Lecture notes on random graphs and probabilistic combinatorial optimization

`!! draft in construction !!`

Charles Bordenave [1]

April 8, 2016

[1]Institut de Mathématiques - Université de Toulouse & CNRS - France. Email: bordenave@math.univ-toulouse.fr

Contents

1 Models of random graphs **9**

1.1 Some graph terminology . . . 9

1.2 Erdős-Rényi random graph . . . 11

1.3 Uniform graph with given degree sequence . . . 12

1.3.1 Definition . . . 12

1.3.2 Degree distribution . . . 12

1.4 The configuration model . . . 15

1.5 Chung-Lu graph . . . 17

1.6 Dynamic graphs . . . 18

2 Subgraph counts and Poisson approximation **19**

2.1 Average subgraph counts . . . 19

2.1.1 Erdős-Rényi graphs . . . 19

2.1.2 Configuration model . . . 20

2.2 Poisson Approximation . . . 22

2.2.1 Method of moments . . . 22

2.2.2 Total variation distance and coupling . . . 23

2.2.3 Basics of Stein's method . . . 24

2.2.4 Stein's method for the Poisson distribution . . . 26

2.3 Cycle counts . . . 28

2.3.1 Erdős-Rényi graphs . . . 28

2.3.2 Configuration model . . . 31

2.4 Graphs with given degree sequence . . . 36

3 Local weak convergence **39**

3.1 Weak convergence in metric spaces 39
3.2 The space of rooted unlabeled networks 41
3.3 Converging graph sequences 43
3.4 Unimodular Galton-Watson trees 46
3.4.1 Galton-Watson trees 46
3.5 Convergence of random graphs 48
3.5.1 Erdős-Rényi graphs 48
3.5.2 Configuration model 52
3.6 Concentration and convergence of random graphs 57
3.6.1 Bounded difference inequality 57
3.6.2 Almost sure convergence of Erdős-Rényi random graphs 59
3.6.3 Concentration inequality on uniform matchings 62
3.6.4 Almost sure convergence in the configuration model 63

4 The giant connected component **65**

4.1 Growth of Galton-Watson trees 65
4.2 Random walks and branching processes 68
4.3 Hitting time for random walks 69
4.4 Emergence of the giant component 72
4.5 Erdős-Rényi graph : proof of theorem 4.13 74
4.5.1 Proof of theorem 4.13(i) 74
4.5.2 Proof of theorem 4.13(ii) 75
4.6 Configuration Model : : proof of theorem 4.14 78
4.6.1 Proof of theorem 4.14(i) 78
4.6.2 Proof of theorem 4.14(ii) 81
4.7 Application to network epidemics 84
4.7.1 A simple SIR dynamic 84
4.7.2 Dynamic on the Erdős-Rényi graph 85
4.7.3 Dynamic on the configuration model 86

5 Continuous length combinatorial optimization 87

5.1 Issues of combinatorial optimization 87

5.2 Limit of random networks 89

5.3 The minimal spanning tree 89

5.4 Maximal weight independent set 89

5.4.1 Proof of theorem 5.1 90

Notation

$\mathbb{N}$	set of positive integers $1, 2, \ldots$.
$\mathbb{Z}_+$	set of non-negative integers $0, 1, 2, \ldots$.
$\mathbb{R}_+$	set of non-negative real numbers $[0, \infty)$.
$\mathcal{P}(\mathcal{X})$	set of probability measures on $\mathcal{X}$.
$\mathbb{G}(V)$	set of locally finite graphs on the vertex set V.
$\widehat{\mathbb{G}}(V)$	set of locally finite multigraphs on the vertex set V.
$\mathcal{G}_*$	set of equivalence classes of locally finite connected rooted graphs.
$\widehat{\mathcal{G}}_*$	set of equivalence classes of locally finite connected rooted multigraphs.
$\lvert S \rvert$	cardinal of a finite set S.
$\mu_n \rightsquigarrow \mu$	the sequence $(\mu_n)_n$ tends weakly to μ for continuous bounded functions.
$X \overset{d}{\sim} \mu$	the random variable X follows the law μ.
$\mathcal{L}(X)$	the law of random variable X (i.e. $X \overset{d}{\sim} \mathcal{L}(X)$).
$X_n \overset{d}{\to} X$	the sequence of random variables $(X_n)_n$ converges in distribution to X (i.e. $\mathcal{L}(X_n) \rightsquigarrow \mathcal{L}(X)$).
$d_G(u, v)$	the graph distance between u and v, with $u, v \in V_G$
$B_G(u, t)$	the set of vertices of V_G at graph distance at most t from $u \in V_G$.

Chapter 1

Models of random graphs

1.1 Some graph terminology

We start with elementary definitions that will be used throughout these notes. Let V be a countable set, and let E be a set of distinct pairs of elements in V. We call an element in V a *vertex* and an element in the image of E an *edge*. The sets V and E define a *graph* $G = (V, E)$. In graph theory, this would rather be called a labeled simple graph but we will stick here to 'graph'. If E is a multi-set of non-necessarily distinct pairs of elements of V, the pair (V, E) is called a *multi-graph*.

In a multi-graph a *loop* is an edge $e \in E$ such that for some vertex $v \in V$, $e = \{v, v\}$. An edge $e \in E$ is said to be *multiple* if e has cardinality larger than 1 in E. Note that a graph is a multigraph with no loop nor multiple edge.

A *network* or *weighted graph* $G = (V, E, \omega)$ is a graph (V, E) together with a complete separable metric space Ω called the *mark space* and a map ω from $V \cup E$ to Ω. Images in Ω are called marks. Note that a multigraph is a network with marks on $\Omega = \mathbb{N} = \{1, 2, \cdots\}$. For $e = \{u, v\} \in E$, $\omega(e)$ is the number of edges between u and v while $\omega(v)$ counts the number of loops on v.

The *degree* of a vertex $v \in V$, $\deg(v)$ or $\deg(v; G)$ is the number of edges incident to v with loops counting twice. A (multi)graph is regular if all vertices have the same degree. if A (multi)graph is *locally finite* if the degree of each vertex is finite. A (multi)graph is finite if the sets V and E are finite.

We will denote by $\mathbb{G}(V)$ and $\widehat{\mathbb{G}}(V)$ the set of locally finite graphs and multigraphs on the vertex set V. If the vertex set is $[n] = \{1, \cdots, n\}$ for some integer n, then we will simply write $\mathbb{G}(n)$ and $\widehat{\mathbb{G}}(n)$ in place of $\mathbb{G}([n])$ and $\widehat{\mathbb{G}}([n])$.

For $W \subset V$, we denote by $G \cap W$ the restriction of G to vertex set W: an edge $e = \{u, v\} \in E$ is in $G \cap W$ if u and v are in W. Similarly, $G \backslash W$ is $G \cap (V \backslash W)$. We say that $G' = (V', E')$ is a *subgraph* of G if $V' \subset V$ and $E' \subset E$.

The symmetric group S_V of V acts naturally on the network: the image of an edge being the pair of the image of its adjacent vertices. The *(vertex)-automorphism group* of a network G, Aut(G), is the subgroup of S_V that leaves the graph invariant. More generally, a bijective map from V to V' defines a *network isomorphism*. Then if $G = (V, E)$ and $G' = (V', E')$ are two networks with common mark space Ω, we say that G' and G are *isomorphic* if G' is the image of G by a network isomorphism. Network isomorphisms define an equivalence relation denoted by "$\simeq$". In graph theory, an equivalence class of simple graphs is called an unlabeled graph. Note that if $G \simeq G'$ then $|\mathrm{Aut}(G)| = |\mathrm{Aut}(G')|$.

For a multi-graphs, there is also a notion of *edge-automorphism group*. Let $G = (V, E)$ with a finite number m of edges, loops counting for two edges. Index its edges in an arbitrary manner from 1 to m, loops being indexed as a set of two indices. We then obtain a network $\bar{G}$ with marks on edge $\{u, v\}$ equal to the set of indices of the edges $\{u, v\}$, and marks on vertex u equal to the set of pairs of indices of loops on u. The permutation group S_m acts on the network $\bar{G}$ by assigning on edge $\{u, v\}$ the image by the permutation of the marks. We may then define the edge-automorphism group of H as the group of permutations on the indices that keeps $\bar{H}$ invariant. We denote by b the cardinal of this group. If G is a graph then $b = 1$. More generally, if $\omega(v)$ is the number of loops attached to v and for and $\omega(e)$ is the multiplicity of e, we have,

$$b = \prod_{v \in V} (2^{\omega(v)} \omega(v)!) \prod_{e \in E} (\omega(e)!).$$

Let $\ell \geq 1$ be an integer. A *path* π of length ℓ from u to v in G is a sequence $(u_0, \cdots, u_\ell)$ of vertices in V such that $u_0 = u$, $u_\ell = v$ and for $i = 1 \cdots, \ell$, $\{u_{i-1}, u_i\} \in E$. A (multi)graph is *connected* if for any u, v in V there exists a path from u to v. A *cycle* $(u_0, \cdots, u_\ell)$ is a path from u to u such that for $0 \leq i \neq j \leq \ell - 1$, $u_i \neq u_j$. A *tree* is a connected graph without cycle. A *forest* is a graph without cycle.

