DNA: Dynamic Social Network Alignment Appendix

I. Full Proof of Lemma (Auxiliary Function)

We give the proof in details on the Lemma (Auxiliary Function), especially for the bounds. Moreover, we present more insights in the derivation for the multiplicative updating rule of $\mathbf{V}^{(\cdot)}$, one of the key updating rules in Algo. 1, while proving this Lemma. We utilize $\mathbf{V}^s$ for demonstration in accordance with Sec 4.2. Note that, the derivation for $\mathbf{V}^s$ is the same as that of $\mathbf{V}^s$ and thus omitted for clarity.

Proof. First, we prove the correctness of the auxiliary function of the nonnegative objective function $\mathcal{J}$ w.r.t. $\mathbf{V}^s$.

According to the definition of the auxiliary function, a function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ is the auxiliary function of objective function $\mathcal{J}$ *w.r.t.* $\mathbf{V}^s$, $\mathcal{J}(\mathbf{V}^s)$, of its nonnegative form *in Eq. (23) of our DNA paper* if and only if

$$Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s) \geq \mathcal{J}(\mathbf{V}^s),$$

$$Z(\mathbf{V}^s, \mathbf{V}^s) = \mathcal{J}(\mathbf{V}^s).$$

Thus, to obtain the auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$, we need to figure out upper bounds of positive terms and lower bounds of negative terms in the nonnegative form of $\mathcal{J}(\mathbf{V}^s)$. Now, we elaborate the derivation of these bounds.

The inequality $2ab \leq a^2 + b^2$ always holds for $\forall a, b > 0$. Since $\mathbf{V}^s$ is nonnegative, we obtain the following inequality:

$$[\mathbf{V}^s]_{ij} \leq \frac{[\mathbf{V}^s]_{ij}^2 + [\tilde{\mathbf{V}}^s]_{ij}^2}{2[\tilde{\mathbf{V}}^s]_{ij}} \tag{1}$$

Utilizing the fact $tr(\mathbf{X}^T\mathbf{Y}) = \sum_{ij}[\mathbf{X}]_{ij}[\mathbf{Y}]_{ij}$ for any $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times m}$, we obtain the upper bound as follows:

$$tr((\mathbf{U}^s\mathbf{Q}^{sT})^-\mathbf{V}^{sT}) \leq \sum_{ij}[\boldsymbol{\Upsilon}]_{ij}\frac{[\mathbf{V}^s]_{ij}^2 + [\tilde{\mathbf{V}}^s]_{ij}}{2[\tilde{\mathbf{V}}^s]_{ij}}, \tag{2}$$

where

$$\boldsymbol{\Upsilon} = (\mathbf{U}^s\mathbf{Q}^{sT})^-. \tag{3}$$

For any nonnegative matrix $\mathbf{A} \in \mathbb{R}_+^{n\times n}$, $\mathbf{B} \in \mathbb{R}_+^{n\times n}$, $\mathbf{V} \in \mathbb{R}_+^{n\times k}$ and $\tilde{\mathbf{V}} \in \mathbb{R}_+^{n\times k}$, the following inequality[1] always holds:

$$\sum_{ij}\frac{[\mathbf{A}\tilde{\mathbf{V}}\mathbf{B}]_{ij}[\mathbf{V}]_{ij}^2}{[\tilde{\mathbf{V}}]_{ij}} \geq tr(\mathbf{V}^T\mathbf{A}\mathbf{V}\mathbf{B}), \tag{4}$$

under the symmetric constraints, *i.e.*, $[\mathbf{A}]_{ij} = [\mathbf{A}]_{ji}$ and $[\mathbf{B}]_{ij} = [\mathbf{B}]_{ji}$. The equality exists if and only if $\mathbf{V} = \tilde{\mathbf{V}}$. Thus, we obtain the upper bounds as follows:

$$\begin{aligned} tr((\mathbf{P}^{st}\mathbf{V}^s)^T\mathbf{P}^{st}\mathbf{V}^s) &\leq \sum_{ij}\frac{[\boldsymbol{\Pi}^s]_{ij}[\mathbf{V}^s]_{ij}^2}{[\tilde{\mathbf{V}}^s]_{ij}}, \\ tr(\mathbf{V}^s(\mathbf{Q}^s\mathbf{Q}^{sT})^+\mathbf{V}^{sT}) &\leq \sum_{ij}\frac{[\tilde{\mathbf{V}}^s\boldsymbol{\Phi}^s]_{ij}[\mathbf{V}^s]_{ij}^2}{[\tilde{\mathbf{V}}^s]_{ij}}, \end{aligned} \tag{5}$$

where

$$\begin{aligned} \boldsymbol{\Pi} &= \mathbf{P}^{st^T}\mathbf{P}^{st}\mathbf{V}^s, \\ \boldsymbol{\Phi} &= (\mathbf{Q}^s\mathbf{Q}^{sT})^+. \end{aligned} \tag{6}$$

Note that, the positive semidefinite matrix $\mathbf{X}^T\mathbf{X}$ for any $\mathbf{X} \in \mathbb{R}^{n\times m}$ is always Hermitian.

