A Convolution Bias-Incorporated Nonnegative Latent Factorization of Tensors Model for Accurate Representation Learning to Dynamic Directed Graphs

Supplementary File

Qu Wang, Hao Wu, *Member, IEEE*, and Xin Luo, *Fellow, IEEE*

I. INTRODUCTION

This is the supplementary file for paper entitled *A Convolution Bias-Incorporated Nonnegative Latent Factorization of Tensors Model for Accurate Representation Learning to Dynamic Directed Graphs*. Supplementary equations and experimental results are put into this file.

II. SUPPLEMENTARY TABLE

Table S1 records the notation descriptions used in this paper.

TABLE S1
SYMBOLS DESCRIPTION

Symbol	Description
I, J, K	Three entity sets.
Λ, Γ	Known and unknown sets of an HDI tensor.
$\lvert\cdot\rvert$	Cardinality of an enclosed set.
$\lVert\cdot\rVert_F$	Compute the Frobenius norm of a tensor
$\mathbb{R}$	Real number domain.
$\mathbf{Y}$	Three-order tensors.
$\hat{\mathbf{Y}}$	Low-rank approximation to $\mathbf{Y}$.
$\mathbf{Z}$	Three-order tensors.
$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$	Three-order bias tensors.
y_{ijk}, $\hat{y}_{ijk}$	Single element in $\mathbf{Y}$ and $\hat{\mathbf{Y}}$.
z_{ijk}, a_{ijk}, b_{ijk}, c_{ijk}	Single element in $\mathbf{Z}$, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.
$\mathbf{G}$	A Core tensor.
$\mathbf{H}$	Shared core tensor.
G_{mnl}, h_{mnl}	Single element in $\mathbf{G}$ and $\mathbf{H}$.
P, Q, T	Latent feature matrices.
$\mathrm{U}^{(*)}$, $\mathrm{V}^{(*)}$, $\mathrm{W}^{(*)}$	Bias matrices
D, E, F	Bias latent feature matrices
R_1, R_2, R_3	Rank of $\hat{\mathbf{Y}}$, LF dimension of each LF matrix, and size of core tensors' each dimension.
m, n, l	Index of R_1, R_2, *and* R_3.
p_{mr}, q_{nr}, t_{lr}	Single element in P, Q, and T.
d_{mr}, e_{nr}, f_{lr}	Single element in D, E, and F.
g_{ijk}, h_{ijk}	Single element in $\mathbf{G}$ and $\mathbf{H}$.
$\Lambda(i)$, $\Lambda(j)$, $\Lambda(k)$	Subsets of Λ linked with $i\in I$, $j\in J$, and $k\in K$.
λ	Regularization coefficient.
η, μ	Learning rate.
t	Current training iteration count.
N	Total number of training rounds.
Φ, Ψ, Ω	Training set, validating set, testing set

III. SUPPLEMENTARY PROOF OF CONVERGENCE

A. Build Lagrangian Function

To establish the relationship between the SLF-NMUT algorithm and the KKT conditions of the learning objective (13), we introduce Lagrangian multipliers for the nonnegativity constraints on core tensors **G**, **H**, latent feature matrices P, Q, T, and bias matrices D, E, F. Then, the Lagrangian function L for (10) is formulated as follows:

$$L=\varepsilon\left(\mathbf{G},\mathrm{P},\mathrm{Q},\mathrm{T},\mathbf{H},\mathrm{D},\mathrm{E},\mathrm{F}\right)-\sum_{m=1}^{R_1}\sum_{n=1}^{R_2}\sum_{l=1}^{R_3}\tilde{g}_{mnl}g_{mnl}-\sum_{i=1}^{|I|}\sum_{m=1}^{R_1}\tilde{p}_{im}p_{im}-\sum_{j=1}^{|J|}\sum_{n=1}^{R_2}\tilde{q}_{jn}q_{jn}-\sum_{k=1}^{|K|}\sum_{l=1}^{R_3}\tilde{t}_{kl}t_{kl}$$
$$-\sum_{m=1}^{R_1}\sum_{n=1}^{R_2}\sum_{l=1}^{R_3}\tilde{h}_{mnl}h_{mnl}-\sum_{i=1}^{|I|}\sum_{m=1}^{R_1}\tilde{d}_{im}d_{im}-\sum_{j=1}^{|J|}\sum_{n=1}^{R_2}\tilde{e}_{jn}e_{jn}-\sum_{k=1}^{|K|}\sum_{l=1}^{R_3}\tilde{f}_{kl}f_{kl}. \tag{S1}$$

Considering that P, Q, T are highly similar, so are D, E, and F. We only take partial derivatives of g_{mnl}, p_{im}, h_{mnl} and d_{im} of L.

$$\begin{cases}\dfrac{\partial L}{\partial g_{mnl}}=\dfrac{\partial\varepsilon}{\partial g_{mnl}}-\tilde{g}_{mnl}=0, & \dfrac{\partial L}{\partial p_{im}}=\dfrac{\partial\varepsilon}{\partial p_{im}}-\tilde{p}_{im}=0,\\ \dfrac{\partial L}{\partial h_{mnl}}=\dfrac{\partial\varepsilon}{\partial h_{mnl}}-\tilde{h}_{mnl}=0, & \dfrac{\partial L}{\partial d_{im}}=\dfrac{\partial\varepsilon}{\partial d_{im}}-\tilde{d}_{im}=0.\end{cases}$$
$$\Downarrow \tag{S2}$$
$$\begin{cases}\tilde{g}_{mnl}=\dfrac{\partial\varepsilon}{\partial g_{mnl}},\ \tilde{p}_{im}=\dfrac{\partial\varepsilon}{\partial p_{im}},\ \tilde{h}_{mnl}=\dfrac{\partial\varepsilon}{\partial h_{mnl}},\ \tilde{d}_{im}=\dfrac{\partial\varepsilon}{\partial d_{im}}.\end{cases}$$

Then, considering the KKT conditions of (13), i.e., $\tilde{g}_{mnl}g_{mnl}=0,\tilde{p}_{im}p_{im}=0,\tilde{h}_{mnl}h_{mnl}=0,\tilde{d}_{im}d_{im}=0,$ we multiply both sides of each equation by their respective factors to get the following formula:

$$\begin{cases}0=g_{mnl}\dfrac{\partial\varepsilon}{\partial g_{mnl}},\ 0=p_{im}\dfrac{\partial\varepsilon}{\partial p_{im}},\ 0=h_{mnl}\dfrac{\partial\varepsilon}{\partial h_{mnl}},\ 0=d_{im}\dfrac{\partial\varepsilon}{\partial d_{im}}.\end{cases} \tag{S3}$$

According to (10), bring the partial derivative into (S3):

$$\begin{cases}g_{mnl}\sum\limits_{y_{ijk}\in\Lambda}\left(\left(y_{ijk}-\hat{y}_{ijk}\right)\left(-p_{im}q_{jn}t_{kl}\right)+\lambda g_{mnl}\right)=0,\\ p_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(\left(y_{ijk}-\hat{y}_{ijk}\right)\left(-\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}g_{mnl}q_{jn}t_{kl}\right)+\lambda p_{im}\right)=0,\\ h_{mnl}\sum\limits_{y_{ijk}\in\Lambda}\left(\left(y_{ijk}-\hat{y}_{ijk}\right)\left(-\left(d_{im}+e_{jn}+f_{kl}\right)\right)+\lambda h_{mnl}\right)=0,\\ d_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(\left(y_{ijk}-\hat{y}_{ijk}\right)\left(-\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}h_{mnl}\right)+\lambda d_{im}\right)=0.\end{cases} \tag{S4}$$

The following derivation is obtained from (S4), which is actually (13).

