7.2 HDP-aMMSB

Methods. Unlike the time-series datasets above, very few relational datasets are accompanied by ground-truth community assignments. To assess the quality of our learned models, we turn to held-out data prediction. Specifically, we hold out 10% of edges and a similar number of non-edges at training time. At various points during inference, we compute the probability of each held-out edge under our approximate posterior $q(\cdot)$:

$$\mathbb{E}_q[x_{ij}] = \mathbb{E}_q[\pi_i]^T \mathbb{E}_q[\phi] \mathbb{E}_q[\pi_j] \tag{29}$$

By thresholding this probability, we can compute the precision and recall of our model as:

$$\text{recall} = \frac{\text{Num. correct } x_{ij} = 1}{\text{True num. } x_{ij} = 1}, \quad \text{precision} = \frac{\text{Num. correct } x_{ij} = 1}{\text{Total predicted } x_{ij} = 1} \tag{30}$$

We compute this curve every twenty iterations and show its progression over the course of inference.

We initialize our model's global parameters by a random initialization for $q(\pi)$: $\hat{\theta}_{i\ell} \sim \text{Gamma}(5, 2)$ and set $q(\phi)$ equal to its prior under $p(\cdot)$. Finally, both of our experiments use the hyperparameter settings described in Appendix D.

Toy Data. We next test our model on the toy dataset from Figure 2 containing $K = 6$ communities with fairly sparse community memberships π_i sampled from Dirichlet$(.05)$. We divide the data into $B = 20$ batches, each containing 10 nodes, and hold out 10% of edges along with a similar number of non-edges. As shown in Figure 10, our algorithm recovers the structure of the graph with few errors.

Co-authorship Network. Finally, we train our model on a co-authorship network of physicists working in quantum cosmology and general relativity[5]. This network contains many disconnected subnetworks; we train on the largest, which contains $N = 4,158$ nodes and 26,850 edges. We run using $K = 20$ and $B = 300$ batches.

Our results shown in Figure 11 demonstrate that our model learns substantial communities within the graph; however, although we obtain high precision, our model fails to reach even 30% recall on the held-out interactions. This is due to the model's high edge probability within the "main" community (middle row of Figure 11) and low probability elsewhere. The main community contains just under a third of the observed edges, accounting for nearly all of our correctly recalled edges.

[5] `https://snap.stanford.edu/data/ca-GrQc.html`