We define the excess as

$$\mathrm{exc}(G) = |E| - |V|.$$

Lemma 1.1 (Excess and trees) *If G is a connected (multi)graph, then*

$$\mathrm{exc}(G) \geq -1.$$

Moreover, G is a tree if and only if $\mathrm{exc}(G) = -1$.

Proof. Let $u \in V$ be a distinguished vertex and consider for all $v \in V \backslash \{u\}$ a shortest path $\{u_0(v), u_1(v), \cdots, u_{k_v}(v)\}$ from v to u : $u_0(v) = v$, $u_{k_v}(v) = u$. Define the mapping φ from $V \backslash \{u\}$ to E by setting $\sigma(v) = \{v, u_1(v)\}$. Since the paths are the shortest possible, σ is an injection, and it follows that $|V \backslash \{u\}| \leq |E|$. In the case of equality $|V \backslash \{u\}| = |E|$, σ is a bijection and it is easy to check that G is a tree. □

Exercise 1.2 *Let $k \geq 3$ be an integer and $G = ([k], \{\{1, 2\}, \cdots \{k-1, k\}, \{k, 1\}\})$ be a cycle of length k. Show that* $|\mathrm{Aut}(G)| = 2k$.

Exercise 1.3 *Let G', G be two finite graphs. Assume that $G' \subset G$ and G connected. Show that* $\text{exc}(G') \le \text{exc}(G)$. *(Hint : adapt the proof of lemma 1.1 by considering shortest paths from $v \in V_G \backslash V_{G'}$ to $V_{G'}$).*

1.2 Erdős-Rényi random graph

Let p be a positive real and n an positive integer, the Erdős-Rényi random graph $\mathcal{G}(n,p)$ is a probability distribution on $\mathbb{G}(n)$ such that each of the $n(n-1)/2$ possible edge is present independently and with probability $\min(p,1)$. In other words, if G is a random graph with distribution $\mathcal{G}(n,p)$, $0 \le p \le 1$, and H is a graph with n vertices and m edges then

$$\mathbb{P}(\{H\}) = \mathbb{P}(G = H) = p^m (1-p)^{\frac{n(n-1)}{2} - m}. \tag{1.1}$$

In particular, $\mathcal{G}(n,1/2)$ is the uniform measure on $\mathbb{G}(n)$. It is important to point out that random graph G is homogeneous : for any permutation $\sigma in \S_n$, $\sigma(G)$ and G have the same distribution (in other words G is exchangeable).

The distribution of $\deg(1;G)$ is a Binomial distribution with parameter $n-1$ and p. In particular, the average degree of vertex 1 is

$$\mathbb{E}\deg(1;G) = (n-1)p.$$

In these notes, we will mainly study the asymptotic properties of random graphs with uniformly bounded average degrees. We will thus be mainly interested by the probability distribution $\mathcal{G}(n, \lambda/n)$ with $\lambda \in \mathbb{R}^+$. In this case, $\deg(1;G)$ is a Binomial distribution with parameter $n-1$ and λ/n. It follows for all integer k

$$\mathbb{P}(\deg(1;G) = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

As n goes to infinity, this converges to $e^{-\lambda}\lambda^k/k!$. In other words, we retrieve the well known fact that the Binomial distribution with parameter n and λ/n converges to a Poisson distribution with parameter λ.

The distribution $\mathcal{G}(n,p)$ was first introduced by Gilbert (1959). It owes its name to an independent celebrated paper of Erdős and Rényi (1959) who had defined the random graph on n vertices and m uniformly distributed edges. The books Bollobás (2001), Janson, Łuczak, and Rucinski (2000) cover a good part of the known properties of this random graph. For a more probabilistic treatment, we refer to Durrett (2007) and van der Hofstad (2012).

1.3 Uniform graph with given degree sequence

1.3.1 Definition

Let $\mathbf{d} = (d_1, \cdots, d_n) \in \mathbb{Z}_+^n$ be a sequence of non-negative integers. We say that $\mathbf{d}$ *is graphic* if $\mathbb{G}(\mathbf{d})$, the set of graphs G on $[n]$ such that for all $i \in [n]$, $\deg(i; G) = d_i$, is not empty. If $\mathbf{d}$ is graphic, we may then define $\mathcal{G}(\mathbf{d})$ as the uniform probability distribution on $\mathbb{G}(\mathbf{d})$.

It is not completely obvious how to characterize graphic sequences. This question has been settled by Erdős and Gallai (1960). Here, we may just notice that if $\mathbf{d}$ is graphic then $\sum_{i=1}^n d_i$ is even (since it is equal to twice the sum of degrees).

An important case is $\mathbf{d} = (d, \cdots, d)$ for some $d \geq 2$. In this case, $\mathbb{G}(\mathbf{d})$ is the set of d-regular graphs on n vertices. If $\mathbf{d}$ is graphic, the probability distribution $\mathcal{G}(\mathbf{d})$ will be usually denoted by $\mathcal{G}(n, d)$. This probability is called the *uniform d-regular graph* on n vertices. Uniform regular graphs are especially interesting structures, for a specific review, see Wormald (1999).

1.3.2 Degree distribution

If G is a graph with degree sequence $\mathbf{d} = (d_1, \cdots, d_n)$, the *degree distribution* of G is defined as the probability measure on $\mathbb{Z}_+$

$$P_{\mathbf{d}} = \frac{1}{n} \sum_{i=1}^{n} \delta_{d_i},$$

where δ is the Dirac distribution. Equivalently, $P_{\mathbf{d}} \in \mathcal{P}(\mathbb{Z}_+)$ is defined for all $k \in \mathbb{Z}_+ = \{0, 1, \cdots\}$ by

$$P_{\mathbf{d}}(\{k\}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(d_i = k).$$

Note that the measure $P_{\mathbf{d}}$ contains less information than $\mathbf{d}$, the labels of the degrees have been lost.

In these notes, we will be mainly interested by large graph asymptotics. Let $P \in \mathcal{P}(\mathbb{Z}_+)$ and $p \in \mathbb{R}_+$. We will often consider that a sequence $\mathbf{d}_n = (d_1(n), \cdots, d_n(n)), n \geq 1$ satisfies some the following hypothesis:

(H_0) $P_{\mathbf{d}_n}$ converges weakly to P with $P(\{0\}) < 1$, i.e. for any $k \in \mathbb{Z}_+$,

$$\lim_{n \to \infty} P_{\mathbf{d}_n}(\{k\}) = P(\{k\}).$$

(H_p) H_0 holds and, if $D(n)$ and D have law $P_{\mathbf{d}_n}$ and P,

$$\lim_{n \to \infty} \mathbb{E} D(n)^p = \mathbb{E} D^p < \infty,$$

equivalently,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} d_i(n)^p = \sum_{k\geq 0} k^p P(\{k\}).$$

The probability distribution P will be called the asymptotic distribution of $\mathbf{d}_n$. In the sequel, we will often use the following lemma.

