SUPPLEMENTARY MATERIAL

Anonymous authors
Paper under double-blind review

1 DERIVATION OF ELBO BOUND

$$
\begin{aligned}
\log(p_\theta(\boldsymbol{A})) &= \log\left(\int \sum_{\boldsymbol{c}} p_\theta(\boldsymbol{Z}, \boldsymbol{c}, \boldsymbol{A}) d\boldsymbol{Z}\right) && (1)\\
&= \log\left(\int \sum_{\boldsymbol{c}} p(\boldsymbol{Z})\, p_\theta(\boldsymbol{c}|\boldsymbol{Z})\, p_\theta(\boldsymbol{A}|\boldsymbol{c}, \boldsymbol{Z}) d\boldsymbol{Z}\right) && (2)\\
&= \log\left(\mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\frac{p(\boldsymbol{Z})\, p_\theta(\boldsymbol{c}|\boldsymbol{Z})\, p_\theta(\boldsymbol{A}|\boldsymbol{c}, \boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{I}) q_\phi(\boldsymbol{c}|\boldsymbol{Z}, \mathcal{I})}\right\}\right) && (3)\\
&\geq \mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\left(\frac{p(\boldsymbol{Z})\, p_\theta(\boldsymbol{c}|\boldsymbol{Z})\, p_\theta(\boldsymbol{A}|\boldsymbol{c}, \boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{I}) q_\phi(\boldsymbol{c}|\boldsymbol{Z}, \mathcal{I})}\right)\right\} && (4)\\
&= \mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\left(\frac{p(\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{I})}\right) + \log\left(\frac{p_\theta(\boldsymbol{c}|\boldsymbol{Z})}{q_\phi(\boldsymbol{c}|\boldsymbol{Z}, \mathcal{I})}\right) + \log\Big(p_\theta(\boldsymbol{A}|\boldsymbol{c}, \boldsymbol{Z})\Big)\right\} && (5)\\
&= \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{I})}\left\{\log\left(\frac{p(\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{I})}\right)\right\}\\
&+ \mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\left(\frac{p_\theta(\boldsymbol{c}|\boldsymbol{Z})}{q_\phi(\boldsymbol{c}|\boldsymbol{Z}, \mathcal{I})}\right)\right\}\\
&+ \mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\Big(p_\theta(\boldsymbol{A}|\boldsymbol{c}, \boldsymbol{Z})\Big)\right\}. && (6)
\end{aligned}
$$

Where (4) follows from Jensen's Inequality. First term of (6) is given by:

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{I})}\left\{\log\left(\frac{p(\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{I})}\right)\right\} &= \sum_{i=1}^{N} \mathbb{E}_{\boldsymbol{z}_i\sim q_\phi(\boldsymbol{z}_i|\mathcal{I})}\left\{\log\left(\frac{p(\boldsymbol{z}_i)}{q_\phi(\boldsymbol{z}_i|\mathcal{I})}\right)\right\} && (7)\\
&= -\sum_{i=1}^{N} D_{KL}(q_\phi(\boldsymbol{z}_i|\mathcal{I}) \,||\, p(\boldsymbol{z}_i)). && (8)
\end{aligned}
$$

Second term of Eq. (6) can be derived as:

$$
\begin{aligned}
&\mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\left(\frac{p_\theta(\boldsymbol{c}|\boldsymbol{Z})}{q_\phi(\boldsymbol{c}|\boldsymbol{Z}, \mathcal{I})}\right)\right\}\\
&= \sum_{i=1}^{N} \mathbb{E}_{(\boldsymbol{z}_i,c_i)\sim q_\phi(\boldsymbol{z}_i,c_i|\mathcal{I})}\left\{\log\left(\frac{p_\theta(c_i|\boldsymbol{z}_i)}{q_\phi(c_i|\boldsymbol{z}_i, \mathcal{I})}\right)\right\} && (9)\\
&\approx \sum_{i=1}^{N} \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}_{c_i\sim q_\phi(c_i|\boldsymbol{z}_i^{(m)}, \mathcal{I})}\left\{\log\left(\frac{p_\theta(c_i|\boldsymbol{z}_i^{(m)})}{q_\phi(c_i|\boldsymbol{z}_i^{(m)}, \mathcal{I})}\right)\right\} && (10)\\
&= -\sum_{i=1}^{N} \frac{1}{M} \sum_{m=1}^{M} D_{KL}(q_\phi(c_i|\boldsymbol{z}_i^{(m)}, \mathcal{I}) \,||\, p_\theta(c_i|\boldsymbol{z}_i^{(m)})) && (11)
\end{aligned}
$$

