Caclulation

1 Calculation

We calculate what the covariance matrix looks like after a single partition/local complementation operation. We assume that the partition is sampled according to the stochastic block model. So the graph is partitioned in to subgraphs V_1 and V_2 and an edge between two vertices in V_i exists with probability p, while an edge between a vertex in V_1 and a vertex in V_2 exists with probability q. We assume $|V_1| = |V_2|$.

1.1 Edge Exists

First calculate $P(\text{edge e exists after partition/complement operation})$. Let e connect vertices i and j. This equals

$$P(\text{e exists after complementing}|(i \text{ and } j \in V_1)) \cdot P(i \text{ and } j \in V_1)$$

$$+P(\text{e exists after complementing}|(i \text{ and } j \in V_2)) \cdot P(i \text{ and } j \in V_2)$$

$$+P(e \text{ exists after complementing } |i \in V_1, j \in V_2) \cdot P(i \in V_1, j \in V_2)$$

$$+P(e \text{ exists after complementing } |i \in V_2, j \in V_1) \cdot P(i \in V_2, j \in V_1).$$

When vertices i and j are in the same partition after a complement, the probability the edge exists after the complement is $1-p$. Complements don't affect edges which go across partitions, so the probability of that edge existing after the complement is still q. So the above equals

$$(1-p) \cdot \frac{1}{4} + (1-p) \cdot 14 + q \cdot \frac{1}{4} + q \cdot \frac{1}{4}.$$

$$= \frac{1-p+q}{2}.$$

1.2 Variance

Let X_{e_i} denote the random variable where $X_{e_i} = 1$ if the edge e_i exists and 0 otherwise. So the covariance matrix we want to calculate looks like

$$\begin{bmatrix} Var(X_{e_1}) & & & Cov(X_{e_1}, X_{e_{n^2}}) \\ Cov(X_{e_2}, X_{e_1}) & Var(X_{e_2}) & & \\ . & . & . & \\ & & & \\ Cov(X_{e_{n^2}}, X_{e_1}) & & & Var(X_{e_{n^2}}) \end{bmatrix}$$

We first calculate the diagonal elements (i.e. the variances). By definition, $Var(X_{e_i}) = E(X_{e_i}^2) - E(X_{e_i})^2$. $E(X_{e_i})$ is just the probability that the edge e_i exists, which we already calculated.

Consider the random variable $Y_{e_i} = x_{e_i}^2$. Then Y_{e_i} is 1 when the edge e_i exists after the complement and zero otherwise. That is, it has the exact same distribution as X_{e_i}. Therefore,

$$Var(X_{e_i}) = E(Y_i) - E(X_{e_i})^2$$

$$= \left(\frac{1-p+q}{2}\right) - \left(\frac{(1-p+q)^2}{4}\right).$$

1.3 Covariance

We calculate $Cov(X_{e_i}, X_{e_j})$. By definition,

$$Cov(X_{e_i}, X_{e_j}) = E(X_{e_i} - \mu_{X_{e_i}})(X_{e_j} - \mu_{X_{e_j}})$$

$$= E(X_{e_i}X_{e_j} - X_{e_i}\mu_{X_{e_j}} - X_{e_j}\mu_{X_{e_i}} + \mu_{X_{e_i}}\mu_{X_{e_j}}).$$

We just need to calculate the product term. Then we can use the linearity to get the whole thing. See that $X_{e_i}X_{e_j}$ is the random variable which takes the value 1 if both edges e_i and e_j exist after complementing and zero otherwise. So $E(X_{e_i}X_{e_j}) = P(X_{e_i} = 1 \text{ and } X_{e_j} = 1)$. Let e_i connect vertices i, j and e_j connect vertices (i', j'). This is

$$P(X_{e_i=1} \wedge X_{e_j=1} | (i,j), (i',j') \in V_1) \cdot P((i,j), (i',j') \in V_1)$$

$$+P(X_{e_i=1} \wedge X_{e_j=1} | (i,j), (i',j') \in V_2) \cdot P((i,j), (i',j') \in V_2)$$

$$+P(X_{e_i=1} \wedge X_{e_j=1} | (i,j) \in V_1, (i',j') \in V_2) \cdot P((i,j) \in V_1, (i',j') \in V_2)$$

$$+P(X_{e_i=1} \wedge X_{e_j=1} | (i,j) \in V_2, (i',j') \in V_1) \cdot P((i,j) \in V_2, (i',j') \in V_1).$$

$$= (1-p)^2 \cdot 1/4 + (1-p)^2 \cdot 1/4 + (1-p)q \cdot 1/4 + (1-p)q \cdot 1/4$$

$$= \frac{(1-p)((1-p)+q)}{2}.$$

So the covariance is equal to $\frac{(1-p)((1-p)+q)}{2}$ plus the (expectation of) the other terms. These values and the variance values give us all entries in the covariance matrix after a single partition/complementing operation.