$$\begin{cases}g_{mnl}\sum\limits_{y_{ijk}\in\Lambda}y_{ijk}p_{im}q_{jn}t_{kl}=g_{mnl}\sum\limits_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}+\lambda g_{mnl}\right),\\ p_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(y_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}g_{mnl}q_{jn}t_{kl}\right)=p_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(\hat{y}_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}g_{mnl}q_{jn}t_{kl}+\lambda p_{im}\right),\\ h_{mnl}\sum\limits_{y_{ijk}\in\Lambda}\left(y_{ijk}\left(d_{im}+e_{jn}+f_{kl}\right)\right)=h_{mnl}\sum\limits_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}\left(d_{im}+e_{jn}+f_{kl}\right)+\lambda h_{mnl}\right),\\ d_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(y_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}h_{mnl}\right)=d_{im}\sum\limits_{y_{ijk}\in\Lambda(i)}\left(\hat{y}_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}h_{mnl}+\lambda d_{im}\right).\end{cases}\Rightarrow\begin{cases}g_{mnl}\leftarrow\dfrac{g_{mnl}\sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right)}{\sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}+\lambda g_{mnl}\right)},\\ p_{im}\leftarrow\dfrac{p_{im}\sum_{y_{ijk}\in\Lambda(i)}\left(y_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}g_{mnl}q_{jn}t_{kl}\right)}{\sum_{y_{ijk}\in\Lambda(i)}\left(\hat{y}_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}g_{mnl}q_{jn}t_{kl}+\lambda p_{im}\right)},\\ h_{mnl}\leftarrow\dfrac{h_{mnl}\sum_{y_{ijk}\in\Lambda}\left(y_{ijk}\left(d_{im}+e_{jn}+f_{kl}\right)\right)}{\sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}\left(d_{im}+e_{jn}+f_{kl}\right)+\lambda h_{mnl}\right)},\\ d_{im}\leftarrow\dfrac{d_{im}\sum_{y_{ijk}\in\Lambda(i)}\left(y_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}h_{mnl}\right)}{\sum_{y_{ijk}\in\Lambda(i)}\left(\hat{y}_{ijk}\sum\limits_{n=1}^{R_2}\sum\limits_{l=1}^{R_3}h_{mnl}+\lambda d_{im}\right)}.\end{cases} \tag{S5}$$

From (S1) to (S5), it becomes evident that the learning scheme based on SLF-NMUT is closely related to the KKT condition of its learning target.

B. Proof of Lemma 1

Utilizing Definition 3, we can deduce that:

$$F\left(x^t\right)=G\left(x^t,x^t\right)\geq G\left(x^{t+1},x^t\right)\geq F\left(x^{t+1}\right). \tag{S6}$$

Note that we have condition $F(x^{t+1}) = F(x^t)$ when x^t ensures the existence of a local minimum of $G(x, x^t)$. Consequently, condition $F'(x^t) = 0$ is valid if function F is differentiable in the vicinity of x^t. As a result, equation (S6) can be extended into the next converging sequence to $x_{\min} = \arg\min_x F(x)$:

$$F\left(x_{\min}\right)\leq\cdots\leq F\left(x^{t+1}\right)\leq F\left(x^t\right)\leq\cdots\leq F\left(x_1\right)\leq F\left(x_0\right). \tag{S7}$$

Next, our objective is to ensure that equation (13) for aligns perfectly with the one in equation (15), utilizing a specially designed G. Given a $g_{mnl} \in \mathbf{G}$, we define Fg_{mnl} as the partial loss derived from (10), which pertains exclusively to g_{mnl}.

