

Supplementary Material

(1) Proof of Claim 1 in Lemma 2 (EDA):

Define $\mathcal{J}(\boldsymbol{\beta}^t, \boldsymbol{\Theta}^t, \mathbf{U}^t)$ as the joint objective function at iteration t . When fix $\boldsymbol{\Theta}^t$, there is

$$\begin{aligned} & C_S \|\mathbf{H}_S \boldsymbol{\beta}^t - \mathbf{T}_S\|_F^2 + C_T \|\mathbf{H}_T \boldsymbol{\beta}^t - \mathbf{T}_T \circ \boldsymbol{\Theta}^t\|_F^2 + \gamma \|\boldsymbol{\Theta}^t - \mathbf{I}\|_F^2 + \tau \|\mathbf{H}_{\mathcal{T}u} \boldsymbol{\beta}^t - \boldsymbol{\Phi}_{p,\mathcal{T}u}\|_F^2 + \lambda \cdot \text{tr}((\boldsymbol{\beta}^t)^T \mathbf{H}^T \mathbf{L} \mathbf{H} \boldsymbol{\beta}^t) \\ & \geq \\ & C_S \|\mathbf{H}_S \boldsymbol{\beta}^{t+1} - \mathbf{T}_S\|_F^2 + C_T \|\mathbf{H}_T \boldsymbol{\beta}^{t+1} - \mathbf{T}_T \circ \boldsymbol{\Theta}^t\|_F^2 + \gamma \|\boldsymbol{\Theta}^t - \mathbf{I}\|_F^2 + \tau \|\mathbf{H}_{\mathcal{T}u} \boldsymbol{\beta}^{t+1} - \boldsymbol{\Phi}_{p,\mathcal{T}u}\|_F^2 + \lambda \cdot \\ & \text{tr}((\boldsymbol{\beta}^{t+1})^T \mathbf{H}^T \mathbf{L} \mathbf{H} \boldsymbol{\beta}^{t+1}) \end{aligned} \quad (3.1)$$

Note that with fixed $\boldsymbol{\Theta}^t$, the model form a ℓ_2 -norm based least square problem with respect to $\boldsymbol{\beta}$, which is always monotonically non-increasing. So, the inequality (3.1) holds.

According to (3), we know that $\|\boldsymbol{\beta}\|_{2,1} = \sum_{i=1}^L \|\boldsymbol{\beta}_i\|_2$, then the summation $\|\boldsymbol{\beta}^t\|_{2,1} + \sum_{i=1}^L \left(\frac{\|\boldsymbol{\beta}_i^t\|_2^2}{2\|\boldsymbol{\beta}_i^t\|_2} - \|\boldsymbol{\beta}_i^t\|_2 \right)$ is a constant. Therefore, there is

$$\begin{aligned} & \|\boldsymbol{\beta}^t\|_{2,1} + \sum_{i=1}^L \left(\frac{\|\boldsymbol{\beta}_i^t\|_2^2}{2\|\boldsymbol{\beta}_i^t\|_2} - \|\boldsymbol{\beta}_i^t\|_2 \right) + C_S \|\mathbf{H}_S \boldsymbol{\beta}^t - \mathbf{T}_S\|_F^2 + C_T \|\mathbf{H}_T \boldsymbol{\beta}^t - \mathbf{T}_T \circ \boldsymbol{\Theta}^t\|_F^2 + \gamma \|\boldsymbol{\Theta}^t - \mathbf{I}\|_F^2 + \tau \|\mathbf{H}_{\mathcal{T}u} \boldsymbol{\beta}^t - \boldsymbol{\Phi}_{p,\mathcal{T}u}\|_F^2 \\ & + \lambda \cdot \text{tr}((\boldsymbol{\beta}^t)^T \mathbf{H}^T \mathbf{L} \mathbf{H} \boldsymbol{\beta}^t) \\ & \geq \\ & \|\boldsymbol{\beta}^{t+1}\|_{2,1} + \sum_{i=1}^L \left(\frac{\|\boldsymbol{\beta}_i^{t+1}\|_2^2}{2\|\boldsymbol{\beta}_i^{t+1}\|_2} - \|\boldsymbol{\beta}_i^{t+1}\|_2 \right) + C_S \|\mathbf{H}_S \boldsymbol{\beta}^{t+1} - \mathbf{T}_S\|_F^2 + C_T \|\mathbf{H}_T \boldsymbol{\beta}^{t+1} - \mathbf{T}_T \circ \boldsymbol{\Theta}^t\|_F^2 + \gamma \|\boldsymbol{\Theta}^t - \mathbf{I}\|_F^2 + \\ & \tau \|\mathbf{H}_{\mathcal{T}u} \boldsymbol{\beta}^{t+1} - \boldsymbol{\Phi}_{p,\mathcal{T}u}\|_F^2 + \lambda \cdot \text{tr}((\boldsymbol{\beta}^{t+1})^T \mathbf{H}^T \mathbf{L} \mathbf{H} \boldsymbol{\beta}^{t+1}) \end{aligned} \quad (3.2)$$