Lemma 1.4 (Convergence of degree sequence) *Let $k \in \mathbb{N}$ and assume that (H_0) holds. Let $(D_1(n), \cdots, D_k(n))$ be a uniformly sampled k-tuple without replacement on $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$. Then, we have the convergence in distribution,*

$$(D_1(n), \cdots D_k(n)) \xrightarrow{d} (D_1, \cdots, D_k),$$

where $(D_1, \cdots, D_k)$ are i.i.d. with law P, i.e. for any subset $A \subset \mathbb{Z}_+^k$,

$$\lim_{n\to\infty} \mathbb{P}\left((D_1(n), \cdots D_k(n)) \in A\right) = \mathbb{P}\left((D_1, \cdots, D_k) \in A\right).$$

Assume further that (H_p) holds for some $p \in \mathbb{N}$, then, for any real $0 \leq p_\ell \leq p$, $1 \leq \ell \leq k$, we have

$$\lim_{n\to\infty} \mathbb{E}\prod_{\ell=1}^{k} D_\ell(n)^{p_\ell} = \prod_{\ell=1}^{k} \mathbb{E}D^{p_\ell}.$$

Proof. The first statement can be proved with a simple coupling argument. Let $(i_1, \cdots, i_k)$ be i.i.d. variables uniformly distributed on $[n]$ and σ be an independent uniformly sampled injection from $[k]$ to $[n]$. Then $(d_{i_1}(n), \cdots, d_{i_k}(n))$ are i.i.d. variables with law $P_{\mathbf{d}_n}$ and $(d_{\sigma(1)}(n), \cdots, d_{\sigma(k)}(n))$ has the same law than $(D_1(n), \cdots, D_k(n))$. Moreover, conditioned on the event E that $(i_1, \cdots, i_k)$ are all distinct, $(i_1, \cdots, i_k)$ has the same law than $(\sigma(1), \cdots, \sigma(k))$. This event E has probability equal to

$$\frac{(n)_k}{n^k},$$

$(n)_k = n(n-1)\cdots(n-k+1)$. The above probability goes to 1 as n goes to infinity. We deduce for any event A that

$$\begin{aligned}
&|\mathbb{P}\left((D_1(n), \cdots, D_k(n)) \in A\right) - \mathbb{P}\left((d_{i_1}(n), \cdots, d_{i_k}(n)) \in A\right)| \\
&\quad \leq \left|\mathbb{P}\left((d_{\sigma(1)}(n), \cdots, d_{\sigma(k)}(n)) \in A \cap E\right) - \mathbb{P}\left((d_{i_1}(n), \cdots, D_{i_k}(n)) \in A \cap E\right)\right| + \mathbb{P}(E^c) \\
&\quad = \mathbb{P}(E^c)
\end{aligned}$$

Now, (H_0) implies that $\mathbb{P}((d_{i_1}(n), \cdots, d_{i_k}(n)) \in A)$ converges to $\mathbb{P}((D_1, \cdots, D_k) \in A)$. We have proved the first statement.

The second statement requires a little more care. With the above notation, we have

$$\begin{aligned}\mathbb{E}\prod_{\ell=1}^{k} d_{i_\ell}(n)^{p_\ell} &= \frac{1}{n^k}\sum_{\tau:[k]\to[n]}\prod_{\ell=1}^{k} d_{\tau(\ell)}(n)^{p_\ell} \\ &= \frac{(n)_k}{n^k}\mathbb{E}\prod_{\ell=1}^{k} D_\ell(n)^{p_\ell} + \frac{1}{n^k}\sum_{*}\prod_{\ell=1}^{k} d_{\tau(\ell)}(n)^{p_\ell},\end{aligned}$$

where the last sum is over all maps $\tau : [k] \to [n]$ which are not injective. We set

$$M(n) = \max(d_1(n), \cdots, d_n(n)).$$

Since the image of such map τ has cardinal at most $k-1$, it follows that

$$\begin{aligned}\frac{1}{n^k}\sum_{*}\prod_{\ell=1}^{k} d_{\tau(\ell)}(n)^{p_\ell} &\le \frac{M(n)^p}{n}\frac{1}{n^{k-1}}\sum_{1\le i_1,\cdots,i_{k-1}\le n}\prod_{\ell=1}^{k-1} d_{i_\ell}(n)^p \\ &= \frac{M(n)^p}{n}\left(\frac{1}{n}\sum_{i=1}^{n} d_i(n)^p\right)^{k-1} \\ &= \frac{M(n)^p}{n}\left(\mathbb{E}D(n)^p\right)^{k-1}.\end{aligned}$$

Now, from lemma 1.5, we have

$$M(n)^p = o(n).$$

It remains to use assumption (H_p) to conclude the proof. □

Lemma 1.5 (Bound of max degree) *Assume that (H_p) holds for some $p \in \mathbb{N}$, then,*

$$\lim_{n\to\infty} n^{-1/p}\max(d_1(n), \cdots, d_n(n)) = 0.$$

Proof. Define $M(n) = \max(d_1(n), \cdots, d_n(n))$. From (H_0), we have for any $t > 0$,

$$\lim_{n\to\infty}\mathbb{E}(D(n)\mathbf{1}_{\{D(n)\le t\}})^p = \mathbb{E}(D\mathbf{1}_{\{D\le t\}})^p.$$

Now, from (H_p), $\lim_{t\to\infty}\mathbb{E}(D\mathbf{1}_{\{D\le t\}})^p. = \mathbb{E}D^p$. It yields to

$$\lim_{t\to\infty}\lim_{n\to\infty}\mathbb{E}(D(n)\mathbf{1}_{\{D(n)>t\}})^p = 0.$$

In particular, for any $\varepsilon > 0$, there exists t, such that for all n large enough,

$$\mathbb{E}(D(n)\mathbf{1}_{\{D(n)>t\}})^p \le \varepsilon^p.$$

However, notice that

$$\mathbb{E}(D(n)\mathbf{1}_{\{D(n)>t\}})^p \ge \frac{M(n)^p\mathbf{1}_{\{M(n)>t\}}}{n}.$$

Hence, either $M(n) \le t$ or $M(n) \le n^{1/p}\varepsilon$. Letting n tending to infinity and then ε to 0 concludes the proof. □

1.4 The configuration model

The configuration model was originally introduced in Bollobás (1980) in the context of regular graphs. More recently, it has drawn a renewed attention after the work Molloy and Reed (1995). For its relevance for real life networks see Chung and Lu (2006). As above, let $\mathbf{d} = (d_1, \cdots, d_n)$ be a sequence of integers. If $\sum_{i=1}^n d_i$ is even then there exists multigraphs with degree sequence $\mathbf{d}$. It is much simpler to build a probability distribution on $\widehat{\mathbb{G}}(\mathbf{d})$, the set of multigraphs on $[n]$ such that for all $i \in [n]$, $\deg(i; G) = d_i$.

It is done explicitly as follows. Let Δ be a finite set with even cardinal. A *matching* of a finite set Δ is an involution of Δ (i.e. a permutation that is its own inverse) with no fixed point (i.e. a derangement). Let $M(\Delta)$ be the set of matchings of the set Δ. If Δ is even, the number of matchings is given by

$$|M(\Delta)| = (|\Delta| - 1)(|\Delta| - 3) \cdots 1 = (|\Delta| - 1)!!.$$

Now, for a sequence of integers $\mathbf{d} = (d_1, \cdots, d_n)$ we define $\Delta = \{(i, j) : 1 \leq i \leq n, 1 \leq j \leq d_i\}$. Let $m \in M(\Delta)$, we define the multigraph $G(m)$ on $[n]$ with edge set

$$E = \{\{i, i'\} : m(i, j) = (i', j'), (i, j) \in \Delta\}.$$

The set Δ is thought as the set of half-edges which are matched to form an edge, see figure 3.1.

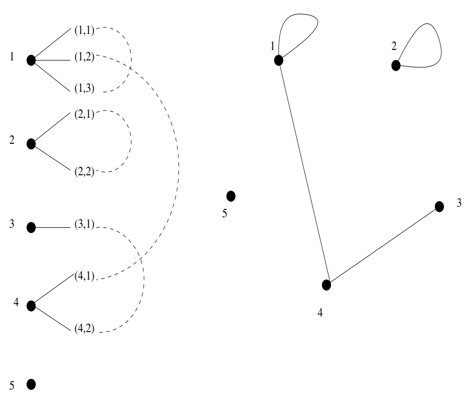

Figure 1.1: A matching and its corresponding multigraph.

If $\sum_{i=1}^n d_i$ is even, then for all $i \in [n]$, $\deg(i; G(m)) = d_i$. Let σ be a random matching of Δ drawn uniformly among all matchings. Then, we may define the random multigraph $G = G(\sigma)$ on $[n]$. We denote by $\widehat{\mathcal{G}}(\mathbf{d})$ the corresponding probability distribution on $\widehat{\mathbb{G}}(\mathbf{d})$, it is called the *configuration model*. By construction, if A is a subset of $\widehat{\mathbb{G}}(n)$, we have

$$\mathbb{P}(G \in A) = \frac{1}{|M(\Delta)|} \sum_{m \in M(\Delta)} \mathbf{1}(G(m) \in A). \tag{1.2}$$

It is possible to compute explicitly the marginal distribution of $\widehat{\mathcal{G}}(\mathbf{d})$. For a graph

Lemma 1.6 (Marginal probability of configuration model) *Let $H \in \widehat{\mathbb{G}}(\mathbf{d})$ with b elements in its edge-automorphism group. Then, if $G \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d})$,*

$$\mathbb{P}(G = H) = \frac{\prod_{i=1}^n (d_i!)}{b\left(\sum_{i=1}^n d_i\right)!!}.$$

Lemma 1.6 implies that $\widehat{\mathcal{G}}(\mathbf{d})$ is not the uniform probability distribution on $\mathbb{G}(\mathbf{d})$. Note however that if $H \in \mathbb{G}(\mathbf{d})$, then H is a graph and $n_j = 0$ for $j \geq 2$. In particular, the above probability is constant on $\mathbb{G}(\mathbf{d})$. Hence lemma 1.6 has a beautiful consequence.