The inequality $z - 1 \leq \log z$ always holds for $\forall z > 0$. Thus, we obtain the following inequality owing to the nonnegativity of matrix $\mathbf{V}^s$:

$$\begin{aligned} \frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}} &\geq 1 + \log\frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}, \\ \frac{[\tilde{\mathbf{V}}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ik}}{[\tilde{\mathbf{V}}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ik}} &\geq 1 + \log\frac{[\mathbf{V}^s]_{ij}[\mathbf{V}^s]_{ik}}{[\tilde{\mathbf{V}}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ik}}. \end{aligned} \tag{7}$$

Thus, we obtain the lower bounds as follows

$$\begin{aligned} tr((\mathbf{U}^s\mathbf{Q}^{sT})^+\mathbf{V}^{sT}) &\geq \sum_{ij}[\boldsymbol{\Psi}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ij}\left(1 + \log\frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}\right), \\ tr((\mathbf{P}^{ts}\mathbf{V}^{(j)})^T\mathbf{P}^{st}\mathbf{V}^s) &\geq \sum_{ij}[\boldsymbol{\Lambda}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ij}\left(1 + \log\frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}\right), \\ tr(\mathbf{V}^s(\mathbf{Q}^s\mathbf{Q}^{sT})^-\mathbf{V}^{sT}) &\geq \sum_{ij}[\boldsymbol{\Gamma}^s]_{jk}[\tilde{\mathbf{V}}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ik} \\ &\qquad \left(1 + \log\frac{[\mathbf{V}^s]_{ij}[\mathbf{V}^s]_{ik}}{[\tilde{\mathbf{V}}^s]_{ij}[\tilde{\mathbf{V}}^s]_{ik}}\right), \end{aligned} \tag{8}$$

where

$$\begin{aligned} \boldsymbol{\Lambda} &= \mathbf{P}^{st^T}\mathbf{P}^{ts}\mathbf{V}^{(j)}, \\ \boldsymbol{\Psi} &= (\mathbf{U}^s\mathbf{Q}^{sT})^+, \\ \boldsymbol{\Gamma} &= (\mathbf{Q}^s\mathbf{Q}^{sT})^-. \end{aligned} \tag{9}$$

Replacing the positive (negative) terms *in Eq. (23) of our DNA paper* with the corresponding upper (lower) bounds

[1] This inequality is proved in the study [Ding et. al., Convex and Semi-Nonnegative Matrix Factorizations, TPAMI 2010, vol. 32, no. 1, pp. 45-55]

above, we obtain the auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ in the Lemma (Auxiliary Function) as follows:

$$
\begin{aligned}
Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s) = & -2\beta \sum_{ij} [\mathbf{\Psi}^s]_{ij} [\tilde{\mathbf{V}}^s]_{ij} \left(1 + \log \frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}\right) \\
& + \beta \sum_{ij} [\mathbf{\Upsilon}^s]_{ij} \frac{[\mathbf{V}^s]_{ij}^2 + [\tilde{\mathbf{V}}^s]_{ij}^2}{[\tilde{\mathbf{V}}^s]_{ij}} \\
& + \beta \sum_{ij} \frac{[\tilde{\mathbf{V}}^s \mathbf{\Phi}^s]_{ij} [\mathbf{V}^s]_{ij}^2}{[\tilde{\mathbf{V}}^s]_{ij}} \\
& - \beta \sum_{ij} [\mathbf{\Gamma}^s]_{jk} [\tilde{\mathbf{V}}^s]_{ij} [\tilde{\mathbf{V}}^s]_{ik} \left(1 + \log \frac{[\mathbf{V}^s]_{ij} [\mathbf{V}^s]_{ik}}{[\tilde{\mathbf{V}}^s]_{ij} [\tilde{\mathbf{V}}^s]_{ik}}\right) \\
& - 2\gamma \sum_{ij} [\mathbf{\Lambda}^s]_{ij} [\tilde{\mathbf{V}}^s]_{ij} \left(1 + \log \frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}\right) \\
& + \gamma \sum_{ij} \frac{[\mathbf{\Pi}^s]_{ij} [\mathbf{V}^s]_{ij}^2}{[\tilde{\mathbf{V}}^s]_{ij}}
\end{aligned} \tag{10}
$$

Next, we prove the convexity of the constructed auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ w.r.t. $\mathbf{V}^s$ as well as the existence of the global optimum.

