=1}^K\|\Delta'_{W,(k)}\|_s + \frac{1}{K}\sum_{k=1}^K\|\Delta''_{W,(k)}\|_s\\ &\leq 4\frac{1}{K}\sum_{k=1}^K\|\Delta'_{W,(k)}\|_s + \frac{1}{K}\sum_{k=1}^K\sum_{r_k>R_k^W}\sigma_{r_k}(W^*_{(k)}),\end{aligned}$$

where the inequalities follow the same discussion with Lemma 2 in Tomioka et al. (2011) and Lemma 1 in Negahban & Wainwright (2011). Let $B := \frac{1}{K}\sum_{k=1}^{K}\sum_{r_k > R_k^W}\sigma_{r_k}(W_{X^*,(k)})$ for brevity. Then, by the Holder-type inequality, we have following inequalities as

$$\begin{aligned}\|\|\Delta'_W\|\|_s &\leq \frac{1}{K}\sum_{k=1}^{K}\sqrt{2R_k^W}\|\Delta_{W,(k)}\|_F \\ &\leq \|\|\Delta_W\|\|_F \frac{1}{K}\sum_{k=1}^{K}\sqrt{2R_k^W}.\end{aligned}$$

Combining the results, we evaluate the inequality (14) as

$$\begin{aligned}&\frac{1}{2n}\|\mathfrak{X}(\Delta_X)\|^2 \\ &\leq \frac{1}{n}\|\|\mathfrak{X}^*(\mathcal{E})\|\|_m\|\|\Delta_X\|\|_s + \lambda_n\|\|\Delta_W\|\|_s + \mu_n(\|\|\overline{\Delta_W}\|\|_F^2 + A) \\ &\leq \lambda^* C_\Gamma\|\|\Delta_W\|\|_F \\ &\quad + \lambda_n\left(\|\|\Delta_W\|\|_F\frac{1}{K}\sum_{k=1}^{K}\sqrt{2R_k^W} + B\right) \\ &\quad + \mu_n(\|\|\overline{\Delta}_W\|\|_F^2 + A).\end{aligned}$$

By the RSC condition and the same result in (15), we have

$$\begin{aligned}\|\|\Delta_X\|\|_F^2 &\leq \frac{1}{2n}\|\mathfrak{X}(\Delta_X)\|^2 \\ &\leq (\lambda^* C'_\Gamma + \lambda_n)\|\|\Delta_W\|\|_F\frac{1}{K}\sum_{k=1}^{K}\sqrt{R_k^W} + \mu_n\|\|\Delta_W\|\|_F^2 + \lambda_n B + \mu_n A \\ &\leq C_P(\lambda^* C'_\Gamma + \lambda_n)\|\|\Delta_X\|\|_F\frac{1}{K}\sum_{k=1}^{K}\sqrt{R_k^W} + C_P\mu_n\|\|\Delta_X\|\|_F^2 + \lambda_n B + \mu_n A.\end{aligned}$$

where $C'_\Gamma = C_\Gamma(\frac{1}{K}\sum_{k=1}^{K}\sqrt{R_k^W})^{-1}$. When $A = 0$ and $B = 0$, using the condition of the constant, we have

$$\begin{aligned}&(1 - C_P\mu_n)\|\|\Delta_X\|\|_F^2 \\ &\leq \kappa_n\|\|\Delta_X\|\|_F\frac{1}{K}\sum_{k=1}^{K}\sqrt{R_k^W},\end{aligned}$$

by the setting of κ_n. Then we have

$$\|\|\Delta_X\|\|_F \lesssim \frac{\kappa_n}{K}\sum_{k=1}^{K}\sqrt{R_k^W}. \tag{17}$$

Then the claim holds. □

A.3. Proof of Lemma 4

Proof. This proof start from the inequality (17) used in the proof of Lemma 2. By the settings of Lemma 4, we have $A = 0$ and $B = 0$. Using the setting of the basis functions, we have

$$\|\|\Delta_W\|\|_F^2 \leq C_P^2\|\|\Delta_X\|\|_F^2. \tag{18}$$

To evaluate the convergence, we bound the following term. For all $g \in [0,1]^K$, we have

$$
\begin{aligned}
&|f_{\hat{X}}(g) - f_{X^*}(g)| \\
&= \left| \sum_{k=1}^{K} \sum_{m_k=1}^{M^{(k)}} (\hat{w}_{m_1 \dots m_K} - w^*_{m_1 \dots m_K}) \phi^{(1)}_{m_1}(g_1) \cdots \phi^{(K)}_{m_K}(g_K) \right| \\
&\leq \left(\sum_{k=1}^{K} \sum_{m_k=1}^{M^{(k)}} (\hat{w}_{m_1 \dots m_K} - w^*_{m_1 \dots m_K})^2 \right)^{1/2} \\
&\quad \times \left(\sum_{k=1}^{K} \sum_{m_k=1}^{M^{(k)}} (\phi^{(1)}_{m_1}(g_1) \cdots \phi^{(K)}_{m_K}(g_K))^2 \right)^{1/2},
\end{aligned}
$$

by the boundedness property of $\{\phi_m\}_m$.

Then, we obtain

$$
\sup_{g \in [0,1]^K} |f_{\hat{X}}(g) - f_{X^*}(g)| \leq C_g |||\Delta_W|||,
$$

where C_g is a positive constant. Combining the result in (18), we obtain the claim.

□

A.4. Proof of Theorem 5 and Theorem 6

Proof. Using Lemma 8, we obtain the regularization parameter bounding $|||\mathfrak{X}^*(\mathcal{E})|||_m$. Then, we substitute the parameter into the result of Lemma 2 and 6, thus we obtain the claim.

□