Corollary 1.7 (Configuration model restricted to graphs) *If $\mathbf{d}$ is graphic, then the configuration model $\widehat{\mathcal{G}}(\mathbf{d})$ conditioned on $\{G \in \mathbb{G}(n)\}$, is $\mathcal{G}(\mathbf{d})$, the uniform probability distribution on $\mathbb{G}(\mathbf{d})$.*

Proof of lemma 1.6. The map $m \mapsto G(m)$ from $M(\Delta)$ to $\widehat{\mathbb{G}}(\mathbf{d})$ is surjective (i.e. each multigraph in $\widehat{\mathbb{G}}(\mathbf{d})$ can be obtained by some matching). In view of equation (1.2), we should prove that

$$\sum_{m \in M(\Delta)} \mathbf{1}(G(m) \in H) = |G^{-1}(\{H\})| = b^{-1} \prod_{i=1}^n (d_i!). \tag{1.3}$$

We fix a matching m such that $G(m) = H$. If $m' \in M(\Delta)$ satisfies $G(m) = G(m')$ then there exists a family of permutations $\alpha = (\alpha_i)_{i \in [n]}$ such that $\alpha_i \in S_{d_i}$ and for all $(i, j) \in \Delta$,

$$m'(i, \alpha_i(j)) = (i', \alpha_{i'}(j')),$$

where $m(i, j) = (i', j')$. Conversely, for any sequence of permutations $(\alpha_i)_{i \in [n]}$ such that $\alpha_i \in S_{d_i}$, the above identity defines a matching $m' = m_\alpha$ such that $G(m_\alpha) = G(m)$.

Assume first than $H \in \mathbb{G}(\mathbf{d})$ is a graph. If the permutations $(\alpha_i)_{i \in [n]}$ are not all the identity, we have $m \neq m_\alpha$. Equivalently, the map $\alpha \to m_\alpha$ is a bijection from $S_{d_1} \times \cdots S_{d_n}$ to $G^{-1}(\{H\})$. We deduce that any $H \in \mathbb{G}(\mathbf{d})$ is obtained by $\prod_{i=1}^n (d_i!)$ different matchings of Δ.

In the general case, if $H \in \widehat{\mathbb{G}}(\mathbf{d})$, each element $m' \in G^{-1}(\{H\})$ can be obtained from b elements of $S_{d_1} \times \cdots S_{d_n}$. Indeed, assume first that H has a multiple edge $\{i, i'\}$ with multiplicity k: $m(i, j_\ell) = (i', j'_\ell)$, for $1 \leq \ell \leq k$. Then, if σ is any permutation on $\{j_1, \cdots, j_k\}$, composing α_i by σ to get $\alpha_i \circ \sigma$ leaves the matching unchanged. Similarly, assume that H has k loops at i and $m(i, j_1) = (i, j_2), \cdots, m(i, j_{2k-1}) = (i, j_{2k})$ with j_ℓ all distinct. Then, if σ is any permutation on $\{j_1, \cdots, j_k\}$ and if we compose the permutation α_i by a product of transpositions of $(j_{2\ell-1}\ j_{2\ell})$ of the form: $\alpha_i \circ (j_{2\sigma(1)}\ j_{2\sigma(1)-1}) \circ \cdots \circ (j_{2\sigma(k)}\ j_{2\sigma(k)-1})$, we leave the matching unchanged.

In summary, there are $\prod_{i=1}^n (d_i!)/b$ matchings such that $G(m) = H$. This proves (1.3). □

We will see in the next chapters that the configuration model $\widehat{\mathcal{G}}(\mathbf{d})$ is a convenient probabilistic tool to analyze $\mathcal{G}(\mathbf{d})$. As already pointed, we will be mainly interested by degree sequence $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$ of n integers with even sum which satisfies property (H_0).

1.5 Chung-Lu graph

Let us mention an *inhomogeneous* version of the Erdős-Rényi graph, namely the Chung-Lu graph, see Chung and Lu (2006). Its level of difficulty ranges between the Erdős-Rényi graph and the configuration model. In these notes it will mostly be used as a source of exercises. Let $\underline{\lambda} = (\lambda_i)_{1\leq i\leq n}$ be collection of non-negative real numbers. For integer $n \geq 1$, let

$$\|\underline{\lambda}\|_1 = \sum_{i=1}^{n} \lambda_i.$$

We assume that $\|\underline{\lambda}\|_1 > 0$. We build a graph G on $[n]$ by putting independently the edge $\{i, j\}$ with $i \neq j$, with probability

$$p_{ij} = \frac{\lambda_i \lambda_j}{\|\underline{\lambda}\|_1} \wedge 1.$$

We denote the corresponding graph ensemble by $\mathcal{G}(n, \underline{\lambda})$. The marginal probability is easy to compute: for any graph $H = ([n], E) \in \mathbb{G}(n)$, we have

$$\mathbb{P}(G = H) = \prod_{1\leq i<j\leq n} \left((1 - p_{ij})\mathbf{1}_{\{i,j\}\notin E} + p_{ij}\mathbf{1}_{\{i,j\}\in E}\right).$$

As usual, we may define the intensity distribution as the empirical measure

$$P_{\underline{\lambda}} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}.$$

It is interesting to consider a sequence of intensities $\underline{\lambda}_n = (\lambda_1(n), \cdots, \lambda_n(n))$ such that the following assumption holds, for $p > 0$,

(H'_0) $P_{\underline{\lambda}_n}$ converges weakly to $P \in \mathcal{P}(\mathbb{R}_+)$, $P(\{0\}) < 1$.

(H'_p) H'_0 holds and, if $\Lambda(n)$ and Λ have law $P_{\underline{\lambda}_n}$ and P,

$$\lim_{n\to\infty} \mathbb{E}\Lambda(n)^p = \mathbb{E}\Lambda^p < \infty.$$

If the sequence of $\underline{\lambda} = (\lambda_i)_{i\in\mathbb{N}}$ is iid with common law Λ on $(0, \infty)$, then we shall denote the distribution of this random graph as $\mathcal{G}(n, \Lambda)$. In the next chapter, we will see how to compute the asymptotic degree distribution of a sequence of graphs $G_n \overset{d}{\sim} \mathcal{G}(n, \underline{\lambda}_n)$ which satisfy the above assumption.

Exercise 1.8 *Assume that for all $i \in [n]$, $\lambda_i = c > 0$. What is then the distribution $\mathcal{G}(n, \underline{\lambda})$?*

Exercise 1.9 *Assume that for any $1 \leq i, j \leq n$, $\lambda_i\lambda_j \leq \|\underline{\lambda}\|_1$. If $G \overset{d}{\sim} \mathcal{G}(n, \underline{\lambda})$, check that the average degree of vertex $i \in [n]$ is $\mathbb{E}\mathrm{deg}(i; G) = \frac{\|\underline{\lambda}\|_1 - \lambda_i}{\|\underline{\lambda}\|_1}\lambda_i$.*

Exercise 1.10 *Check that (H'_p) implies that $\max_{1\leq i\leq n} \lambda_i(n) = o(n^{1/p})$ and that (H'_2) implies that for all n large enough, any $1 \leq i, j \leq n$, $\lambda_i(n)\lambda_j(n) \leq \|\underline{\lambda}_n\|_1$.*

1.6 Dynamic graphs

Of course, there are many models of random graphs besides the above defined models : Erdős-Rényi graph, uniform graph with given degree sequence or Chung-Lu graphs. In this manuscript, to keep the exposition clear, we shall restrict to the study ourselves to these 3 models. Roughly speaking, there are two main ways of defining a random graph. First way: the random graph is defined for fixed n according to some random connectivity rule (like our 3 models). Second way: the graph is defined iteratively by a random aggregation rule, the most studied being arguably the preferential attachment model (introduced in Barabási and Albert (1999)), for another interesting direction, we may mention the Kronecker graphs (refer to Leskovec et al. (2010)). The focus there is to use a simple aggregation dynamics as an explanation of phenomena in 'real world' graphs (e.g. power law degree distribution, clustering, or small world phenomenon).