where (10) follows from Eq. (9) by replacing the expectation over $\boldsymbol{z}_i$ with sample mean by generating M samples $\boldsymbol{z}_i^{(m)}$ from distribution $q(\boldsymbol{z}_i|\mathcal{I})$. Assuming $a_{ij} \in [0, 1]\ \forall i$, the third term of Eq. (6) is the negative of binary cross entropy (BCE) between observed and predicted edges.

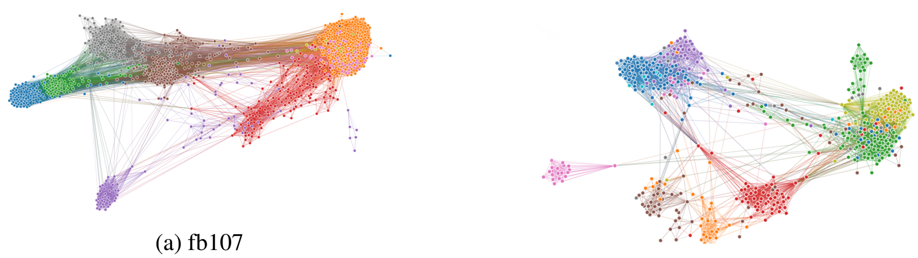
$$\mathbb{E}_{(\boldsymbol{Z},\boldsymbol{c})\sim q_\phi(\boldsymbol{Z},\boldsymbol{c}|\mathcal{I})}\left\{\log\left(p_\theta(\boldsymbol{A}|\boldsymbol{c},\boldsymbol{Z})\right)\right\}$$
$$= \sum_{(i,j)\in\mathcal{E}} \mathbb{E}_{(\boldsymbol{z}_i,\boldsymbol{z}_j,c_i,c_j)\sim q_\phi(\boldsymbol{z}_i,\boldsymbol{z}_j,c_i,c_j|\mathcal{I})}\left\{\log\left(p_\theta(a_{ij}|c_i,c_j,\boldsymbol{z}_i,\boldsymbol{z}_j)\right)\right\} \tag{12}$$

Hence, by substituting Eq. (8) and Eq. (11) in Eq. (6), we get the ELBO bound as:

$$\begin{aligned}\mathcal{L}_{ELBO} \approx & -\sum_{i=1}^{N} D_{KL}(q_\phi(\boldsymbol{z}_i|\mathcal{I}) \,||\, p(\boldsymbol{z}_i)) \\ & -\sum_{i=1}^{N}\frac{1}{M}\sum_{m=1}^{M} D_{KL}(q_\phi(c_i|\boldsymbol{z}_i^{(m)},\mathcal{I}) \,||\, p_\theta(c_i|\boldsymbol{z}_i^{(m)})) \\ & +\sum_{(i,j)\in\mathcal{E}} \mathbb{E}_{(\boldsymbol{z}_i,\boldsymbol{z}_j,c_i,c_j)\sim q_\phi(\boldsymbol{z}_i,\boldsymbol{z}_j,c_i,c_j|\mathcal{I})}\left\{\log\left(p_\theta(a_{ij}|c_i,c_j,\boldsymbol{z}_i,\boldsymbol{z}_j)\right)\right\}\end{aligned} \tag{13}$$

2 VISUALIZATION

Our experiments demonstrate that a single community-aware node embedding is sufficient to aid in both the node representation and community assignment tasks. This is also qualitatively demonstrated by graph visualizations of node embeddings (obtained via t-SNE (**?**)) and inferred communities for two datasets, fb107 and fb3437, presented in Fig. 1.



(a) fb107

(b) fb3437

Figure 1: Graph visualization with community assignments (better viewed in color)