Problem sheet 4

Problem 4.1. Show that the function $h\colon \mathscr{G} \to [0,1]$ defined as $h(\mathcal{G},o) = |\mathcal{G}|^{-1}\mathbb{1}_{\{|\mathcal{G}|<\infty\}}$ is continuous in the weak local topology.

Proof. We have to show that the inverse image $h^{-1}(\mathcal{B})$ of each open subset $\mathcal{B} \subset [0,1]$ is open in $\mathscr{G}$. Since the topology of open sets in $[0,1]$ is generated by open intervals, it suffices to consider $\mathcal{B} = (a,b) \subset [0,1]$. Let

$$n(a) = \max\{n \in \mathbb{N}\colon \tfrac{1}{n} > a\} \quad \text{and} \quad n(b) = \min\{n \in \mathbb{N}\colon \tfrac{1}{n} < b\}.$$

By definition, each finite rooted graph $(\mathcal{G},o)$ with $|\mathcal{V}| \in [n(b), n(a)]$ yields $h(\mathcal{G},o) \in (a,b)$. Next, write $\mathscr{G}_n$ for the subset of all rooted graph on exactly n vertices. Then, $\mathscr{G}_n$ is open as

$$\mathscr{G}_n = \bigcup_{\mathcal{G}\in\mathscr{G}_m} \{\mathcal{G}\},$$

which is a finite union of open sets (recall that singletons of finite graphs are open in the local topology). Therefore,

$$h^{-1}(a,b) = \bigcup_{n=n(b)}^{n(a)} \mathscr{G}_n$$

is a union of open sets and thus open itself. This finishes the proof. □

Problem 4.2. Consider the Stochastik block model on the vertices $\mathcal{V}_n = \mathcal{V}_n^{(1)} \cup \mathcal{V}_n^{(2)}$, where $\mathcal{V}_n = [n]$ and $\mathcal{V}_n^{(1)}$ contains of all even labels and $\mathcal{V}_n^{(2)}$ of all uneven labels. We connect each vertex of the same type independently with probability a/n and each pair of different types with probability b/n. Identify the Poisson BGW tree $\mathcal{T}$ that describes the weak local limit and identify, for which a,b the tree has positive probability of survival.

Proof. First of all, recall that the stochastik block model is an IRG with type space $\{0,1\}$ with limiting type distribution $\mu(1) = \mu(2) = 1/2$ (in the current situation). Moreover, the connection kernel is given as $\kappa_n(x_i,x_j) = \kappa(x_i,x_j) = a + (b-a)\mathbb{1}_{\{a\neq b\}}$. Hence, the limiting BGW tree $\mathcal{T}$ has Poisson offspring with mean

$$\lambda_{i,j} = \kappa(i,j)\mu(j) = \tfrac{1}{2}\Big(a\mathbb{1}_{\{i=j\}} + b\mathbb{1}_{\{i\neq j\}}\Big)$$

by Theorem 3.1. This yields the expectation matrix

$$M = \tfrac{1}{2}\begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

In order to determine the supercritical regime, we have to identify the largest eigenvalue ρ. To this end, we calculate

$$0 = \det(M - \lambda\mathbb{1}) = (\tfrac{a}{2} - \lambda)^2 - \tfrac{b^2}{4} = \lambda^2 - a\lambda + \tfrac{a^2-b^2}{4}.$$

This equation has the solutions

$$\lambda_{1,2} = \frac{a}{2} \pm \sqrt{\frac{a^2}{4} - \frac{a^2 - b^2}{4}} = \frac{a \pm b}{2}.$$

Thus, $\rho = (a+b)/2$ and the BGW tree is supercritical if $\rho > 1 \Leftrightarrow a + b > 2$. □

Problem 4.3 (Theorem 4.6)**.** Prove Theorem 4.6 of the lecture notes: Let $(\mathcal{G}_n)$ be the inhomogeneous random graph associated with the kernels (κ_n) that converge graphically to κ. Then the empirical degree-degree distribution converges to the probability measure ν_2 given by

$$\nu_2(k,l) = \frac{1}{c} \int \frac{\lambda(x)^{k-1}}{(k-1)!} e^{-\lambda(x)} \Big(\int \frac{\lambda(y)^{l-1}}{(l-1)!} e^{-\lambda(y)} \kappa(x,y) \mu(\mathrm{d}y) \Big) \mu(\mathrm{d}x),$$

where $c = \int \int \kappa(x,y)\mu(\mathrm{d}y)\mu(\mathrm{d}x)$.