Chapter 2

Subgraph counts and Poisson approximation

2.1 Average subgraph counts

2.1.1 Erdős-Rényi graphs

In this chapter, we will count the number of times a given subgraph appears in a random graph. More precisely, let $G \in \widehat{\mathbb{G}}(V)$ and $H \in \widehat{\mathbb{G}}(k)$ be finite multigraphs on V and $[k]$, with $k \leq |V|$. We define

$$X(H;G) = \sum_{F \subset G} \mathbf{1}(F \simeq H),$$

where the sum is over all subgraphs of G (of k elements).

If n is a non-negative integer and k a positive integer, we define

$$(n)_k = n(n-1)\cdots(n-k+1) \quad \text{and} \quad (n)_0 = 1.$$

Similarly, for n even we define,

$$(\!(n)\!)_k = (n-1)(n-3)\cdots(n-2k+1) \quad \text{and} \quad (\!(n)\!)_0 = 1$$

If G is an Erdős-Rényi random graph, it is easy to compute the first moment of X.

Proposition 2.1 (Subgraph count in Erdős-Rényi graph) *Let $1 \leq k \leq n$, $H \in \mathbb{G}(k)$ with m edges and c elements in its automorphism group. If G is a random graph with distribution $\mathcal{G}(n, \lambda/n)$, $\lambda \leq n$, then*

$$\mathbb{E}X(H;G) = c^{-1}(n)_k \left(\frac{\lambda}{n}\right)^m \sim_{n\to\infty} c^{-1}\lambda^m n^{-\mathrm{exc}(H)}.$$

Proof. By assumption,

$$X(H;G) = \frac{1}{c}\sum_{\tau} \mathbf{1}(\tau(H) \subset G),$$

where the sum is over all injective maps from $[k]$ to $[n]$. There are $(n)_k$ such injective maps. Now, if τ is an injective map from $[k]$ to V, from Equation (1.1),

$$\mathbb{P}(\tau(H) \subset G) = \left(\frac{\lambda}{n}\right)^m.$$

The conclusion follows. □

This lemma implies that the structure of the Erdős-Rényi graph is far from the lattice graph $\mathbb{Z}^d$. For example, the lattice graph $\mathbb{Z}^d \cap [1,m]^d$ on $n = m^d$ vertices has subgraphs of any excess in number of order n. For an Erdős-Rényi graph, the only connected subgraphs in number of order n are trees. Proposition 2.1 gives the convergence of the average of subgraph counts. We will also give a deviation inequality for $\mathbb{P}(|X(H;G) - \mathbb{E}X(H;G)| \geq t)$ in the forthcoming chapter 3 which will be meaningful when H is a tree.

Corollary 2.2 (Large excess subgraph in Erdős-Rényi graph) *Let $k \geq 4$ be an integer and H be a graph in $\mathbb{G}(k)$ with $\mathrm{exc}(H) \geq 1$. For each $n \in \mathbb{N}$, let G_n be an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$. Then, in probability, $X(H;G_n) \to 0$.*

Proof. From Markov inequality $\mathbb{P}(X(H;G_n) \geq 1) \leq \mathbb{E}X(H;G_n)$. Then by lemma 2.1 we have $\mathbb{P}(X(H;G_n) \geq 1) = O(n^{-\mathrm{exc}(H)})$. □

As an simple corollary, we also have

Corollary 2.3 (Cycle count in Erdős-Rényi graph) *Let $H = ([k], \{\{1,2\},\{2,3\},\cdots\{k,1\}\})$ be a cycle of length $k \geq 3$, we have*

$$\lim_{n\to\infty} \mathbb{E}X(H;G) = \frac{\lambda^k}{2k}.$$

2.1.2 Configuration model

We now turn to the configuration model. We consider a array of integers $(d_1(n), \cdots, d_n(n))$ satisfying condition (H_0) and such that for all integer n, $\sum_{i=1}^n d_i(n)$ is even. We define the random variable

$$D \overset{d}{\sim} P.$$

Proposition 2.4 (Subgraph count for configuration model) *Let $1 \leq k \leq n$, $H \in \widehat{\mathbb{G}}(k)$ with m edges and maximal degree $p \geq 1$. Assume that H has b and c elements in its edge- and (vertex)-automorphism groups. Let $G \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_p), then*

$$\mathbb{E}X(H;G) =\sim_{n\to\infty} n^{-\mathrm{exc}(H)} \frac{\prod_{i=1}^k \mathbb{E}\left[(D)_{\deg(i;H)}\right]}{bc(\mathbb{E}D)^m},$$

where D has distribution P.

As a corollary, we get immediately,

Corollary 2.5 (Cycle count for configuration model) *Assume that $G \stackrel{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_2). If $H_1 = (\{1\}, \{\{1,1\}\})$ is a single loop then*

$$\lim_{n\to\infty} \mathbb{E}X(H_1; G) = \frac{\mathbb{E}(D)_2}{2\mathbb{E}D}.$$

If $H_2 = (\{1,2\}, \{\{1,2\},\{1,2\}\})$ is a single multi-edge then

$$\lim_{n\to\infty} \mathbb{E}X(H_2; G) = \frac{(\mathbb{E}(D)_2)^2}{4(\mathbb{E}D)^2}.$$

If $k \geq 3$ and $H_k = ([k], \{\{1,2\},\{2,3\},\cdots\{k,1\}\})$ is a cycle of length k then

$$\lim_{n\to\infty} \mathbb{E}X(H_k; G) = \frac{(\mathbb{E}(D)_2)^k}{2k(\mathbb{E}D)^k}.$$

As in corollary, 2.2, we get:

Corollary 2.6 (Large excess subgraph in configuration model) *Let $k \geq 1$ be an integer and $H \in \widehat{\mathbb{G}}(k)$ with $\mathrm{exc}(H) \geq 1$ and maximal degree p. Let $G \stackrel{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_p). Then, in probability, $X(H; G_n) \to 0$.*

Proof of proposition 2.4. For ease of notation, let us skip the parameter n. Let $S = \sum_{i=1}^n d_i$. From (H_p), for all real a,

$$\lim_{n\to\infty} \frac{S-a}{n} = \mathbb{E}D > 0. \tag{2.1}$$

Arguing as in the proof of proposition 2.1,

$$\mathbb{E}X(H; G) = c^{-1} \sum_{\tau} \mathbb{E}Y(\tau(H); G), \tag{2.2}$$

where the sum is over all injective maps from $[k]$ to $[n]$ and $Y(H; G)$ is the number of times that $H \subset G$. Note that since G is a multigraph $Y(H; G)$ may be larger than 1. We think as $\Delta_i = \{(i,j) : 1 \leq j \leq d_i\}$ as a set of half-edges adjacent to vertex i. These half-edges are matched to other half-edges by a uniformly drawn matching of $\Delta = \{(i,j) : 1 \leq i \leq n, 1 \leq j \leq d_i\}$. Let $m_i = \deg(i; H)$, we have

$$\mathbb{E}Y(H; G) = \frac{\prod_{i=1}^k (d_i)_{m_i}}{b\,(\!(S)\!)_m}. \tag{2.3}$$

Indeed, arguing as in the proof of lemma 1.6, there are $b^{-1} \prod_{i=1}^k (d_i)_{m_i}$ ways of choosing the half-edges to be matched in order to give the subgraph H. Then, given the choice of the half-edges, the probability that they are effectively matched is $1/(S-1)(S-3)\cdots(S-2m+1)$.

From (2.2), we deduce that

$$\mathbb{E}X(H;G) = \frac{1}{bc\,(\!(S)\!)_m} \sum_{\tau} \prod_{i=1}^{k} (d_\tau(i))_{m_i} = \frac{(n)_k}{bc\,(\!(S)\!)_m} \mathbb{E} \prod_{i=1}^{k} (D_i(n))_{m_i}, \tag{2.4}$$

where $(D_1(n), \cdots, D_k(n))$ is uniformly sampled without replacement on $\mathbf{d}_n = (d_1(n), \cdots, d_k(n))$.

Now, from (2.1), we have

$$(\!(S)\!)_m \sim n^m (\mathbb{E}D)^m.$$

On the other hand, lemma 1.4 implies that

$$\mathbb{E} \prod_{i=1}^{k} (D_i(n))_{m_i} \to \prod_{i=1}^{k} \mathbb{E}\left[(D)_{m_i}\right].$$

This concludes the proof. □

2.2 Poisson Approximation

2.2.1 Method of moments

In the next Section, we will give a closer look at the random variable $X(H;G)$ when $\mathrm{exc}(H) = 0$. From propositions 2.1, 2.4 we know that the expectation $\mathbb{E}X(H;G)$ has a non-degenerate limit when the size of the graph tends to infinity. We will see in the next section that if H is simple enough, we can actually prove that $X(H;G)$ converges weakly to a Poisson random variable.