According to Lemma 1, we know that

$$\sum_{i=1}^L \left(\frac{\|\boldsymbol{\beta}_i^t\|_2^2}{2\|\boldsymbol{\beta}_i^t\|_2} - \|\boldsymbol{\beta}_i^t\|_2 \right) \leq \sum_{i=1}^L \left(\frac{\|\boldsymbol{\beta}_i^{t+1}\|_2^2}{2\|\boldsymbol{\beta}_i^{t+1}\|_2} - \|\boldsymbol{\beta}_i^{t+1}\|_2 \right) \quad (3.3)$$

Therefore, combine (3.2) with (3.3), we can obtain that

$$\mathcal{J}(\boldsymbol{\beta}^t, \boldsymbol{\Theta}^t, \mathbf{U}^t) \geq \mathcal{J}(\boldsymbol{\beta}^{t+1}, \boldsymbol{\Theta}^t, \mathbf{U}^{t+1}) \quad (3.4)$$

Then, Claim 1 is proven.

Note that \mathbf{U}^t is completely determined when fix $\boldsymbol{\beta}^t$ according to (15), thus \mathbf{U}^{t+1} is also determined when $\boldsymbol{\beta}^{t+1}$ is fixed.

(2) Proof of Claim 2 in Lemma 2 (EDA):

When fix $\boldsymbol{\beta}^{t+1}$, \mathbf{U}^{t+1} is also fixed, and the objective function is convex with respect to $\boldsymbol{\Theta}$. As can be

seen from (18), the update rule of $\boldsymbol{\Theta}$ can be obtained by setting $\frac{d\mathcal{J}(\boldsymbol{\beta}^{t+1}, \boldsymbol{\Theta}, \mathbf{U}^{t+1})}{d\boldsymbol{\Theta}} = 0$, then there is

$$\frac{d\mathcal{J}(\boldsymbol{\beta}^{t+1}, \boldsymbol{\Theta}, \mathbf{U}^{t+1})}{d\boldsymbol{\Theta}} = -2C_T \mathbf{T}_T^T \mathbf{H}_T \boldsymbol{\beta}^{t+1} + 2C_T \mathbf{T}_T^T \mathbf{T}_T \boldsymbol{\Theta}^{t+1} + 2\gamma \boldsymbol{\Theta}^{t+1} - 2\gamma \mathbf{I} = 0 \quad (3.5)$$

Therefore, from (3.5) the update rule of $\boldsymbol{\Theta}$ is obtained as

$$\boldsymbol{\Theta}^{t+1} = (C_T \mathbf{T}_T^T \mathbf{T}_T + \gamma \mathbf{I})^{-1} (C_T \mathbf{T}_T^T \mathbf{H}_T \boldsymbol{\beta}^{t+1} + \gamma \mathbf{I}) \quad (3.6)$$

Additionally, since the second order derivative of the objective function with respect to Θ is

$$\frac{d^2 \mathcal{J}(\beta^{t+1}, \Theta, \mathbf{U}^{t+1})}{d\Theta^2} = 2C_T \mathbf{T}_T^T \mathbf{T}_T + 2\gamma > 0 \quad (3.7)$$

From (3.7), we know that the objective function is convex with respect to Θ , so the update rule (3.6) can minimize the objective function, there is

$$\mathcal{J}(\beta^{t+1}, \Theta^t, \mathbf{U}^{t+1}) \geq \mathcal{J}(\beta^{t+1}, \Theta^{t+1}, \mathbf{U}^{t+1}) \quad (3.8)$$

Then, *Claim 2* is proven.

(3) Proof of Claim 3 in Lemma 2 (MvEDA):

The proof of *Claim 3* is similar with the proof of *Claim 1*, which is shown as follows.

Define $\mathcal{J}(\beta_v^t, \Theta_v^t, \alpha_v^t, \mathbf{U}_v^t)$ as the joint objective function at iteration t . When fix Θ_v^t, α_v^t , there is