To study the convexity of $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ *w.r.t.* $\mathbf{V}^s$, we first drive the matrix of partial differential. It is a matrix of dimension $N^s \times d_i$ and its element is given as follows:

$$
\begin{aligned}
\frac{\partial Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)}{\partial [\mathbf{V}^s]_{ij}} = & -2 \left[\mathbf{\Psi}^s + \tilde{\mathbf{V}}^s \mathbf{\Gamma}^s + \mathbf{\Lambda}^{st}\right]_{ij} \frac{[\tilde{\mathbf{V}}^s]_{ij}}{[\mathbf{V}^s]_{ij}} \\
& + 2 \left[\mathbf{\Upsilon}^s + \tilde{\mathbf{V}}^s \mathbf{\Phi}^s + \mathbf{\Pi}^{st}\right]_{ij} \frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}.
\end{aligned} \tag{11}
$$

Then, we derive the Hessian matrix $\mathbf{H}_{\mathbf{V}}^s$ of $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ *w.r.t.* $\mathbf{V}^s$:

$$
\frac{\partial^2 Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)}{\partial [\mathbf{V}^s]_{ij} \partial [\mathbf{V}^s]_{kl}} = \delta_{xk} \delta_{jk} \mathcal{K}_{ij}. \tag{12}
$$

It is a diagonal matrix of dimension $N^s d_i$ with positive elements,

$$
\begin{aligned}
\mathcal{K}_{ij} = & \, 2 \left[\beta \mathbf{\Psi}^s + \beta \tilde{\mathbf{V}}^s \mathbf{\Gamma}^s + \gamma \mathbf{\Lambda}^{st}\right]_{ij} \frac{[\tilde{\mathbf{V}}^s]_{ij}}{[\mathbf{V}^s]_{ij}^2} \\
& + 2 \left[\beta \mathbf{\Upsilon}^s + \beta \tilde{\mathbf{V}}^s \mathbf{\Phi}^s + \gamma \mathbf{\Pi}^{st}\right]_{ij} \frac{[\mathbf{V}^s]_{ij}}{[\tilde{\mathbf{V}}^s]_{ij}}.
\end{aligned} \tag{13}
$$

That is, $\mathbf{H}_{\mathbf{V}^s}$ is positive semidefinite. Now, we prove that the auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ is convex *w.r.t.* $\mathbf{V}^s$ and thus has the global optimum. □

The derivation of the multiplicative updating rule of $\mathbf{V}^s$ is inherently coupled with the construction of the auxiliary function. The auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ approaches its global optimum $\mathbf{V}^s = \arg\min_{\mathbf{V}^s} Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ when the zero partial differential condition is satisfied, *i.e.*,

$$
\frac{\partial Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)}{\partial [\mathbf{V}^s]_{ik}} = 0 \tag{14}
$$

TABLE I: Main hyperparameters

Hyperparameters	Value
DNA model parameter α_1	0.667
DNA model parameter α_2	0.333
DNA model parameter β	0.5
DNA model parameter γ	10^3
Lerning rate of the employed optimizer	0.001
Dropout rate in both pretraining and training	0.2
Number of pretraining iterations	1000
Number of main loop iterations	900

Let $\mathbf{V}^s$ be the optimal solution of $\min_{\mathbf{V}^s} Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$. Incorporating Eq. (11), we derive the multiplicative updating rule of $\mathbf{V}^s$ with some algebraic operations as follows:

$$
\mathbf{V}^s = \mathbf{V}^s \odot \sqrt{\frac{\beta(\mathbf{\Psi}^s + \mathbf{V}^s \mathbf{\Gamma}^s) + \gamma \mathbf{\Lambda}^s}{\beta(\mathbf{\Upsilon}^s + \mathbf{V}^s \mathbf{\Phi}^s) + \gamma \mathbf{\Pi}^s}}. \tag{15}
$$

With the auxiliary function $Z(\mathbf{V}^s, \tilde{\mathbf{V}}^s)$ *w.r.t.* $\mathbf{V}^s$, we give Theorem (Correctness) and Theorem (Convergence) of the multiplicative updating rule, and finally obtain the Corollary (KKT Convergence).

II. Experimental Settings and Hyperparameters

To enhance the reproducibility, we give more details in the experimental settings and hyperparameters. In both pretraining of the LSTM Autoencoder and training of the proposed DNA framework, we utilize RMSProp optimizer and employ the dropout to alleviate over fitting. It is worth mentioning that the number of parameters, *i.e.*, the weight matrices and bias, is of small-scale in the proposed deep neural architecture regardless of the scale of the input dynamic social network tensors. The empirical setting of hyperparemeters is summarized in Table I. Definitions in details are listed as follows.

1) Model parameters α_1 and α_2 weigh the significance of local and global dynamics, respectively.
2) Model parameter β weigh the significance of the transformation.
3) Model parameter γ is the penalty coefficient of the equality constraint to construct common subspace.
4) The learning rate is the hyperparameter of RMSProp, whose other hyperparameters are the default value in TensorFlow.
5) The dropout rate is the hyperparameter when we employ the dropout trick, which is set to an empirical value.
6) The number of pretrain iterations is the number of times we conduct parameter updating in pretraining (line 2 in Algo. 1) for each dynamic graph tensor.
7) The number of main loop iterations is the number of times we conduct line $5-10$ in Algo. 1 to train the proposed DNA.