Symmetry and Graph Bi-regularized Non-negative Matrix Factorization for Precise Community Detection: Supplementary File

Zhigang Liu, Xin Luo, *Senior Member, IEEE*, and MengChu Zhou, *Fellow, IEEE*

I. Introduction

THIS is the supplementary file for the paper entitled "*Symmetry and Graph Bi-regularized Non-negative Matrix Factorization for Precise Community Detection*". We have put the *Symbol Table* in Section II, as well as the proofs of *Theorems* 1, 2 related to Y in Section III.

II. Symbol Appointment

Table I summarizes important symbols of this paper.

Table I. Symbol Appointment

Symbol	Description
G	Target graph describing an LUN.
V, E	Node set and edge set of the target network.
v_i, e_{ij}	Single node and edge in V and E, respectively.
n, m	Number of nodes in V and edges in E, respectively.
C_k	Community k.
$A^{n\times n}$, $\hat{A}^{n\times n}$	Adjacency matrix and its low-rank approximation.
W	The weight matrix related to A.
D	The diagonal matrix of A as $D_{ii}=\sum_l W_{il}$.
L	The Laplacian matrix of A.
$X^{n\times K}$, $Y^{n\times K}$	The latent factor matrices of A.
x_{ik}, y_{jk}	Single entries in X and Y.
$x_{ik}^{(t)}$, $y_{jk}^{(t)}$	The state values of x_{ik} and y_{jk} in the t-th learning iteration.
Ψ, Φ	The Lagrangian multiplier matrices.
ψ_{ik}, ϕ_{jk}	Single entries in Ψ and Φ.
α, λ	The symmetry and graph regularization coefficients.
K	The number of communities in G.
X^{T}	The transpose of a matrix X.
I	An identity matrix.
$\|\|\cdot\|\|_F$	The Frobenius norm of a matrix
$\mathrm{Tr}(\cdot)$	The trace of a matrix.

III. Proofs of Theorems 1 and 2 Related to Y

A. The Proof of Theorem 1 Related to Y

Considering an arbitrary element y_{jk} in Y, we adopt $F(y_{jk})$ to denote the corresponding component from $O_{B\text{-}NMF}$, that is relevant to y_{jk} only. Similarly, for the second term of (18), we let $u=j$ without loss of generality. Thus, the following function is achieved:

$$F\left(y_{jk}\right)=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\sum_{k=1}^{K}x_{ik}y_{jk}-a_{ij}\right)^2+\frac{\alpha}{2}\sum_{j=1}^{N}\sum_{k=1}^{K}\left(x_{jk}-y_{jk}\right)^2+\frac{\lambda}{2}\sum_{j=1}^{N}\sum_{s=1}^{N}\sum_{k=1}^{K}y_{jk}L_{js}y_{sk}. \tag{S1}$$

The first and second-order derivatives of F with y_{jk} in (S1) is formulated as:

$$F'\left(y_{jk}\right)=\left(YX^{\mathrm{T}}X-A^{\mathrm{T}}X+\lambda LY-\alpha X+\alpha Y\right)_{jk}, \tag{S2}$$

$$F''\left(y_{jk}\right)=\left(X^{\mathrm{T}}X\right)_{kk}+\left(\lambda L+\alpha I\right)_{jj}. \tag{S3}$$

Based on (S1)-(S3), we are able to build an auxiliary function $Z(y_{jk}, y_{jk}^{(t)})$ to $F(y_{jk})$, as presented in *Lemma* S1.