$$\begin{aligned}Fg_{mnl}=&\frac{1}{2}\sum_{y_{ijk}\in\Lambda}\Bigg(\left(y_{ijk}-g_{mnl}p_{im}q_{jn}t_{kl}-\hat{y}_{ijk}^{\sim}\right)^2\\&+\lambda\left(\sum_{m=1}^{R_1}\sum_{n=1}^{R_2}\sum_{l=1}^{R_3}g_{mnl}^2+\sum_{m=1}^{R_1}p_{im}^2+\sum_{n=1}^{R_2}q_{jn}^2+\sum_{l=1}^{R_3}t_{kl}^2+\sum_{m=1}^{R_1}\sum_{n=1}^{R_2}\sum_{l=1}^{R_3}h_{mnl}^2+\sum_{m=1}^{R_1}d_{im}^2+\sum_{n=1}^{R_2}e_{jn}^2+\sum_{l=1}^{R_3}f_{kl}^2\right)\Bigg),\end{aligned} \tag{S8}$$

where $\hat{y}_{ijk}^{\sim}$ is the remaining term after $\hat{y}_{ijk}$ removes a certain m, n, l, and $\hat{y}_{ijk}^{\sim} = \hat{y}_{ijk} - g_{mnl}p_{im}q_{jn}t_{kl}$. Consequently, we have computed the first-order and second-order derivatives of Fg_{mnl} with respect to g_{mnl}.

$$F'g_{mnl}=\sum_{y_{ijk}\in\Lambda}\left(\left(y_{ijk}-\hat{y}_{ijk}\right)\left(-p_{im}q_{jn}t_{kl}\right)+\lambda g_{mnl}\right), \tag{S9}$$

$$F''g_{mnl}=\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)^2+\lambda\right). \tag{S10}$$

C. Proof of Proposition 1

With (16), $G(x, x) = Fg_{mnl}(x)$ holds. Next, our goal is to demonstrate $G(x, g_{mnl}^{(t)}) \geq Fg_{mnl}(x)$. To accomplish this, we will commence by deriving the quadratic approximation to $Fg_{mnl}(x)$ at $g(t)$ mnl.

$$\begin{aligned}&Fg_{mnl}\left(x\right)\\&=Fg_{mnl}\left(g_{mnl}^{(t)}\right)+F'g_{mnl}\left(g_{mnl}^{(t)}\right)\left(x-g_{mnl}^{(t)}\right)\\&+\frac{1}{2}F''g_{mnl}\left(g_{mnl}^{(t)}\right)\left(x-g_{mnl}^{(t)}\right)^2.\end{aligned} \tag{S11}$$

By combining (S10)-(S11) and Proposition 1, we find that $G(x, g_{mnl}^{(t)})$ is an auxiliary function of $Fg_{mnl}(x)$ if the following inequality holds:

$$\frac{\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\hat{y}_{ijk}+\lambda g_{mnl}^{(t)}\right)}{g_{mnl}^{(t)}}\geq\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)^2+\lambda\right). \tag{S12}$$

Note that we have $y_{ijk} \geqslant 0$ according to **Y**'s nonnegativity, and LFs ≥ 0 with SLF-NMUT. So (S12) is equal to

$$\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\hat{y}_{ijk}\right)\geq g_{mnl}^{(t)}\sum_{y_{ijk}\in\Lambda}\left(p_{im}q_{jn}t_{kl}\right)^2. \tag{S13}$$

Next, we reconfigure the left term of (S13) as follows:

$$
\begin{aligned}
&\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\hat{y}_{ijk}\right)=\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\left(g_{mnl}^{(t)}p_{im}q_{jn}t_{kl}+\hat{y}_{ij\tilde{k}}\right)\right)\\
&=g_{mnl}^{(t)}\sum_{y_{ijk}\in\Lambda}\left(p_{im}q_{jn}t_{kl}\right)^2+\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\hat{y}_{ij\tilde{k}}\right)\geq g_{mnl}^{(t)}\sum_{y_{ijk}\in\Lambda}\left(p_{im}q_{jn}t_{kl}\right)^2
\end{aligned}
\tag{S14}
$$

Note that (S12) holds with (S14), making $G(x, g_{mnl}^{(t)})$ be an auxiliary function of Fg_{mnl}.

