A Proofs of Theorems 1 and 2

A.1 ML-IALM

In the subsection we give the missing proof of theorem 1. First let us remind a few definitions. At each iteration of IALM algorithm we solve the following subproblem

$$\min_{L \in \mathbb{R}^{m \times n}} \{ \|L\|_* + \frac{1}{2\tau} \|M - L\|_F^2 \}. \quad (1)$$

The minimiser of (1) is given in closed form via the singular value thresholding operator as

$$\hat{L} = D_\tau[M], \quad (2)$$

where $D_\tau[M]$ is defined for the SVD of $M = U\Sigma V^\top$ as

$$D_\tau[M] = US_\tau[\Sigma]V^\top, \quad (3)$$

where $S_\tau[\cdot]$ is the soft thresholding operator defined element-wise:

$$S_\tau[x] = (|x| - \tau)_+ \text{sgn}(x). \quad (4)$$

Then for a full rank restriction operator $R \in \mathbb{R}^{n \times n}$ we define the multilevel SVT (ML-SVT) operator

$$D_H^T[M] = U_H S_\tau[\Sigma_H]V_H^\top R, \quad (5)$$

where $MR = U_H \Sigma_H V_H^\top$ is the SVD of the coarse $M_H = MR$. Then Theorem 1 shows that $D_H^T$ gives an approximate solution to the problem (1).

**Theorem 1.** For any $R$, such that $\|R\|_2 \leq 1$ and $0 < \tau \leq \sigma_{H,1}$, the ML-SVT operator $D_H^T[M]$ gives a $(\frac{2\tau}{\sigma_{H,1}}(\sigma_1 + \sigma_{H,1} - \tau))$-approximate solution to the problem (1), where $\sigma_{H,1}$ is the largest singular value of $MR$.

**Proof.** The proof follows the steps of the proof of Theorem 2.1 of [1]. First note that (1) is a convex problem with optimality criteria

$$M - \hat{L} \in \tau \partial \|L\|_*, \quad (6)$$

where $\partial \|L\|_*$ is the set of subgradients of the nuclear norm. Let $L \in \mathbb{R}^{m \times n}$ be an arbitrary matrix and $U\Sigma V^\top$ be its SVD. It is known [2] that

$$\partial \|L\|_* = \{ UV^\top + W : W \in \mathbb{R}^{m \times n}, U^\top W = \mathbf{0}, WV = \mathbf{0}, \|W\|_2 \leq 1 \}. \quad (7)$$

Next we set $L^H = D_H^T[M]$ and find an upper bound for the distance from $M - L^H$ to the set $\tau \partial \|L\|_*$. Then we decompose the SVDs of $M$ and $MR$ as $M = U_0 \Sigma_0 V_0^\top + U_1 \Sigma V_1^\top$ and $MR = U_{H,0} \Sigma_{H,0} V_{H,0}^\top + U_{H,1} \Sigma_{H,1} V_{H,1}^\top$, where $U_0, U_{H,0}, V_0$ and $V_{H,0}$ (respectively, $U_1, U_{H,1}, V_1$ and $V_{H,1}$) are the corresponding singular vectors associated with the singular values greater than...
\( \tau \) (respectively, smaller than or equal to \( \tau \)). Then we have \( L^H = U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T \), and therefore using \( U_0 U_0^T + U_1 U_1^T = I \) and \( U_0^T U_0 = I \) we get

\[
M - L^H = U_0 \Sigma_0 V_0^T + U_1 \Sigma_1 V_1^T - U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T
\]

\[
= U_0 \Sigma_0 V_0^T + U_1 \Sigma_1 V_1^T - (U_0^T U_0 + U_1^T U_1) U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T
\]

\[
= \tau^{-1} U_1 (\Sigma_1 V_1^T - U_1^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T)
\]

\[
+ \tau^{-1} U_0 (\Sigma_0 V_0^T - U_0^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T)
\]

\[
:= \tau [W + UV^T],
\]

where

\[
W := \tau^{-1} U_1 (\Sigma_1 V_1^T - U_1^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T),
\]

\[
V^T := \tau^{-1} (\Sigma_0 V_0^T - U_0^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T)
\]

and \( U := U_0 \). Then \( U^T W = 0 \) and since \( \| \Sigma_1 \|_2 \leq \tau \) and \( \sigma_{H,1} \geq \tau \) we also have

\[
\| W \|_2
\]

\[
= \tau^{-1} \| \Sigma_1 V_1^T - U_1^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T \|_2
\]

\[
\leq \tau^{-1} (\| \Sigma_1 \|_2 + \| \Sigma_{H,0} - \tau I \|_2 \| R^T \|_2)
\]

\[
\leq \tau^{-1} (\tau + \sigma_{H,1} - \tau)
\]

\[
= \frac{\sigma_{H,1}}{\tau}
\]

