Supplemental Materials for "Popularity Adjusted Block Models are Generalized Random Dot Product Graphs"

1 Sparsity Simulation Study

In this simulation study, we use the same setup as in the simulations for balanced communities but with fixed $K = 3$ and $n = 2^{11}$ and vary the sparsity parameter $\rho \in (0, 1]$ (this was fixed at $\rho_n = 1$ for the previous simulations. More specifically,

- Number of vertices $n = 2048$
- Number of underlying communities $K = 3$
- Sparsity parameter $\rho_n = 0.1, 0.3, 0.5, 0.7, 0.9$
- Mixture parameters $\alpha_k = 1/K$ for $k = 1, ..., K$ (i.e., each community label has an equal probability of being drawn)
- Community labels $z_k \overset{\text{iid}}{\sim} \text{Multinomial}(\alpha_1, ..., \alpha_K)$
- Within-group popularities $\lambda^{(kk)} \overset{\text{iid}}{\sim} \text{Beta}(2, 1)$
- Between-group popularities $\lambda^{(k\ell)} \overset{\text{iid}}{\sim} \text{Beta}(1, 2)$ for $k \neq \ell$

Fifty simulations were performed for each combination of n and K. The results for community detection and parameter estimation error are shown in Fig. 1. Community detection error is defined as the number of misclustered vertices. Parameter estimation error is defined as $\frac{1}{n}\|P - \hat{P}\|_F$ where $\hat{P}$ is the estimated edge probability matrix.

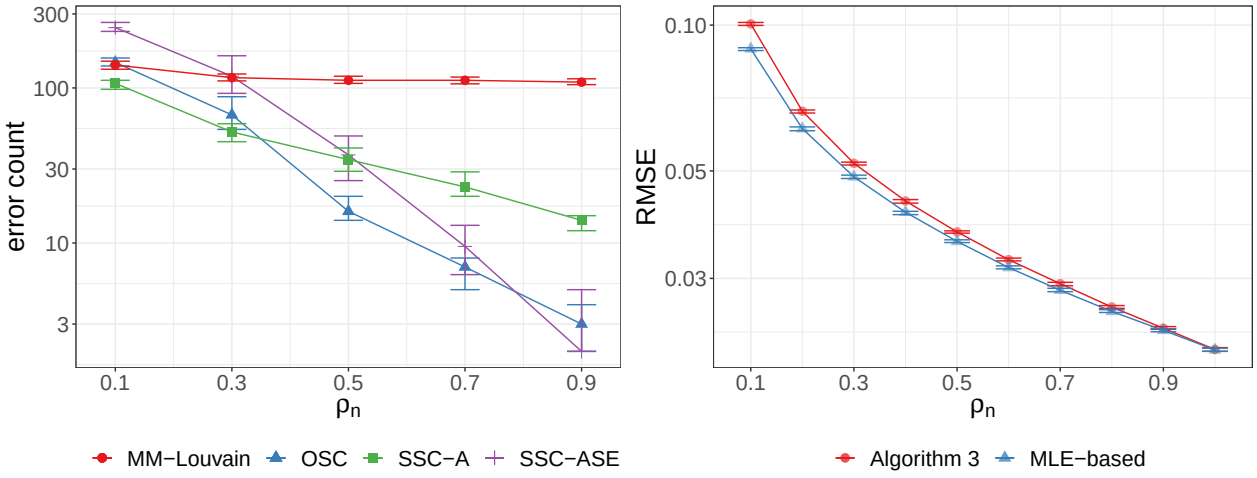

Figure 1: Community detection error counts (left) and popularity parameter RMSEs (right) for $n = 2048$ and $K = 3$ and varying ρ_n. Simulations were repeated 50 times for each ρ_n.