A Supplementary Material

A.1 Experimental Settings

For all our experiments, we fix the variance across node community memberships $\alpha = 1$ and set our hyperparameters for w_k to $\tau_a = 10$ and $\tau_b = 1$ across communities. We set an agressive learning rate so that $\mu_0 = 1$ and $\kappa = .5$. We use a restricted stratified node-sampling technique for all our experiments with the non-link partition set $m = 10$, unless stated otherwise. All experiments were run for 250,000 iterations from 5 random initializations with 10% of the links randomly held out along with an equal amount of non-links for testing. For the aMMSB, we used the same settings. The aMMSB uses a random initialization for $\theta_{ik} \sim \mathrm{Gam}(100, .01)$ with hyperparameters over w_k set to the expected number of link/non-links across K uniformly distributed communities. The learning rate was set to $\mu_0 = 1024$ and $\kappa = .5$. We found these settings gave the best advantage for the aMMSB on these datasets that were optimized for its original experiments, with the exception of changing the Dirichlet prior to be uniform over its mixed-membership distributions ($\alpha = 1$), which we found to improve convergence for the aMMSB across our experiments.

A.2 aHDPR ELBO

A more detailed representation of our ELBO for the aHDPR model can be seen here. Note that since we do not estimate $\phi_{ijk\ell}$, the ELBO needs to be computed in a more efficient manner:

$$\mathcal{L}(q) = \sum_{ij}^{E}\sum_{k=1}^{K}\left[\phi_{ijkk}\log f(w_k, y_{ij})\right] + \sum_{ij}^{E}\left[1 - (\sum_{k=1}^{K}\phi_{ijkk})\log f(\epsilon, y_{ij})\right] \tag{1}$$

$$+ \sum_{ij}^{E}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\left[\phi_{ijk\ell}(\mathbb{E}_q[\log(\pi_{ik})] + \mathbb{E}_q[\log(\pi_{j\ell})])\right] \tag{2}$$

$$+ \sum_{k=1}^{K}(\gamma - 1)\log(1 - v_k) + \sum_{i=1}^{N}\left[\log\Gamma(\sum_{k=1}^{K}\alpha\beta_k) - \sum_{k=1}^{K}\log\Gamma(\alpha\beta_k) + \sum_{k=1}^{K}(\alpha\beta_k - 1)\mathbb{E}_q[\log\pi_{ik}]\right] \tag{3}$$

$$+ \sum_{k=1}^{K}\left[\log\left(\frac{\Gamma(\tau_a + \tau_b)}{\Gamma(\tau_a)\Gamma(\tau_b)}\right) + (\tau_a - 1)\mathbb{E}_q[\log(w_k)] + (\tau_b - 1)\mathbb{E}_q[\log(1 - w_k)]\right] \tag{4}$$

$$- \sum_{ij}^{E}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\phi_{ijk\ell}\log(\phi_{ijk\ell}) \tag{5}$$

$$- \sum_{i=1}^{N}\left[\log\Gamma(\sum_{k=1}^{K}\theta_{ik}) - \sum_{k=1}^{K}\log\Gamma(\theta_{ik}) + \sum_{k=1}^{K}(\theta_{ik} - 1)\mathbb{E}_q[\log\pi_{ik}]\right] \tag{6}$$

$$- \sum_{k=1}^{K}\left[\log\left(\frac{\Gamma(\lambda_{ka} + \lambda_{kb})}{\Gamma(\lambda_{ka})\Gamma(\lambda_{kb})}\right) + (\lambda_{ka} - 1)\mathbb{E}_q[\log(w_k)] + (\lambda_{kb} - 1)\mathbb{E}_q[\log(1 - w_k)]\right] \tag{7}$$

where $\beta_k = v_k \prod_{\ell=1}^{k-1}(1 - v_\ell)$ and $\beta_{K+} = 1 - \sum_{k=1}^{K}\beta_k$. Since we no longer estimate $\phi_{ijk\ell}$ directly, we can show how our ELBO is modified with this optimized inference procedure. In particular, we focus on equations 21, and 24:

$$\sum_{ij}^{E}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\left[\phi_{ijk\ell}(\mathbb{E}_q[\log(\pi_{ik})] + \mathbb{E}_q[\log(\pi_{j\ell})])\right] = \sum_{ij}^{E}\left[\sum_{k=1}^{K}\mathbb{E}_q[\log(\pi_{ik})]\sum_{\ell=1}^{K}\phi_{ijk\ell} + \sum_{\ell=1}^{K}\mathbb{E}_q[\log(\pi_{j\ell})]\sum_{k=1}^{K}\phi_{ijk\ell}\right]$$

$$\sum_{ij}^{E}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\left[\phi_{ijk\ell}\log(\phi_{ijk\ell})\right] = \sum_{ij}^{E}\left[\sum_{k=1}^{K}\mathbb{E}_q[\log(\pi_{ik})]\sum_{\ell=1}^{K}\phi_{ijk\ell} + \sum_{\ell=1}^{K}\mathbb{E}_q[\log(\pi_{j\ell})]\sum_{k=1}^{K}\phi_{ijk\ell}\right.$$
$$\left. + \log f(\epsilon, y_{ij}) - \log f(\epsilon, y_{ij})\sum_{k=1}^{K}\phi_{ijkk} + \sum_{k=1}^{K}\log f(w_k, y_{ij})\phi_{ijkk} - \log(Z_{ij})\right] \tag{8}$$

For an efficient calculation of our ELBO the terms that we need to simplify are $\sum_{k=1}^{K}\mathbb{E}_q[\log(\pi_{ik})]\sum_{\ell=1}^{K}\phi_{ijk\ell}$ and $\sum_{\ell=1}^{K}\mathbb{E}_q[\log(\pi_{j\ell})]\sum_{k=1}^{K}\phi_{ijk\ell}$. From Equation 12, note that

$$\sum_{k=1}^{K}\mathbb{E}_q[\log(\pi_{ik})]\sum_{\ell=1}^{K}\phi_{ijk\ell} = \sum_{k=1}^{K}\tilde{\pi}_{ik}\left[\phi_{ijkk} + \frac{1}{Z_{ij}}\tilde{\pi}_{ik}f(\epsilon, y_{ij})(\tilde{\pi}_j - \tilde{\pi}_{jk})\right]$$

$$\sum_{\ell=1}^{K}\mathbb{E}_q[\log(\pi_{j\ell})]\sum_{k=1}^{K}\phi_{ijk\ell} = \sum_{\ell=1}^{K}\tilde{\pi}_{j\ell}\left[\phi_{ij\ell\ell} + \frac{1}{Z_{ij}}\tilde{\pi}_{j\ell}f(\epsilon, y_{ij})(\tilde{\pi}_i - \tilde{\pi}_{i\ell})\right]$$