D. Proof of Theorem 1

Based on (15), (16), and (S9), we obtain

$$
\begin{aligned}
&g_{mnl}^{(t+1)}=\arg\min_x G\left(x,g_{mnl}^{(t)}\right)\\
&\Rightarrow F'g_{mnl}\left(g_{mnl}^{(t)}\right)+\frac{\sum_{y_{ijk}\in\Lambda}\left(\left(p_{im}q_{jn}t_{kl}\right)\hat{y}_{ijk}+\lambda g_{mnl}^{(t)}\right)}{g_{mnl}^{(t)}}\times\left(x-g_{mnl}^{(t)}\right)=0\\
&\Rightarrow g_{mnl}^{(t+1)}\leftarrow g_{mnl}^{(t)}\frac{\sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right)}{\sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}+\lambda g_{mnl}\right)}.
\end{aligned}
\tag{S15}
$$

Based on (S15), it is evident that Fg_{mnl} exhibits a nonincreasing behavior with respect to (13). Naturally, (S15) holds $\forall i\in I$, $j\in J$, $k\in K$, m, n, $l\in\{1, 2, \ldots, R\}$. Hence, Theorem 1 holds.

E. Proof of Theorem 2

Based on (19), a sequence converges with (S1). Let $\mathbf{G}^{(*)}$ represent a stationary point of $\mathbf{G}$, that is:

$$
0\leq g_{mnl}^{(*)}=\lim_{t\to\infty}g_{mnl}^{(t)}<+\infty,\forall i\in I,m,n,l\in\{1,2,...,R\}.
\tag{S16}
$$

As a consequence, if $\mathbf{G}^{(*)}$ serves as one of the equilibrium points of (10), the following KKT conditions pertaining to $\mathbf{G}$ must be satisfied:

$$
\begin{aligned}
&\forall i\in I, j\in J,k\in K,m,n,l\in\{1,\ 2,\ldots,R\}:\\
&(\mathrm{a})\left.\frac{\partial L}{\partial g_{mnl}}\right|_{g_{mnl}=g_{mnl}^{(*)}}=\sum_{y_{ijk}\in\Lambda}\left(\lambda g_{mnl}^{(*)}-\left(y_{ijk}-\hat{y}_{ijk}\right)p_{im}q_{jn}t_{kl}\right)-\phi_{mnl}^{(*)}=0;\\
&(\mathrm{b})\phi_{mnl}^{(*)}\cdot g_{mnl}^{(*)}=0;\quad(\mathrm{c})g_{mnl}^{(*)}\geq 0;\quad(\mathrm{d})\phi_{mnl}^{(*)}\geq 0.
\end{aligned}
\tag{S17}
$$

Note that following (S1)-(S5), Condition (a) is inherently met when we consider (13). Consequently, we can assert that:

$$
\phi_{mnl}^{(*)}=\sum_{y_{ijk}\in\Lambda}\left(\lambda g_{mnl}^{(*)}-\left(y_{ijk}-\hat{y}_{ijk}\right)p_{im}q_{jn}t_{kl}\right).
\tag{S18}
$$

Hence, our primary focus is on Conditions (c) and (d). We begin by constructing θ (t) mnl

$$
\theta_{mnl}^{(t)}=\frac{\sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right)}{\sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}+\lambda g_{mnl}\right)}.
\tag{S19}
$$

Clearly, (S19) is bounded by nonnegative LFs and y_{ijk}.

$$
0\leq\theta_{mnl}^{(*)}=\lim_{t\to+\infty}\theta_{mnl}^{(t)}=\frac{\sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right)}{\sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}+\lambda g_{mnl}\right)}.
\tag{S20}
$$