Proof. We apply Proposition 4.4 and note that uniform integrability in the assumptions was only used to ensure sparsity. Hence, it is not required here by Lemma 1.5 and we have

$$\nu_2(k,l) = \frac{k}{\mathbb{E}\mathcal{D}(o)} \mathbb{P}(\mathcal{D}(o) = k, \mathcal{D}(V) = l).$$

As $c = \mathbb{E}(\mathcal{D}(o))$, it hence remains to calculate $k\mathbb{P}(\mathcal{D}(o) = k, \mathcal{D}(V) = l)$ for the limiting BGW tree. Recall that the number of offsprings of a type x vertex in the limiting BGW tree is Poisson distributed with parameter $\lambda(x) = \int \kappa(x,y)\mu(\mathrm{d}y)$. Thus, denoting (o,x) for the root of type x,

$$\begin{aligned}
\mathbb{P}(\mathcal{D}(o) = k, \mathcal{D}(V) = l) &= \int_{\mathscr{S}} \mathbb{P}(\mathcal{D}(o,x) = k, \mathcal{D}(V) = l) \mu(\mathrm{d}x) \\
&= \int_{\mathscr{S}} \mathbb{P}(\mathcal{D}(o,x) = k) \mathbb{P}(\mathcal{D}(V) = l \mid \mathcal{D}(o,x) = k) \mu(\mathrm{d}x) \\
&= \int_{\mathscr{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)} \mathbb{P}(\mathcal{D}(V) = l \mid \mathcal{D}(o,x) = k) \mu(\mathrm{d}x)
\end{aligned}$$

To calculate the remaining probability, recall that, given $\mathcal{D}(o,x) = k$, the k children of the root are independently distributed according to the normalised intensity measure $\kappa(x,y)\mu(\mathrm{d}y)/\lambda(x)$, which follows from the Poisson point process structure of the neighbourhood. As V is chosen uniformly at random among these k vertices, it follows the same distribution. Hence,

$$\mathbb{P}(\mathcal{D}(V) = l \mid \mathcal{D}(o,x) = k) = \int_{\mathscr{S}} \mathbb{P}(\mathcal{D}(v,y) = \ell) \frac{\kappa(x,y)\mu(dy)}{\lambda(x)},$$

where (v,y) is a root's child of type y. In order to have degree l, the vertex v that is already connected to the root must have $l-1$ children itself. As each vertex produces its offspring independently, we have

$$\mathbb{P}(\mathcal{D}(v,y) = l) = \frac{\lambda(y)^{l-1}}{(l-1)!} e^{-\lambda(y)}.$$

Putting everything together, we finally infer,

$$\begin{aligned} k\mathbb{P}(\mathcal{D}(o)=k,\mathcal{D}(V)=l) &= \int_{\mathscr{S}} \frac{\lambda(x)^k}{(k-1)!} e^{-\lambda(x)} \mathbb{P}(\mathcal{D}(V)=l \mid \mathcal{D}(o,x)=k)\mu(\mathrm{d}x) \\ &= \int \frac{\lambda(x)^k}{(k-1)!} e^{-\lambda(x)} \Big(\int \frac{\lambda(y)^{l-1}}{(l-1)!} e^{-\lambda(y)} \frac{\kappa(x,y)\mu(\mathrm{d}y)}{\lambda(x)} \Big) \mu(\mathrm{d}x), \end{aligned}$$

as claimed. □

Problem 4.4 (Global clustering)**.** In the lecture, we studied the *local clustering coefficient*, describing the probability that two neighbours of a uniformly chosen vertex are neighbours themselves and thus form a triangle. In contrast the *global clustering coefficient* is defined as

$$\boldsymbol{c}_n^{\text{glob}} = \frac{\sum_{v\in\mathcal{G}_n} \Delta(\mathcal{G}_n, v)}{\frac{1}{2}\sum_{v\in\mathcal{G}_n} \mathcal{D}(v)(\mathcal{D}(v)-1)},$$

i.e. the overall proportion of wedges that complete a triangle. Assume $(\mathcal{G}_n)_n$ converges weakly locally to $(\mathcal{G},o)\in\mathscr{G}$ and that the sequence $(\mathcal{D}(O_n)^2)_n$ is uniformly integrable. Show that

$$\boldsymbol{c}_n^{\text{glob}} \longrightarrow \frac{\mathbb{E}[\Delta(\mathcal{G},o)]}{\frac{1}{2}\mathbb{E}[\mathcal{D}(o)(\mathcal{D}(o)-1)]},$$

in probability, as $n\to\infty$.