Let X be a real random variables with all its moments finite : for any integer $k \geq 1$, $\mathbb{E}[X^k] = m_k < \infty$. Assume further that there exists a unique probability measure P on $\mathbb{R}$ such that for all integer $k \geq 1$, $\int x^k dP = m_k$. In the latter case, we say that P is uniquely determined by its moments. Carleman's theorem asserts that it is indeed the case if

$$\sum_{k \geq 1} m_{2k}^{-\frac{1}{2k}} = \infty.$$

If the random variable has bounded support, the Carleman condition is satisfied.

A commonly used method to prove that a sequence of real random variables $(X_n)_{n\geq 1}$ converges in distribution to the random variable X is to show that for all integer $k \geq 1$, $\lim_{n\to\infty} \mathbb{E}[X_n^k] = \mathbb{E}[X^k] = m_k$. More formally, the method of moments is contained in the next lemma.

Lemma 2.7 (Method of moments) *Let $(P_n)_{n\geq 1}$ be a sequence of real probability measures. Assume that $P \in \mathcal{P}(\mathbb{R})$ is uniquely determined by its moments. If for all $k \geq 1$,*

$$\lim_{n\to\infty} \int x^k dP_n(x) = \int x^k dP$$

then $P_n \rightsquigarrow P$.

Proof. We have $\int x^2 dP_n = m_2 + o(1) \leq c$ for some c. In particular, from Markov inequality $P_n([-t,t]^2) \leq c/t^2$. Hence, from Prohorov's theorem $\{P_n, n \geq 1\}$ is relatively compact. Let Q be a weak accumulation point of P_n, $P_{n_\ell} \rightsquigarrow Q$ along some subsequence.

Now, since $\int x^{2k} dP_n = m_{2k} + o(1) \leq c_k$ for some c_k, the function $x \mapsto x^k$ is uniformly integrable for $(P_n)_{n\geq 1}$. It implies that $\int x^k dQ(x) = \lim_{\ell\to\infty} \int x^k dP_{n_\ell}$. However, by assumption, the latter is equal to $\int x^k dP(x)$. Since the law of P is uniquely determined by its moments, we have $Q = P$. □

If X is a Poisson random variable with intensity $\lambda > 0$, there is a variant of this method. For integer $k \geq 1$, the k-th factorial moment of X has a simple expression: $\mathbb{E}[(X)_k] = \lambda^k$. Hence, in order to prove that $(X_n)_{n\in\mathbb{N}}$ converges weakly to X is sufficient to show that for all integer $k \geq 1$, $\lim_n \mathbb{E}[(X_n)_k] = \lambda^k$.

There are many drawbacks to this method. First, the random variable X_n needs to have finite moments of any order for all n large enough. Secondly, the computation of moments can be tedious. This method is usually used when no other method actually works and we shall not use it here.

2.2.2 Total variation distance and coupling

The *total variation distance* between two probability measures P and Q on a common σ-field $(S, \mathcal{S})$ is

$$d_{TV}(P,Q) = \sup_{A\in\mathcal{S}} |P(A) - Q(A)|\,.$$

Since $P(A^c) - Q(A^c) = -(P(A) - Q(A))$, we note that the absolute value can be removed in the definition. If S is a countable set, the supremum is reached for $A = \{x \in S : P(x) \geq Q(x)\}$. We have $P(A) - Q(A) = \sum_{x\in A} |P(x) - Q(x)|$ and $P(A^c) - Q(A^c) = -\sum_{x\in A^c} |P(x) - Q(x)|$. Since $P(A^c) - Q(A^c) = -(P(A) - Q(A))$, we get the simple formula:

$$d_{TV}(P,Q) = \frac{1}{2}\sum_{x\in S} |P(x) - Q(x)|\,.$$

A *coupling* of two probability measures P and Q on $(S, \mathcal{S})$ is a probability measure Π on $(S^2, \mathcal{S}^2)$ such that $P = \Pi\pi_1^{-1}$ and $Q = \Pi\pi_2^{-1}$, where $\pi_1(x,y) = x$, $\pi_2(x,y) = y$ for $(x,y) \in S^2$. In a more probabilistic rephrasing, a coupling of two probability measures P and Q is the distribution of a pair of random variables (X,Y) on S^2 such that X has law P and Y has law Q. For example the product measure $P \otimes Q$ is a coupling of P and Q. For an introduction to coupling, we refer to Lindvall (1992).

Lemma 2.8 (Coupling inequality) *Let P and Q be two probability measures on a common σ-field $(S, \mathcal{S})$. For any coupling (X,Y) of P and Q, we have*

$$d_{TV}(P,Q) \leq \mathbb{P}(X \neq Y).$$

Proof. For $A \in \mathcal{S}$, we write,

$$P(A) - Q(A) = \mathbb{E}\left[\mathbf{1}(X \in A) - \mathbf{1}(Y \in A)\right] = \mathbb{E}\left[(\mathbf{1}(X \in A) - \mathbf{1}(Y \in A))\mathbf{1}(X \neq Y)\right] \leq \mathbb{E}\left[\mathbf{1}(X \neq Y)\right].$$

□

The coupling inequality calls for a converse statement.

Theorem 2.9 (Maximal coupling) *Let P and Q be two probability measures on a common σ-field $(S, \mathcal{S})$. There exists a coupling (X, Y) of P and Q such that*

$$d_{TV}(P, Q) = \mathbb{P}(X \neq Y).$$

Proof. Consider the measure $\lambda = P + Q$, we denote by $f = dP/d\lambda$ and $g = dQ/d\lambda$ the Radon-Nikodym derivatives of P and Q with respect to λ. Considering the set $A = \{x \in S : f(x) \geq g(x)\}$. We deduce as above that

$$d_{TV}(P, Q) = \int_A (f - g) d\lambda = \frac{1}{2} \int |f - g| \, d\lambda.$$

Now, writing $|f - g| = (f - f \wedge g) + (g - f \wedge g)$, we get

$$\frac{1}{2} \int |f - g| \, d\lambda = 1 - \int f \wedge g \, d\lambda.$$

Let $\gamma = \int f \wedge g \, d\lambda$. In order to prove the statement it is thus sufficient to find a coupling (X, Y) such that $\mathbb{P}(X = Y) \geq \gamma$. If P and Q are mutually singular measures, there is nothing to prove, indeed, $d_{TV}(P, Q) = 1$ and the product coupling $P \otimes Q$ achieves the bound. Assume otherwise that P and Q are not mutually singular, then $\gamma > 0$. We may also assume $\gamma < 1$ otherwise, $P = Q$ and the coupling (X, X) where X has law P achieves the bound. We define (X_1, Y_1, Z, U) a quadruple of independent random variables, X_1 has distribution $(f - f \wedge g)d\lambda/(1 - \gamma)$, X_2 has distribution $(g - f \wedge g)d\lambda/(1 - \gamma)$, Z has distribution $f \wedge g d\lambda/\gamma$, and U is a Bernoulli random variables with parameter $\mathbb{P}(U = 1) = \gamma$. Then we may define the coupling (X, Y) where $X = (1 - U)X_1 + UZ$ and $Y = (1 - U)X_2 + UZ$. We have $\mathbb{P}(X = Y) \geq \mathbb{P}(U = 1) = \gamma$. □

2.2.3 Basics of Stein's method

There is a powerful technique to compare a probability measure to another one. This method is called the *Stein's Method* by the name of its author. We will sketch briefly the general idea and then apply it to the Poisson distribution in the next paragraph. The seminal paper on the topic is Stein (1972). For an introduction, we refer the reader to Barbour and Chen (2005).

Let $(S, \mathcal{S})$ be a complete metric space. We consider two probability measures P and P_0 on $(S, \mathcal{S})$. Let $\mathcal{H}$ be a set of measurable functions from S to $\mathbb{R}$. We assume that all functions in $\mathcal{H}$ are P and P_0 integrable. The goal of Stein's method is to estimate the difference over all $h \in \mathcal{H}$,

$$\int h dP - \int h dP_0.$$

The measure P_0 is thought as being a good approximation of P and $\mathcal{H}$ is thought as a set of test functions. In most applications, we shall assume that

$$d_{\mathcal{H}}(P,Q) = \sup_{h\in\mathcal{H}} \left| \int h dP - \int h dQ \right|$$

is a distance on the set of probability measures on $(S,\mathcal{S})$. In this setting, the goal of Stein's method is to find good bounds for the distance $d_{\mathcal{H}}(P,P_0)$. For example, if $\mathcal{H} = \{\mathbf{1}_A : A \text{ measurable}\}$, then $d_{\mathcal{H}} = d_{TV}$ is the total variation distance:

$$d_{TV}(P,Q) = \sup_A |P(A) - Q(A)|.$$

If $S = \mathbb{R}$ and $\mathcal{H} = \{\mathbf{1}_{(-\infty,x]} : x \in \mathbb{R}\}$ then $d_{\mathcal{H}}$ is the Kolmogorov-Smirnov distance. If $S = \mathbb{R}$ and $\mathcal{H} = \{h : \mathbb{R} \to \mathbb{R}; \|h'\| \leq 1\}$, where $\|h\| = \sup_{x\in\mathbb{R}} |h(x)|$ then $d_{\mathcal{H}}$ is the L_1-Wasserstein distance.