$$\begin{aligned} & C_S \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{S,v} \beta_v^t - \mathbf{T}_S\|_F^2 + C_T \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{T,v} \beta_v^t - \mathbf{T}_T \circ \Theta_v^t\|_F^2 + \gamma \sum_{v=1}^V \alpha_v^t \|\Theta_v^t - \mathbf{I}\|_F^2 + \tau \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{\mathcal{T}u,v} \beta_v^t - \\ & \Phi_{p,\mathcal{T}u}^{k,v}\|_F^2 + \lambda \cdot \sum_{v=1}^V \alpha_{v,t}^r \text{tr}((\beta_v^t)^T \mathbf{H}_v^T \mathcal{L}_v \mathbf{H}_v \beta_v^t) \\ & \geq \\ & C_S \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{S,v} \beta_v^{t+1} - \mathbf{T}_S\|_F^2 + C_T \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{T,v} \beta_v^{t+1} - \mathbf{T}_T \circ \Theta_v^t\|_F^2 + \gamma \sum_{v=1}^V \alpha_v^t \|\Theta_v^t - \mathbf{I}\|_F^2 + \\ & \tau \sum_{v=1}^V \alpha_v^t \|\mathbf{H}_{\mathcal{T}u,v} \beta_v^{t+1} - \Phi_{p,\mathcal{T}u}^{k,v}\|_F^2 + \lambda \cdot \sum_{v=1}^V \alpha_{v,t}^r \text{tr}((\beta_v^{t+1})^T \mathbf{H}_v^T \mathcal{L}_v \mathbf{H}_v \beta_v^{t+1}) \end{aligned} \quad (4.1)$$

Note that with fixed Θ_v^t, α_v^t , the model form a ℓ_2 -norm based least square problem with respect to β , which is always monotonically non-increasing. So, the inequality (4.1) holds.

According to (3), we know that $\sum_{v=1}^V \|\beta_v\|_{2,1} = \sum_{v=1}^V \sum_{i=1}^L \|\beta_{i,v}\|_2$, then the summation

$$\begin{aligned} & \sum_{v=1}^V \|\beta_v^t\|_{2,1} + \sum_{v=1}^V \sum_{i=1}^L \left(\frac{\|\beta_{i,v}^t\|_2^2}{2\|\beta_{i,v}^t\|_2} - \|\beta_{i,v}^t\|_2 \right) \text{ is a constant. Therefore, there is} \\ & \sum_{v=1}^V \|\beta_v^t\|_{2,1} + \sum_{v=1}^V \sum_{i=1}^L \left(\frac{\|\beta_{i,v}^t\|_2^2}{2\|\beta_{i,v}^t\|_2} - \|\beta_{i,v}^t\|_2 \right) + \sum_{v=1}^V C_S \alpha_v^t \|\mathbf{H}_{S,v} \beta_v^t - \mathbf{T}_S\|_F^2 + \sum_{v=1}^V C_T \alpha_v^t \|\mathbf{H}_{T,v} \beta_v^t - \mathbf{T}_T \circ \\ & \Theta_v^t\|_F^2 + \sum_{v=1}^V \gamma \alpha_v^t \|\Theta_v^t - \mathbf{I}\|_F^2 + \sum_{v=1}^V \tau \alpha_v^t \|\mathbf{H}_{\mathcal{T}u,v} \beta_v^t - \Phi_{p,\mathcal{T}u}^{k,v}\|_F^2 + \lambda \cdot \sum_{v=1}^V \alpha_{v,t}^r \text{tr}((\beta_v^t)^T \mathbf{H}_v^T \mathcal{L}_v \mathbf{H}_v \beta_v^t) \\ & \geq \\ & \sum_{v=1}^V \|\beta_v^{t+1}\|_{2,1} + \sum_{v=1}^V \sum_{i=1}^L \left(\frac{\|\beta_{i,v}^{t+1}\|_2^2}{2\|\beta_{i,v}^{t+1}\|_2} - \|\beta_{i,v}^{t+1}\|_2 \right) + \sum_{v=1}^V C_S \alpha_v^t \|\mathbf{H}_{S,v} \beta_v^{t+1} - \mathbf{T}_S\|_F^2 + \sum_{v=1}^V C_T \alpha_v^t \|\mathbf{H}_{T,v} \beta_v^{t+1} - \\ & \mathbf{T}_T \circ \Theta_v^t\|_F^2 + \sum_{v=1}^V \gamma \alpha_v^t \|\Theta_v^t - \mathbf{I}\|_F^2 + \sum_{v=1}^V \tau \alpha_v^t \|\mathbf{H}_{\mathcal{T}u,v} \beta_v^{t+1} - \Phi_{p,\mathcal{T}u}^{k,v}\|_F^2 + \lambda \cdot \sum_{v=1}^V \alpha_{v,t}^r \text{tr}((\beta_v^{t+1})^T \mathbf{H}_v^T \mathcal{L}_v \mathbf{H}_v \beta_v^{t+1}) \end{aligned} \quad (4.2)$$

According to Lemma 1, we know that

$$\sum_{v=1}^V \sum_{i=1}^L \left(\frac{\|\beta_{i,v}^t\|_2^2}{2\|\beta_{i,v}^t\|_2} - \|\beta_{i,v}^t\|_2 \right) \leq \sum_{v=1}^V \sum_{i=1}^L \left(\frac{\|\beta_{i,v}^{t+1}\|_2^2}{2\|\beta_{i,v}^{t+1}\|_2} - \|\beta_{i,v}^{t+1}\|_2 \right) \quad (4.3)$$

Therefore, combine (4.2) with (4.3), we can obtain that

$$\mathcal{J}(\beta_v^t, \Theta_v^t, \alpha_v^t, \mathbf{U}_v^t) \geq \mathcal{J}(\beta_v^{t+1}, \Theta_v^t, \alpha_v^t, \mathbf{U}_v^{t+1}) \quad (4.4)$$

Then, *Claim 3* is proven.