Furthermore,

\[
\| W V \|_2
\]

\[
= \tau^{-2} \left\| U_1 (\Sigma_1 V_1^T - U_1^T U_{H,0}(\Sigma_{H,0} - \tau I) V_{H,0}^T R^T)
\right\|_2
\]

\[
\leq \tau^{-2} (\| \Sigma_1 \|_2 + \| \Sigma_{H,0} - \tau I \|_2) (\| \Sigma_0 \|_2 + \| \Sigma_{H,0} - \tau I \|_2)
\]

\[
\leq \tau^{-2} (\tau + \sigma_{H,1} - \tau) (\sigma_1 + \sigma_{H,1} - \tau)
\]

\[
= \frac{\sigma_{H,1}}{\tau^2} (\sigma_1 + \sigma_{H,1} - \tau),
\]

where \( \sigma_{H,1} = \| \Sigma_{H,0} \|_2 \) is the largest singular value of \( MR \) and for the last inequality we used the assumptions that \( \sigma_{H,1} \geq \tau \) and \( \| R \|_2 \leq 1 \). Therefore, since \( \frac{\sigma_{H,1}}{\tau^2} (\sigma_1 + \sigma_{H,1} - \tau) \geq \frac{\sigma_{H,1}}{\tau} \) then \( D^H(M) \) is at most \( \frac{\sigma_{H,1}}{\tau} (\sigma_1 + \sigma_{H,1} - \tau) \) away from a zero subdifferential of \( (1) \). \( \Box \)
A.2 ML-AltProj

In this section we give the proof of Theorem 2. Here the optimisation problem in question is given as

\[
\min_{L \in \mathbb{R}^{m \times n}} \| D - L - S \|_2 \quad s.t. \quad \text{rank}(L) \leq l. \tag{13}
\]

Then we solve (13) using the multilevel hard thresholding operator defined as

\[ L^H = U^H \mathcal{H}_k | \Sigma_H | V^T_H R^T, \tag{14} \]

where \( MR = M_H = U_H \Sigma_H V_H^T \) is a SVD of the coarse model, \( \mathcal{H} \) is the hard thresholding operator and \( R \) is the restriction operator.

**Theorem 2.** The multilevel low rank approximation procedure given in (14) gives a \((\sigma_1 + \sigma_{H,1})\)-approximate solution to the problem (13), where \( \sigma_{H,1} \) is the largest singular value of \( M_H = MR \).

**Proof.** First note that for any \( B = XY^T \) (\( X, Y \in \mathbb{R}^{m \times k} \)) we can find a vector \( \omega = \sum_{i=1}^{k+1} \gamma_i v_i \) so that \( \omega^T X = 0 \) and \( ||\omega||_2 = \sum_{i=1}^{k+1} \gamma_i^2 = 1 \), where \( v_i \) are the columns of \( V \). Then

\[
\| M - B \|_2^2 \geq \| \omega^T (M - B) \|_2^2 = \| \omega^T M \|_2^2 = \sum_{i=1}^{k+1} \gamma_i^2 \sigma_i^2 \geq \sigma_{k+1}^2. \tag{15}
\]

On the other hand

\[
\| M - L^H \|_2 = \| U_0 \Sigma_0 V_0^T + U_1 \Sigma_1 V_1^T - U_{H,0} \Sigma_{H,0} V_{H,0}^T R^T \|_2 \tag{16}
\]

\[
\leq \sigma_1 + \sigma_{k+1} + \sigma_{H,1},
\]

where \( U_0, V_0, U_{H,0} \) and \( V_{H,0} \) (respectively, \( U_1, V_1, U_{H,1} \) and \( V_{H,1} \)) are correspondingly the singular vectors associated with the largest \( k \) (respectively, smallest \( n - k \) and \( n_H - k \)) singular values of \( M \) and \( MR \). From these two inequalities we have that for any \( B \)

\[
\| M - L^H \|_2 - \| M - B \|_2 \leq \sigma_1 + \sigma_{H,1}. \tag{17}
\]

\[\square\]

**References**