***Lemma* S1.** The following function is an auxiliary function for $F(y_{jk})$.

$$Z\left(y_{jk}, y_{jk}^{(t)}\right) = F\left(y_{jk}^{(t)}\right) + F'\left(y_{jk}^{(t)}\right)\left(y_{jk} - y_{jk}^{(t)}\right) + \frac{1}{2}\left(\left(\left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} + \alpha y_{jk}^{(t)}\right) \Big/ y_{jk}^{(t)}\right)\left(y_{jk} - y_{jk}^{(t)}\right)^2, \tag{S4}$$

Proof. Note that $Z(y_{jk}, y_{jk})=F(y_{jk})$ evidently holds. Hence, we only need to prove that $Z(y_{jk}, y_{jk}^{(t)})\geq F(y_{jk})$. To achieve this objective, we expend $F(y_{jk})$ to its second-order Taylor series at the state point $y_{jk}^{(t)}$ as:

$$F\left(y_{jk}\right) = F\left(y_{jk}^{(t)}\right) + F'\left(y_{jk}^{(t)}\right)\left(y_{jk} - y_{jk}^{(t)}\right) + \frac{1}{2}\left(\left(X^{\mathrm{T}}X\right)_{kk} + \left(\lambda L + \alpha I\right)_{jj}\right)\left(y_{jk} - y_{jk}^{(t)}\right)^2. \tag{S5}$$

By combining (S4) and (S5), we see that the desired condition of $Z(y_{jk}, y_{jk}^{(t)})\geq F(y_{jk})$ is equivalent to the following condition:

$$\left(\left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} + \alpha y_{jk}^{(t)}\right) \Big/ y_{jk}^{(t)} \geq \left(X^{\mathrm{T}}X\right)_{kk} + \left(\lambda L + \alpha I\right)_{jj}. \tag{S6}$$

Since $y_{jk}^{(t)}\geq 0$, (S6) can be reduced into the following form:

$$\begin{aligned} &\left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} + \alpha y_{jk}^{(t)} \geq y_{jk}^{(t)}\left(X^{\mathrm{T}}X\right)_{kk} + \left(\lambda L + \alpha I\right)_{jj} y_{jk}^{(t)} \\ &\Leftrightarrow \left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} \geq y_{jk}^{(t)}\left(X^{\mathrm{T}}X\right)_{kk} + \lambda L_{jj} y_{jk}^{(t)}. \end{aligned} \tag{S7}$$

To achieve (S6), we have the following inferences:

$$\left(YX^{\mathrm{T}}X\right)_{jk} = \sum_{l=1}^{K} y_{jl}^{(t)}\left(X^{\mathrm{T}}X\right)_{lk} = \sum_{l=1, l\neq k}^{K} y_{jl}^{(t)}\left(X^{\mathrm{T}}X\right)_{lk} + y_{jk}^{(t)}\left(X^{\mathrm{T}}X\right)_{kk} \geq y_{jk}^{(t)}\left(X^{\mathrm{T}}X\right)_{kk}, \tag{S8}$$

and

$$\lambda\left(DY\right)_{jk} = \lambda\sum_{v=1}^{N} D_{jv} y_{vk}^{(t)} = \lambda\left(\sum_{v=1, v\neq j}^{N} D_{jv} y_{vk}^{(t)} + D_{jj} y_{jk}^{(t)}\right) \geq \lambda D_{jj} y_{jk}^{(t)} \geq \lambda\left(D - A\right)_{jj} y_{jk}^{(t)} = \lambda L_{jj} y_{jk}^{(t)}. \tag{S9}$$

By combining (S7)-(S9), (S6) holds, making $Z(y_{jk}, y_{jk}^{(t)})$ be an auxiliary function of $F(y_{jk})$. □

Based on the above inferences, we present the following proof of *Theorem* 1 regarding to Y.

Proof of *Theorem* 1. By replacing $Z(y_{jk}, y_{jk}^{(t)})$ in (17) in the manuscript with (S4), we achieve the following learning rule:

$$\begin{aligned} &y_{jk}^{(t+1)} = \arg\min_{x} H\left(y_{jk}, y_{jk}^{(t)}\right) \Rightarrow \\ &F'\left(y_{jk}^{(t)}\right) + \frac{\left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} + \alpha y_{jk}^{(t)}}{y_{jk}^{(t)}}\left(y_{jk} - y_{jk}^{(t)}\right) = 0 \\ &\Rightarrow y_{jk}^{(t+1)} = y_{jk}^{(t)} - y_{jk}^{(t)} \frac{F'\left(y_{jk}^{(t)}\right)}{\left(YX^{\mathrm{T}}X\right)_{jk} + \lambda\left(DY\right)_{jk} + \alpha y_{jk}^{(t)}} \\ &= y_{jk}^{(t)} \frac{\left(A^{\mathrm{T}}X + \alpha X\right)_{ik}}{\left(YX^{\mathrm{T}}X + \lambda LY + \alpha Y\right)_{jk}}. \end{aligned} \tag{S10}$$