We assume that there exists a set $\mathcal{F}$ of measurable $S \to \mathbb{R}$ functions, and a linear mapping $T : \mathcal{F} \to \mathcal{H}$ such that for all $h \in \mathcal{H}$, there exists a function $f = f_h \in \mathcal{F}$ such that

$$Tf = h - \int h dP_0$$

Then we obviously get

$$\int h dP - \int h dP_0 = \int T f_h dP \tag{2.5}$$

T is called a *Stein operator* of the measure of P_0 and f_h is the *Stein transform* of h. In particular, we note that for all $h \in \mathcal{H}$,

$$\int T f_h dP_0 = 0,$$

and

$$d_{\mathcal{H}}(P,P_0) = \sup_{h\in\mathcal{H}} \left| \int T f_h dP \right|.$$

There are general procedures to find Stein operators. The goal being to find an operator where we can estimate nicely $|\int T f dP|$. It is not in the scope of these notes to develop further in this direction.

We should however mention that if $P_0(dx) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ is the standard Gaussian distribution $\mathcal{N}(0,1)$ and all functions in $\mathcal{H}$ are bounded, then

$$Tf : x \mapsto f'(x) - x f(x)$$

is a Stein operator for P_0 and

$$f_h(x) = e^{x^2/2} \int_{-\infty}^{x} \left(h(t) - \int h dP_0 \right) e^{-t^2/2} dt.$$

This operator was the starting point of Stein's work and much can be said about it.

2.2.4 Stein's method for the Poisson distribution

Chen (1975) has found a Stein operator for the Poisson distribution. Recall that the Poisson distribution with intensity $\lambda \in \mathbb{R}_+$, Poi_λ, is the probability measure on $\mathbb{N}$ defined by, for $n \in \mathbb{N}$,

$$\mathrm{Poi}_\lambda(\{n\}) = e^{-\lambda}\frac{\lambda^n}{n!}.$$

To fit with the above framework, we consider the space $S = \mathbb{N}$ and define $\mathcal{H} = \{\mathbf{1}_A : A \subset \mathbb{N}\}$. Then, as already mentioned, $d_{\mathcal{H}} = d_{TV}$ is the total variation distance. Let $\mathcal{F} \approx \mathbb{R}^{\mathbb{N}}$ be the set of real bounded functions on $\mathbb{N}$, we define the operator from $\mathcal{F}$ to $\mathcal{F}$,

$$Tf : k \mapsto \lambda f(k+1) - kf(k).$$

It is easy to check that if $f \in \mathcal{F}$ and Y is a random variable with Poisson distribution with intensity λ, then

$$\mathbb{E}[Tf(Y)] = \mathbb{E}[\lambda f(Y+1) - Yf(Y)] = 0.$$

Moreover, for all $h \in \mathcal{F}$, there exists a unique $f = f_h$ such that $f(0) = 0$ and

$$Tf = h - \mathbb{E}[h(Y)].$$

Indeed, the sequence

$$\lambda f(n+1) = nf(n) + h(n) - \mathbb{E}[h(Y)]$$

is easily solved by recursion. We find

$$f(n+1) = \sum_{k=0}^{n} \frac{(n)_k}{\lambda^{k+1}} \left(h(n-k) - \mathbb{E}[h(Y)]\right) = \frac{n!}{\lambda^{n+1}} \sum_{k=0}^{n} \frac{\lambda^k}{k!} \left(h(k) - \mathbb{E}[h(Y)]\right).$$

For $h = \mathbf{1}_A$, we define the function $f_{\lambda,A} = f_{\mathbf{1}_A}$ that we shall often simply denote by f, we get

$$f(0) = 0 \quad \text{and} \quad f(n+1) = \frac{\mathrm{Poi}_\lambda(A \cap [0,n]) - \mathrm{Poi}_\lambda(A)\mathrm{Poi}_\lambda([0,n])}{\lambda \mathrm{Poi}_\lambda(\{n\})}.$$

Theorem 2.10 (Properties of Chen-Stein operator) *The function $f = f_{\lambda,A}$ has the following properties:*

(i) For any random variable X on $\mathbb{N}$:

$$\mathbb{E}\left[\lambda f(X+1) - Xf(X)\right] = \mathbb{P}(X \in A) - \mathrm{Poi}_\lambda(A).$$

(ii) $\sup_n |f(n)| \le \min\left(1, \sqrt{\frac{2}{e\lambda}}\right)$.

(iii) $\sup_n |f(n+1) - f(n)| \le \lambda^{-1}(1 - e^{-\lambda}) \le 1$.

Proof. Point (i) follows from (2.5). The proof of (ii)-(iii) is performed in (Barbour and Eagleson, 1983, lemma 4), we omit it. □

Corollary 2.11 (Distance to Poisson) *For any random variable X on $\mathbb{N}$ and $\lambda > 0$,*

$$\mathrm{d}_{TV}(\mathcal{L}(X), \mathrm{Poi}_\lambda) = \max_{A \subset \mathbb{N}} \mathbb{E}\left[\lambda f_{\lambda,A}(X+1) - X f_{\lambda,A}(X)\right].$$

In order to illustrate the strength of Stein's method, consider $(Y_1, \cdots, Y_n)$ a sequence of independent Bernoulli variable with $\mathbb{P}(Y_i = 1) = 1 - \mathbb{P}(Y_i = 0) = p_i$ and set $\lambda = \sum_{i=1}^n p_i$. We define $X = \sum_{i=1}^n Y_i$ and $X_i = \sum_{j \neq i} Y_j = X - Y_i$. We write

$$\begin{aligned} \lambda f(X+1) - X f(X) &= \sum_{i=1}^n p_i \left(f(X+1) - f(X_i + 1) \right) \\ &\quad + \sum_{i=1}^n (p_i - Y_i) f(X_i + 1) + \sum_{i=1}^n Y_i \left(f(X_i+1) - f(X) \right). \end{aligned} \tag{2.6}$$

From theorem 2.10(iii), $|f(X+1) - f(X_i+1)| \leq \lambda^{-1}(1 - e^{-\lambda}) Y_i$. We notice also that $Y_i f(X) = Y_i f(X_i + 1)$, and X_i and Y_i are independent variables. Hence, taking expectation,

$$\mathbb{E}[\lambda f(X+1) - X f(X)] \leq \lambda^{-1}(1 - e^{-\lambda}) \sum_{i=1}^n p_i^2 \leq \frac{\sum_{i=1}^n p_i^2}{\sum_{i=1}^n p_i}.$$

In conclusion, from corollary 2.11, we thus deduce that

$$\mathrm{d}_{TV}(\mathcal{L}(X), \mathrm{Poi}_\lambda) \leq \frac{\sum_{i=1}^n p_i^2}{\sum_{i=1}^n p_i} \leq \max_{1 \leq i \leq n} p_i.$$

We have found without much effort a striking formula. If all p_i are equal and $0 \leq \lambda \leq n$, we obtain

$$\mathrm{d}_{TV}\left(\mathrm{Bin}\left(n, \frac{\lambda}{n}\right), \mathrm{Poi}_\lambda \right) \leq \frac{\lambda}{n}. \tag{2.7}$$

Exercise 2.12 *Let λ and μ be two positive real, show that $d_{TV}(\mathrm{Poi}_\lambda, \mathrm{Poi}_\mu) \leq |\lambda - \mu|$. (Hint: first bound $\mathrm{d}_{TV}\left(Bin\left(n, \frac{\lambda}{n}\right), Bin\left(n, \frac{\mu}{n}\right)\right)$ by using the coupling inequality).*

Exercise 2.13 *Let $\underline{\lambda}_n = (\lambda_1(n), \cdots, \lambda_n(n))$ be an array of positive real numbers satisfying (H_2'). Let G_n be a Chung-Lu graph with distribution $\mathcal{G}(\underline{\lambda}_n)$. Show that there exists a constant $c > 0$ such that for all integers n and any $i \in [n]$,*

$$d_{TV}(\deg(i; G_n), \mathrm{Poi}_{\lambda_i}) \leq c \frac{\lambda_i(n)}{n}.$$

(Hint: use exercises 1.10 and 2.12).

