Practical session 5

Exercise 1 For two vertices u and v in a graph G, we denote by $dist_G(u, v)$ the length of a shortest path connecting them. The diameter of a graph is defined to be $\mathrm{diam}\,(G) = \max_{u,v} dist_G(u, v)$.

Let $G \sim G(n, m)$, where $m = \left\lceil \frac{1}{2}\binom{n}{2} \right\rceil$. Prove that

$$\lim_{n\to\infty} \Pr\left(\mathrm{diam}\,(G(n, m)) > 2\right) = 0.$$

Solution

Denote $N = \binom{n}{2}$ and for every integer $1 \le i < j \le n$ let $A_{i,j}$ denote the event "$dist_G(i, j) > 2$". Observe that $A_{i,j}$ occurs if and only if $ij \notin E(G)$, and $ik \notin E(G)$ or $jk \notin E(G)$ holds for every vertex $k \in [n] \setminus \{i, j\}$. For every integer $0 \le t \le n-2$ let $A^t_{i,j}$ denote the event "$dist_G(i, j) > 2$ and $d_G(i) + d_G(j) = t$". These events are pairwise disjoint and

$$A_{i,j} = \bigcup_{t=0}^{n-2} A^t_{i,j}.$$

Since $G(n, m)$ is a uniform probability space, it follows that

$$\Pr\left(A_{i,j}\right) = \sum_{t=0}^{n-2} \Pr\left(A^t_{i,j}\right) = \sum_{t=0}^{n-2} \frac{|A^t_{i,j}|}{\binom{N}{m}}.$$

$|A^t_{i,j}|$ can be calculated as follows. First, we choose the t vertices that are adjacent to either i or j (but not to both). Then, for each of the t chosen vertices we choose whether it is adjacent to i or to j. The remaining $n - 2 - t$ vertices will not be in $N_G(i) \cup N_G(j)$. Since we determined exactly the edges and non-edges incident to i and j, it remains to choose $m - t$ edges out of $(N - 1) - 2t - 2(n - 2 - t) = N - 2n + 3$ potential edges (equivalently, we may choose the $m - t$ edges from the $\binom{n-2}{2}$ pairs of $[n] \setminus \{i, j\}$).

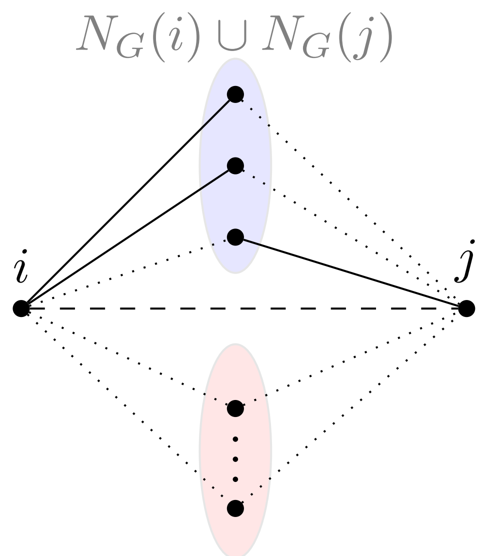

Figure 1: An example of $A^3_{i,j}$: three vertices are adjacent to either i or j (but not both) and the remaining $n-5$ vertices are connected to neither.

Therefore,

$$\begin{aligned}
\Pr\left(A_{i,j}\right) &= \sum_{t=0}^{n-2} \frac{\binom{n-2}{t} 2^t \cdot \binom{N-2n+3}{m-t}}{\binom{N}{m}} \\
&= \sum_{t=0}^{n-2} \binom{n-2}{t} 2^t \cdot \frac{m \cdot \ldots \cdot (m-t+1)}{N \cdot \ldots \cdot (N-t+1)} \cdot \frac{(N-2n+3) \cdot \ldots \cdot (N-2n+3-m+t+1)}{(N-t) \cdot \ldots \cdot (N-m+1)} \\
&\leq \sum_{t=0}^{n-2} \binom{n-2}{t} 2^t \cdot \left(\frac{m}{N}\right)^t \cdot \left(\frac{N-2n+3}{N-t}\right)^{m-t} \\
&\leq \sum_{t=0}^{n-2} \binom{n-2}{t} 2^t \cdot \left(\frac{1}{2} + \frac{1}{N}\right)^t \cdot \left(1 - \frac{2n-3-t}{N-t}\right)^{m-t} \\
&\leq \sum_{t=0}^{n-2} \binom{n-2}{t} \left(1 + \frac{2}{N}\right)^t \cdot e^{-\frac{(2n-3-t)(m-t)}{N-t}}, \qquad (1)
\end{aligned}$$

where the third inequality is due to the fact that $1-x \leq e^{-x}$ for every $x \in \mathbb{R}$. For every $0 \leq t \leq n-2$ it holds that

$$\begin{aligned}
e^{-\frac{(2n-3-t)(m-t)}{N-t}} &\leq e^{-\frac{(2n-3-t)(m-n)}{N}} \\
&= e^{-\frac{m(2n-3)}{N} + \frac{n(2n-3-t)}{N} + \frac{tm}{N}} \\
&\leq e^{-\frac{N/2\cdot(2n-3)}{N} + \frac{2n^2}{N} + \frac{tm}{N}} \\
&\leq e^{-n+3/2+\frac{2n^2}{N} + \frac{t(N/2+1)}{N}} \\
&\leq e^{-n+2+8+t/2+1} \\
&= e^{11} \cdot e^{-n} \cdot (\sqrt{e})^t. \qquad (2)
\end{aligned}$$

Combining (1) and (2) we obtain

$$\begin{aligned}\Pr(A_{i,j}) &\le \sum_{t=0}^{n-2} \binom{n-2}{t} \left(1+\frac{2}{N}\right)^t \cdot e^{-\frac{(2n-3-t)(m-t)}{N-t}} \\ &\le e^{11} \cdot e^{-n} \cdot \sum_{t=0}^{n-2} \binom{n-2}{t} \left(1+\frac{2}{N}\right)^t \cdot (\sqrt{e})^t \\ &= e^{11} \cdot e^{-n} \cdot \left(1+\sqrt{e}(1+2/N)\right)^{n-2},\end{aligned}$$

where the above equality holds by Newton's binomial formula. For sufficiently large values of n, the quantity $e^{11} \cdot e^{-n} \cdot \left(1+\sqrt{e}(1+2/N)\right)^{n-2}$ is of the form $c \cdot a^n$ for some constants $c > 0$ and $0 < a < 1$. It follows that $\Pr(A_{i,j}) = o(1/N)$ for every $1 \le i < j \le n$. A union bound argument then implies that

$$\Pr(\operatorname{diam}(G) > 2) \le \sum_{1\le i<j\le n} \Pr(A_{i,j}) \le N \cdot e^{11} \cdot e^{-n} \cdot (1+\sqrt{e}(1+2/N))^{n-2} = o(1).$$