Therefore, the update rule for g_{mnl} can be reformulated using the SLF-NMUT:

$$g_{mnl}^{(t+1)} = g_{mnl}^{(t)}\theta_{mnl}^{(t)}. \tag{S21}$$

By combining (19) and (S21), we have

$$\lim_{t\to+\infty}\left|g_{mnl}^{(t+1)} - g_{mnl}^{(t)}\right| = 0 \Rightarrow g_{mnl}^{(*)}\theta_{mnl}^{(*)} - g_{mnl}^{(*)} = 0. \tag{S22}$$

Note that in accordance with the update rule (S5), $g_{mnl}^{(*)}$ is greater than or equal to zero given a nonnegative initial hypothesis. Consequently, we make the following inferences:

1) When $g_{mnl}^{(*)} > 0$. Based on (S19) and (S22), we have

$$g_{mnl}^{(*)}\theta_{mnl}^{(*)} - g_{mnl}^{(*)} = 0, g_{mnl}^{(*)} > 0 \Rightarrow \theta_{mnl}^{(*)} = 1 \Rightarrow \lambda|\Lambda|g_{mnl}^{(*)} + \sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}\right) - \sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right) = 0. \tag{S23}$$

Combining (S18) and (S23), we obtain Condition (b) in (S17)

$$\phi_{mnl}^{(*)} = \lambda|\Lambda|g_{mnl}^{(*)} + \sum\nolimits_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}\right) - \sum\nolimits_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right) = 0 \Rightarrow \phi_{mnl}^{(*)}\cdot g_{mnl}^{(*)} = 0, \tag{S24}$$

when $\Phi_{mnl}^{(*)}$ and $g_{mnl}^{(*)} > 0$, conditions (c) and (d) are naturally satisfied. Consequently, (S17) is fulfilled when $g_{mnl}^{(*)} > 0$.

2) When $g_{mnl}^{(*)} = 0$. It's worth noting that conditions (b) and (c) in (S17) are naturally met. Therefore, we aim to substantiate condition (d). To accomplish this, we reconfigure $g_{mnl}^{(*)}$ into

$$g_{mnl}^{(*)} = g_{mnl}^{(0)}\lim_{t\to+\infty}\prod_{r=1}^{t}\theta_{mnl}^{(r)}. \tag{S25}$$

Based on (S25), we can deduce the following inferences:

$$\begin{aligned}
& g_{mnl}^{(0)} > 0, g_{mnl}^{(0)}\lim_{t\to+\infty}\prod_{r=1}^{t}\theta_{mnl}^{(r)} = g_{mnl}^{(*)} = 0 \\
& \Rightarrow \lim_{t\to+\infty}\prod_{r=1}^{t}\theta_{mnl}^{(r)} = 0 \\
& \Rightarrow \lim_{t\to+\infty}\theta_{mnl}^{(t)} = \theta_{mnl}^{(*)} = \frac{\sum\nolimits_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right)}{\sum\nolimits_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}\right) + \lambda|\Lambda|g_{mnl}^{(*)}} \le 1 \\
& \Rightarrow \phi_{mnl}^{(*)} = \lambda|\Lambda|g_{mnl}^{(*)} + \sum_{y_{ijk}\in\Lambda}\left(\hat{y}_{ijk}p_{im}q_{jn}t_{kl}\right) - \sum_{y_{ijk}\in\Lambda}\left(y_{ijk}p_{im}q_{jn}t_{kl}\right) \ge 0.
\end{aligned} \tag{S26}$$

Hence, it follows that (S17) is satisfied even when $g_{mnl}^{(*)} > 0$. Similarly, we can demonstrate that the sequences $\{p_{im}^{(t)}\}$, $\{q_{jn}^{(t)}\}$, $\{t_{kl}^{(t)}\}$, $\{h_{mnl}^{(t)}\}$, $\{d_{im}^{(t)}\}$, $\{e_{jn}^{(t)}\}$, $\{f_{kl}^{(t)}\}$ also converge to a stable equilibrium point of (10). As a result, Theorem 2 is validated.

