APPENDIX

We provide the gradients used for ASGD in the proposed algorithms.

MVN2VEC-CON:

$$\frac{\partial O_{\text{CON}}}{\partial \mathbf{f}_u^v} = \left(1 - \sigma(\mathbf{f}_u^v \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_n^{v'}\}_{v' \in \mathcal{V}}))\right) \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_n^{v'}\}_{v' \in \mathcal{V}})$$
$$- \sum_{i=1}^{K} \sigma(\mathbf{f}_u^v \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_{n_i'}^{v'}\}_{v' \in \mathcal{V}})) \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_{n_i'}^{v'}\}_{v' \in \mathcal{V}}), \tag{9}$$

$$\frac{\partial O_{\text{CON}}}{\partial \tilde{\mathbf{g}}_n^{\hat{v}}} = \left(1 - \sigma(\mathbf{f}_u^v \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_n^{v'}\}_{v' \in \mathcal{V}}))\right) \cdot \mathbf{f}_u^v \cdot \begin{cases} \theta + \frac{1-\theta}{|\mathcal{V}|}, & \hat{v} = v, \\ \theta, & \hat{v} \neq v, \end{cases} \tag{10}$$

$$\frac{\partial O_{\text{CON}}}{\partial \tilde{\mathbf{g}}_{n_i'}^{\hat{v}}} = -\sigma(\mathbf{f}_u^v \cdot \varphi_\theta^v(\{\tilde{\mathbf{g}}_{n_i'}^{v'}\}_{v' \in \mathcal{V}})) \cdot \mathbf{f}_u^v \cdot \begin{cases} \theta + \frac{1-\theta}{|\mathcal{V}|}, & \hat{v} = v, \\ \theta, & \hat{v} \neq v. \end{cases} \tag{11}$$

MVN2VEC-REG:

$$\frac{\partial O_{\text{REG}}}{\partial \mathbf{f}_u^v} = \left(1 - \sigma(\mathbf{f}_u^v \cdot \tilde{\mathbf{f}}_n^v)\right) \cdot \tilde{\mathbf{f}}_n^v - \sum_{i=1}^{K} \sigma(\mathbf{f}_u^v \cdot \tilde{\mathbf{f}}_{n_i'}^v) \cdot \tilde{\mathbf{f}}_{n_i'}^v$$
$$+2\gamma\left(K+1\right)\left(1 - \frac{1}{|\mathcal{V}|}\right) \cdot \left(\mathbf{f}_u^v - \frac{1}{|\mathcal{V}|}\sum_{v' \in \mathcal{V}} \mathbf{f}_u^{v'}\right), \tag{12}$$

$$\frac{\partial O_{\text{REG}}}{\partial \tilde{\mathbf{f}}_n^v} = \left(1 - \sigma(\mathbf{f}_u^v \cdot \tilde{\mathbf{f}}_n^v)\right) \cdot \mathbf{f}_u^v + 2\gamma\left(1 - \frac{1}{|\mathcal{V}|}\right) \cdot \left(\tilde{\mathbf{f}}_n^v - \frac{1}{|\mathcal{V}|}\sum_{v' \in \mathcal{V}} \tilde{\mathbf{f}}_n^{v'}\right), \tag{13}$$

$$\frac{\partial O_{\text{REG}}}{\partial \tilde{\mathbf{f}}_{n_i'}^v} = -\,\sigma(\mathbf{f}_u^v \cdot \tilde{\mathbf{f}}_{n_i'}^v) \cdot \mathbf{f}_u^v + 2\gamma\left(1 - \frac{1}{|\mathcal{V}|}\right) \cdot \left(\tilde{\mathbf{f}}_{n_i'}^v - \frac{1}{|\mathcal{V}|}\sum_{v' \in \mathcal{V}} \tilde{\mathbf{f}}_{n_i'}^{v'}\right). \tag{14}$$

Note that in implementation, $|V|$ should be the number of views in which u is associated with at least one edge.