(27) in the manuscript and (S10) indicate that the learning scheme of B-NMF given in (15b) in the manuscript can be denoted by the auxiliary function-based learning scheme (17) in the manuscript with the auxiliary function (S4). Note that y_{jk} is an arbitrary entry of Y. Hence, $\forall j\in\{1, 2, \ldots, n\}$, $k\in\{1, 2, \ldots, K\}$, $F(y_{jk})$ is non-increasing with the learning scheme (15b) as X alternatively fixed. Besides, as proven in Section III.B, we have achieved that $\forall i\in\{1, 2, \ldots, n\}$, $k\in\{1, 2, \ldots, K\}$, $F(x_{ik})$ is non-increasing with the learning scheme (15a) as Y alternatively fixed. Thus, based on the inferences above, we conclude that $O_{B\text{-}NMF}$ is non-increasing with B-NMF's learning scheme (15). Hence, *Theorem* 1 stands. ■

B. The Proof of Theorem 2 Related to Y

To achieve this objective, we make $y_{jk}^{(*)} = \lim_{t\to+\infty} y_{jk}^{(t)}$. If $y_{jk}^{(*)}$ is one of $O_{B\text{-}NMF}$'s KKT stationary points, the following conditions are desired:

$$\begin{aligned} &\forall j \in \{1, \ldots, N\}, k \in \{1, \ldots, K\}: \\ &(\mathrm{a})\ \left.\frac{\partial L}{\partial y_{jk}}\right|_{y_{jk}=y_{jk}^{(*)}} = \left(Y^{(*)}X^{\mathrm{T}}X - A^{\mathrm{T}}X - \alpha X + \alpha Y^{(*)} + \lambda LY^{(*)} + \Phi^{(*)}\right)_{jk} = 0, \\ &(\mathrm{b})\ \phi_{jk}^{(*)} y_{jk}^{(*)} = 0, \\ &(\mathrm{c})\ y_{jk}^{(*)} \geq 0, \\ &(\mathrm{d})\ \phi_{jk}^{(*)} \geq 0. \end{aligned} \tag{S11}$$

Note that based on (8)-(15) in the manuscript, Condition (a) is constantly satisfied with (15b) in the manuscript. Thus, the following equation is achieved:

$$\phi_{jk}^{(*)} = -\left(Y^{(*)}X^{\mathrm{T}}X - A^{\mathrm{T}}X - \alpha X + \alpha Y^{(*)} + \lambda LY^{(*)}\right)_{jk}. \tag{S12}$$

Thus, to perform the analyses with conditions (b)-(d), we build the following auxiliary variable:

$$\sigma_{jk}^{(t)} = \frac{\left(A^{\mathrm{T}}X + \alpha X\right)_{jk}}{\left(Y^{(t)}X^{\mathrm{T}}X + \lambda LY^{(t)} + \alpha Y^{(t)}\right)_{jk}}. \tag{S13}$$

Note that we have A, X, $Y \geq 0$ in (9) in the manuscript, yielding the following bound of $\sigma_{jk}^{(t)}$:

$$0 \leq \sigma_{jk}^{(*)} = \lim_{t\to+\infty}\sigma_{jk}^{(t)} = \frac{\left(A^{\mathrm{T}}X + \alpha X\right)_{jk}}{\left(Y^{(t)}X^{\mathrm{T}}X + \lambda LY^{(t)} + \alpha Y^{(t)}\right)_{jk}} < \infty. \tag{S14}$$