Proof. (a): Consider the function $h(G,o)=\Delta(G,o)$, which is continuous but unbounded. Note that

$$\frac{1}{n}\sum_v h(\mathcal{G}_n,v) = h(\mathcal{G}_n,O_n) \le \tfrac{1}{2}\mathcal{D}(O_n)(\mathcal{D}(O_n)-1) \le \mathcal{D}(O_n)^2.$$

Therefore, we can choose $\varepsilon,\delta>0$ and $K>1$, such that for all $k>K$

$$\begin{aligned} \mathbb{P}\Big(\tfrac{1}{n}\sum_v \Delta(\mathcal{G}_n,v)\mathbb{1}_{\{\mathcal{D}(v)>k\}} > \delta\Big) &\le \tfrac{1}{\delta}\mathbb{E}[\Delta(\mathcal{G}_n,O_n)\mathbb{1}_{\{\mathcal{D}(O_n)>k\}}] \\ &\le \tfrac{1}{\delta}\mathbb{E}[\mathcal{D}(O_n)^2\mathbb{1}_{\{\mathcal{D}(O_n)^2>k\}}] \\ &< \tfrac{\varepsilon}{2}, \end{aligned}$$

by the uniform integrability of $\mathcal{D}(O_n)^2$. Next, for $k>K$, we have by local convergence

$$\mathbb{P}\Big(\Big|\Delta(\mathcal{G}_n,O_n)\mathbb{1}_{\{\mathcal{D}(O_n)\le k\}} - \Delta(\mathcal{G},o)\mathbb{1}_{\{\mathcal{D}(o\le k)\}}\Big| > \delta\Big) \le \tfrac{\varepsilon}{2}.$$

Finally, we have $\mathbb{E}[\mathcal{D}(o)^2]<\infty$ by uniform integrability. Therefore,

$$\mathbb{E}[\Delta(\mathcal{G},o)] - \mathbb{E}[\Delta(\mathcal{G},o))\mathbb{1}_{\{\mathcal{D}(o)\le k\}}] < \delta,$$

for sufficiently large $k > K$. Combined, this yields

$$\begin{aligned}
&\mathbb{P}\Big(\Big|\tfrac{1}{n}\sum_v \Delta(\mathcal{G}_n, v) - \mathbb{E}[\Delta(\mathcal{G}, o))\Big| > 3\delta\Big) \\
&\quad\le \mathbb{P}\Big(\Big|\tfrac{1}{n}\sum_v \Delta(\mathcal{G}_n, v) - \mathbb{E}[\Delta(\mathcal{G}, o))\mathbb{1}_{\{\mathcal{D}(o)\le k\}}\Big| > 2\delta\Big) \\
&\quad\le \mathbb{P}\Big(\Big|\tfrac{1}{n}\sum_v \Delta(\mathcal{G}_n, v)\mathbb{1}_{\{\mathcal{D}(v)\le k\}} - \mathbb{E}[\Delta(\mathcal{G}, o))\mathbb{1}_{\{\mathcal{D}(o)\le k\}}\Big| > \delta\Big) \\
&\qquad + \mathbb{P}\Big(\Big|\tfrac{1}{n}\sum_v \Delta(\mathcal{G}_n, v)\mathbb{1}_{\{\mathcal{D}(v)> k\}} > \delta\Big) \\
&\quad\le \varepsilon.
\end{aligned}$$

Put differently,

$$\frac{1}{n}\sum_{v\in\mathcal{G}_n} \Delta(\mathcal{G}_n, v) \longrightarrow \mathbb{E}[\Delta(\mathcal{G}, o)], \quad \text{in probability.}$$

With exactly the same arguments, we also have

$$\frac{1}{n}\sum_{v\in\mathcal{G}_n} \mathcal{D}(v)(\mathcal{D}(v)-1) \longrightarrow \mathbb{E}[\mathcal{D}(o)(\mathcal{D}(o)-1)], \quad \text{in probability.}$$

Combined, we infer

$$2\boldsymbol{c}^{\text{glob}} = \frac{\sum_v \Delta(\mathcal{G}_n, v)}{n}\cdot\frac{n}{\sum_v \mathcal{D}(v)(\mathcal{D}(v)-1)} \longrightarrow \frac{\mathbb{E}\Delta(\mathcal{G}, o)}{\mathbb{E}[\mathcal{D}(o)(\mathcal{D}(o)-1)]},$$

in probability. □

Problem 4.5 (Proposition 4.11)**.** Finish the proof of Proposition 4.11: Let $(\mathcal{G}_n)$ be a sequence of random graphs converging weakly locally to a random rooted graph $(\mathcal{G}, o)$. Then, in probability,

$$\frac{\mathcal{K}_n}{n} \longrightarrow \mathbb{E}\Big[\tfrac{1}{|\mathcal{G}|}\Big] \ge 0.$$