IV. Supplementary Procedure

These are the procedures **Update_G**, **Update**_P, **Update_H** and **Update**_D.

Procedure Update_G

Operation	Cost
Initialize $up\sim\mathbf{G}^{R_1\times R_2\times R_3}$, $down\sim\mathbf{G}^{R_1\times R_2\times R_3} = 0$	$\Theta(R_1\times R_2\times R_3)$
for each $\forall y_{ijk}\in\Lambda$	$\times\|\Lambda\|$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$up\sim g_{mnl} = y\cdot p_{im}\cdot q_{jn}\cdot t_{kl}$	$\Theta(1)$
$down\sim g_{mnl}$ += $\tilde{y}\cdot p_{im}\cdot q_{jn}\cdot t_{kl}+\lambda\cdot p_{im}$	$\Theta(1)$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$g_{mnl} = g_{mnl}\cdot up\sim g_{mnl}\,/\,down\sim g_{mnl}$	$\Theta(1)$

Procedure Update_P

Operation	**Cost**
Initialize $up\sim \mathrm{P}^{\lvert I\rvert\times R1}$, $down \sim \mathrm{P}^{\lvert I\rvert\times R1} = 0$	$\Theta(\lvert I\rvert\times R_1)$
for each $\forall y_{ijk}\in\Lambda$	$\times\lvert\Lambda\rvert$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$up\sim p_{im}$+= $y\cdot g_{mnl}\cdot q_{jn}\cdot t_{kl}$	$\Theta(1)$
$down\sim p_{im}$+= $\tilde{y}\cdot g_{mnl}\cdot q_{jn}\cdot t_{kl}$ +$\lambda\cdot p_{im}$	$\Theta(1)$
for i=1 to $\lvert I\rvert$ **do**	$\times\lvert I\rvert$
for m=1 to R_1 **do**	$\times R_1$
$p_{im} = p_{im}\cdot up\sim p_{im} / down\sim p_{im}$	$\Theta(1)$

Procedure Update_H

Operation	**Cost**
Initialize $up\sim\mathbf{H}^{R_1\times R_2\times R_3}$, $down\sim\mathbf{H}^{R_1\times R_2\times R_3} = 0$	$\Theta(R_1\times R_2\times R_3)$
for each $\forall y_{ijk}\in\Lambda$	$\times\lvert\Lambda\rvert$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$up\sim h_{mnl} = y\cdot(p_{im}+q_{jn}+t_{kl})$	$\Theta(1)$
$down\sim h_{mnl}$+= $\tilde{y}\cdot(p_{im}+q_{jn}+t_{kl})$ +$\lambda\cdot h_{mnl}$	$\Theta(1)$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$h_{mnl} = h_{mnl}\cdot up\sim h_{mnl} / down\sim h_{mnl}$	$\Theta(1)$

Procedure Update_D

Operation	**Cost**
Initialize $up\sim \mathrm{D}^{\lvert I\rvert\times R1}$, $down \sim \mathrm{D}^{\lvert I\rvert\times R1} = 0$	$\Theta(\lvert I\rvert\times R_1)$
for each $\forall y_{ijk}\in\Lambda$	$\times\lvert\Lambda\rvert$
for m=1 to R_1 **do**	$\times R_1$
for n=1 to R_2 **do**	$\times R_2$
for l=1 to R_3 **do**	$\times R_3$
$up\sim d_{im}$+= $y\cdot h_{mnl}$	$\Theta(1)$
$down\sim d_{im}$+= $\tilde{y}\cdot h_{mnl}\cdot$ +$\lambda\cdot d_{im}$	$\Theta(1)$
for i=1 to $\lvert I\rvert$ **do**	$\times\lvert I\rvert$
for m=1 to R_1 **do**	$\times R_1$
$d_{im} = d_{im}\cdot up\sim d_{im} / down\sim d_{im}$	$\Theta(1)$