With (S13), the learning rule for Y can be simplified as:

$$y_{jk}^{(t+1)} = y_{jk}^{(t)}\frac{\left(A^{\mathrm{T}}X + \alpha X\right)_{jk}}{\left(Y^{(t)}X^{\mathrm{T}}X + \lambda LY^{(t)} + \alpha Y^{(t)}\right)_{jk}} = y_{jk}^{(t)}\sigma_{jk}^{(t)}. \tag{S15}$$

By combining (30) in the manuscript and (S13)-(S15), we achieve:

$$\lim_{t\to+\infty}\left|y_{jk}^{(t+1)} - y_{jk}^{(t)}\right| = 0 \Rightarrow y_{jk}^{(*)}\sigma_{jk}^{(*)} - y_{jk}^{(*)} = 0. \tag{S16}$$

Note that $y_{jk}^{(*)} \geq 0$ with the learning scheme (15b) in the manuscript. Thus, we discuss the following cases:

Case a) $\boldsymbol{y_{jk}^{(*)} > 0}$**.** Based on (S14) and (S16), we have:

$$\begin{aligned} & y_{jk}^{(*)}\sigma_{jk}^{(*)} - y_{jk}^{(*)} = 0, y_{jk}^{(*)} > 0 \Rightarrow \sigma_{jk}^{(*)} = 1 \\ & \Rightarrow \left(A^{\mathrm{T}}X + \alpha X\right)_{jk} - \left(Y^{(*)}X^{\mathrm{T}}X + \lambda LY^{(*)} + \alpha Y^{(*)}\right)_{jk} = 0 \\ & \Rightarrow \phi_{jk}^{(*)}y_{jk}^{(*)} = 0, \end{aligned} \tag{S17}$$

which indicates that condition (b) in (S11) is fulfilled. Moreover, based on (S17), we further have $y_{jk}^{(*)}>0$ and $\phi_{jk}^{(*)}=0$, which fulfill conditions (c) and (d) in (S11). Hence, when $y_{jk}^{(*)}>0$, KKT conditions in (S11) are all fulfilled.

Case b) $\boldsymbol{y_{jk}^{(*)} = 0}$**.** Note that in this case, conditions (b) and (c) in (S11) are constantly satisfied. Thus, we aim to prove that condition (d) in (S11) is fulfilled. To do this, we reformulate $y_{jk}^{(*)}$ as:

$$y_{jk}^{(*)} = y_{jk}^{(0)}\lim_{t\to+\infty}\prod_{r=1}^{t}\sigma_{jk}^{(r)}. \tag{S18}$$

Note that if $y_{jk}^{(0)}=0$, it keeps zero constantly with the learning scheme (15b) in the manuscript, which leads to poor representation to A. Hence, without loss of generality, we make $y_{jk}^{(0)}>0$. Thus, we have the following inferences:

$$y_{jk}^{(*)} = y_{jk}^{(0)}\lim_{t\to+\infty}\prod_{r=1}^{t}\sigma_{jk}^{(r)} = 0 \Rightarrow \lim_{t\to+\infty}\prod_{r=1}^{t}\sigma_{jk}^{(r)} = 0. \tag{S19}$$

On the other hand, part A's entries should be one to describe a meaningful LUN, making $(AX+\alpha X)_{jk}>0$ with $\alpha>0$. Thus, with (S19), we have the following inferences:

$$\begin{aligned} & \lim_{t\to+\infty}\prod_{r=1}^{t}\tau_{ik}^{(r)} = 0 \Rightarrow \lim_{t\to+\infty}\tau_{ik}^{(t)} = \tau_{ik}^{(*)} < 1 \\ & \Rightarrow \left(X^{(*)}Y^{\mathrm{T}}Y + \alpha X^{(*)}\right)_{ik} - \left(AY + \alpha Y\right)_{ik} = \psi_{ik}^{(*)} > 0, \end{aligned} \tag{S20}$$

which indicates that condition (d) in (S11) is fulfilled. Hence, we have proven that KKT conditions in (S11) are all fulfilled when $y_{jk}^{(*)}=0$. ■