Proof. Define the function $h\colon \mathscr{G} \to [0,1]$ via $h(G,o) = |\mathcal{C}(o)|^{-1}\mathbb{1}_{\{|\mathcal{C}(o)|<\infty\}}$. Recall that, in the space $\mathscr{G}$ of rooted connected locally finite graphs, we identify the graph G with the connected component $\mathcal{C}(o)$ of its root. Hence, $h(G,o) = |\mathcal{G}|^{-1}\mathbb{1}_{\{|G|<\infty\}}$ and it is continuous and bounded by Problem 4.1. Moreover, in a finite graph all components are finite. Note further

$$\mathcal{K}_n = \sum_{v\in\mathcal{G}_n} |\mathcal{C}(v)|^{-1},$$

as each component appears in the sum precisely as often as its size. Therefore, local convergence yields

$$\frac{\mathcal{K}_n}{n} = \frac{1}{n}\sum_{v\in\mathcal{G}_n} h(\mathcal{G}_n, v) \longrightarrow \mathbb{E}[h(\mathcal{G}, o)] = \mathbb{E}[|\mathcal{G}|^{-1}\mathbb{1}_{\{\mathcal{G}<\infty\}}]$$

However, as $1/\infty = 0$, we also have

$$\mathbb{E}|\mathcal{G}|^{-1} = \mathbb{E}[|\mathcal{G}|^{-1}\mathbb{1}_{\{\mathcal{G}<\infty\}}] + \mathbb{E}[0\cdot\mathbb{1}_{\{\mathcal{G}=\infty\}}] = \mathbb{E}[|\mathcal{G}|^{-1}\mathbb{1}_{\{\mathcal{G}<\infty\}}],$$

in probability, which concludes the proof. □

Problem 4.6. Let $(\mathcal{G}_n)_n$ be a sequence of random graphs converging weakly locally to a rooted graph $(\mathcal{G}, o)$, and $\mathcal{V}_n^{(2)}$ be the set of all vertices of degree at least two. Show that

$$\frac{n}{|\mathcal{V}_n^{(2)}|} \longrightarrow \frac{1}{\mathbb{P}(\mathcal{D}(o) \geq 2)} \tag{1}$$

in probability. Using (1), show that $\frac{1}{n}\sum_{v\in\mathcal{V}_n^{(2)}} \boldsymbol{c}_n^{\text{loc}}(v) \longrightarrow \mathbb{E}\Big[\frac{\Delta(\mathcal{G},o))}{\frac{1}{2}\mathcal{D}(o)(\mathcal{D}(o)-1)}\mathbb{1}_{\{\mathcal{D}(o)\geq 2\}}\Big]$ in probability implies

$$\boldsymbol{c}_n^{\text{av}} \longrightarrow \mathbb{E}\Big[\frac{\Delta(\mathcal{G},o)}{\frac{1}{2}\mathcal{D}(o)(\mathcal{D}(o)-1)}\,\Big|\,\mathcal{D}(o) \geq 2\Big],$$

in probability.

Proof. Clearly,

$$\frac{|\mathcal{V}_n^{(2)}|}{n} = \frac{1}{n}\sum_{v\in\mathcal{G}_n} \mathbb{1}_{\{\mathcal{D}(v)\geq 2\}} \longrightarrow \mathbb{P}(\mathcal{D}(o) \geq 2),$$

in probability by local convergence. This proves (1). Hence, in probability,

$$\begin{aligned}\boldsymbol{c}_n^{\text{av}} &= \frac{n}{|\mathcal{V}_n^{(2)}|}\cdot\frac{\sum_{v\in\mathcal{V}_n^{(2)}}\boldsymbol{c}_n^{\text{loc}}(v)}{n} \longrightarrow \frac{1}{\mathbb{P}(\mathcal{D}(o)\geq 2)}\mathbb{E}\Big[\frac{\Delta(\mathcal{G},o))}{\frac{1}{2}\mathcal{D}(o)(\mathcal{D}(o)-1)}\mathbb{1}_{\{\mathcal{D}(o)\geq 2\}}\Big] \\ &= \mathbb{E}\Big[\frac{\Delta(\mathcal{G},o)}{\frac{1}{2}\mathcal{D}(o)(\mathcal{D}(o)-1)}\,\Big|\,\mathcal{D}(o)\geq 2\Big].\end{aligned}$$

□