Note that \mathbf{U}_v^t is completely determined when fix β_v^t according to (15), thus \mathbf{U}_v^{t+1} is also determined

when β_v^{t+1} is fixed.

(4) Proof of Claim 4 in Lemma 2 (MvEDA):

The proof of *Claim 4* is similar with the proof of *Claim 2*.

When fix β_v^{t+1} , \mathbf{U}_v^{t+1} is also fixed, and the objective function is convex with respect to Θ_v . As can be seen from (29), the update rule of Θ_v can be obtained by setting $\frac{d\mathcal{J}(\beta_v^{t+1}, \Theta_v, \alpha_v^{t+1}, \mathbf{U}_v^{t+1})}{d\Theta_v} = 0$, then there is

$$\frac{d\mathcal{J}(\beta_v^{t+1}, \Theta_v, \alpha_v^{t+1}, \mathbf{U}_v^{t+1})}{d\Theta_v} = -2C_{\mathcal{T}}\alpha_v^{t+1}\mathbf{T}_{\mathcal{T}}^T\mathbf{H}_{\mathcal{T},v}\beta_v^{t+1} + 2C_{\mathcal{T}}\alpha_v^{t+1}\mathbf{T}_{\mathcal{T}}^T\mathbf{T}_{\mathcal{T}}\Theta_v^{t+1} + 2\gamma\alpha_v^{t+1}\Theta_v^{t+1} - 2\gamma\alpha_v^{t+1}\mathbf{I} = 0 \quad (4.5)$$

Therefore, from (3.5) the update rule of Θ_v is obtained as

$$\Theta_v^{t+1} = (C_{\mathcal{T}}\alpha_v^{t+1}\mathbf{T}_{\mathcal{T}}^T\mathbf{T}_{\mathcal{T}} + \gamma\alpha_v^{t+1}\mathbf{I})^{-1}(C_{\mathcal{T}}\alpha_v^{t+1}\mathbf{T}_{\mathcal{T}}^T\mathbf{H}_{\mathcal{T},v}\beta_v^{t+1} + \gamma\alpha_v^{t+1}\mathbf{I}) \quad (4.6)$$

Additionally, since the second order derivative of the objective function with respect to Θ_v is

$$\frac{d^2\mathcal{J}(\beta_v^{t+1}, \Theta_v, \alpha_v^{t+1}, \mathbf{U}_v^{t+1})}{d\Theta_v^2} = 2C_{\mathcal{T}}\alpha_v^{t+1}\mathbf{T}_{\mathcal{T}}^T\mathbf{T}_{\mathcal{T}} + 2\gamma\alpha_v^{t+1} > 0 \quad (\alpha_v > 0) \quad (4.7)$$

From (4.7), we know that the objective function is convex with respect to Θ , so the update rule (4.6) can minimize the objective function, there is

$$\mathcal{J}(\beta_v^{t+1}, \Theta_v^t, \alpha_v^t, \mathbf{U}_v^{t+1}) \geq \mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v^t, \mathbf{U}_v^{t+1}) \quad (4.8)$$

Then, *Claim 4* is proven.

(5) Proof of Claim 5 in Lemma 3 (MvEDA).

As can be seen from (31), the update rule of α_v is obtained by setting $\frac{d\mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v, \mathbf{U}_v^{t+1}, \eta)}{d\alpha_v} = 0$ and

$\frac{d\mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v, \mathbf{U}_v^{t+1}, \eta)}{d\eta} = 0$. Additionally, when r is set as 2, the 2nd order derivative of the objective function with respect to α_v is

$$\frac{d^2\mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v, \mathbf{U}_v^{t+1}, \eta)}{d\alpha_v^2} = 2 \cdot \text{tr}(\beta_v^T \mathbf{H}_v^T \mathcal{L}_v \mathbf{H}_v \beta_v) \geq 0 \quad (4.9)$$

Therefore, we know that the objective function is convex with respect to α_v , there is

$$\mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v^t, \mathbf{U}_v^{t+1}) \geq \mathcal{J}(\beta_v^{t+1}, \Theta_v^{t+1}, \alpha_v^{t+1}, \mathbf{U}_v^{t+1}) \quad (4.10)$$

Then, *Claim 5* is proven.