2.3 Cycle counts

2.3.1 Erdős-Rényi graphs

We now compute the limit of $X(H;G)$ when H is a cycle of length k. We start with the simpler case of Erdős-Rényi graphs.

Theorem 2.14 (Poisson asymptotic for cycles in Erdős-Rényi graphs) *Let $H = ([k], \{\{1,2\},\{2,3\},\cdots\{k$ be a cycle of length $k \geq 3$. Let $\lambda \in \mathbb{R}^+$ and for $n \geq 1$, let G_n be an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$. Then, with $\mu = \frac{\lambda^k}{2k}$,*

$$X(H;G_n) \xrightarrow{d} \mathrm{Poi}_\mu,$$

Proof. We have

$$X(H;G_n) = \sum_{F\in\mathcal{H}} Y_F \qquad \text{where} \qquad Y_F = \mathbf{1}(F \subset G_n),$$

and $\mathcal{H} = \{F : \text{ graph with } V_F \subset [n] \text{ and } F \simeq H\}$. Recall that $|\mathcal{H}| = (n)_k/(2k)$. We define

$$X_F = \sum_{F'\in\mathcal{H}: F\cap F'=\emptyset} Y_{F'}.$$

Let $f = f_{\mu,A}$ be as in theorem 2.10 and $\mu_n = \mathbb{E}X(H;G_n) = |\mathcal{H}|p_n$ where $p_n = \mathbb{P}(H \subset G_n) = (\lambda/n)^k$. As in (2.6), we write

$$\begin{aligned}\mu f(X+1) - Xf(X) &= (\mu - \mu_n)f(X+1) + \sum_{F\in\mathcal{H}} p_n\left(f(X+1) - f(X_F+1)\right) \qquad (2.8)\\ &\quad + \sum_{F\in\mathcal{H}}(p_n - Y_F)f(X_F+1) + \sum_{F\in\mathcal{H}} Y_F\left(f(X_F+1) - f(X)\right).\end{aligned}$$

By theorem 2.10(ii) and proposition 2.1, the first term of (2.8) goes to 0 uniformly over the choice of A. For the second term, we notice that $X - X_F = \sum_{F':F'\cap F\neq\emptyset} Y_{F'}$. Note also that for $F \in \mathcal{H}$, by construction

$$|\{F' \in \mathcal{H} : F' \cap F \neq \emptyset\}| \leq k(n-1)_{k-1} = 2k^2 n^{-1}|\mathcal{H}|.$$

Indeed, to each element in $\{F' \in \mathcal{H} : F' \cap F \neq \emptyset\}$ we may associate injectively, one element in F and $k-1$ distinct elements in $[n]$. Thus, by theorem 2.10(iii),

$$\begin{aligned}\mathbb{E}\sum_{F\in\mathcal{H}} p_n\left(f(X+1) - f(X_F+1)\right) &\leq \sum_{F\in\mathcal{H}} p_n \sum_{F':F'\cap F\neq\emptyset} \mathbb{P}(F' \subset G_n)\\ &\leq p_n^2|\mathcal{H}|^2 2k^2 n^{-1}\\ &= 2k^2\mu_n^2 n^{-1}.\end{aligned}$$

It follows that the second term of (2.8) goes to 0 (uniformly over the choice of A). The event $\{F \subset G_n\}$ is measurable with respect to the filtration generated by the events $(\{\{i,j\} \in E\}, i,j \in F)$, while X_F is measurable with respect to the filtration generated by the events $(\{\{i,j\} \in E\}, i,j \in [n] \backslash F)$. Hence, the variables Y_F and X_F are independent, it follows

$$\mathbb{E} \sum_{F \in \mathcal{H}} (p_n - Y_F) f(X_F + 1) \quad = \quad 0.$$

For the last term of (2.8), we note that

$$Y_F(X - X_F - 1) \quad = \quad Y_F \sum_{F' \neq F, F' \cap F \neq \emptyset} Y_{F'}.$$

We obtain by theorem 2.10(iii), with $c = \lambda^{-1}(1 - e^{-\lambda})$,

$$\begin{aligned} \sum_{F \in \mathcal{H}} Y_F \left(f(X_F + 1) - f(X) \right) &\leq c \sum_{F \in \mathcal{H}} \sum_{F' \neq F, F' \cap F \neq \emptyset} Y_{F \cup F'} \\ &= c \sum_L X(L; G_n), \end{aligned}$$

where the sum is over all equivalence classes of graphs L such that $L \simeq F \cup F'$ with $F, F' \in \mathcal{H}$, $F' \neq F$ and $F' \cap F \neq \emptyset$. There is a finite number of such classes. Fix such a graph $L = F \cup F'$. If F and F' have 1 vertex in common, then L has is a union of two cycles glued at a single vertex. In such case, $\mathrm{exc}(L) = 1$ and by proposition 2.1, we get for some new constant c,

$$\mathbb{E} X(L; G_n) \leq c n^{-1}. \tag{2.9}$$

Otherwise, L has a subgraph L' which is formed by a cycle and a line with $1 \leq \ell \leq k - 1$ edges, connecting two vertices the cycle. In such case, $\mathrm{exc}(L') = 1$. Since $X(L; G_n) \leq X(L'; G_n)$ (or using $\mathrm{exc}(L) \geq \mathrm{exc}(L')$, see exercice 1.3), we may again apply proposition 2.1 : for some new constant c, (2.9) still holds.

So finally the fourth term of (2.8) goes to 0 (uniformly over the choice of A). We may then conclude by applying corollary 2.11.

$\square$

We note that in the proof of theorem 2.14, we could have given an upper bound for $\mu f(X + 1) - X f(X)$ as a function of k and n. We may obtain, for some constant $C > 0$ independent of A,

$$\mu f(X + 1) - X f(X) \leq (Ck)^k n^{-1}.$$

Then, from corollary 2.11, we get an explicit bound for $d_{TV}(\mathcal{L}(X(H; G_n)), \mathrm{Poi}_\mu)$. One of the strength of the Stein method is precisely to give explicit upper bounds for the rates of convergence. We will not however pursue seriously this goal here.

There is a multivariate version of the previous theorem.

Theorem 2.15 (Poisson asymptotic for joint cycles in Erdős-Rényi graphs) *For integers $k \geq 3$ and $3 \leq \ell \leq k$, let H_ℓ be a cycle of length ℓ. Let $\lambda \in \mathbb{R}_+$ and for $n \geq 1$, let G_n be an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$. Then, with $\mu_\ell = \frac{\lambda^\ell}{2\ell}$ and any $(a_1, \cdots, a_k) \in \{0,1\}^k$,*

$$\sum_{\ell=3}^{k} a_\ell X(H_\ell; G_n) \xrightarrow{d} \mathrm{Poi}_{\sum_{\ell=3}^{k} a_\ell \mu_\ell}.$$

Obviously, this result hints loudly that in fact $(X(H_3; G_n), \cdots, X(H_k; G_n))$ converges to $\bigotimes_{\ell=3}^{k} \mathrm{Poi}_{\mu_\ell}$. To prove this stronger result with Stein's method, we should define a Stein operator for Poisson compound distributions, we will not pursue this goal here. Another possibility would be to use a multivariate method of moments.

Proof of theorem 2.15. For $1 \leq \ell \leq k$, let $\mathcal{H}_\ell = \{F :$ graph with $V_F \subset [n]$ and $F \simeq H_\ell\}$, $Y_F = \mathbf{1}(F \subset G_n)$ and $\mathcal{H} = \cup_{\ell=3}^{k} \mathcal{H}_\ell$. We have $|\mathcal{H}_\ell| = (n)_\ell/(2\ell)$ and

$$X = \sum_{\ell=3}^{k} a_\ell X(H_\ell; G_n) = \sum_{\ell=3}^{k} a_\ell \sum_{F \in \mathcal{H}_\ell} Y_F = \sum_{F \in \mathcal{H}} a_F Y_F$$

where if $F \in \mathcal{H}_\ell$, $a_F = a_\ell$. As in the proof of theorem 2.14, for $F \in \mathcal{H}$, we define

$$X_F = \sum_{F' \in \mathcal{H} : F \cap F' = \emptyset} Y_{F'}.$$

Let $f = f_{\mu,A}$ be as in theorem 2.10, $\mu = \sum_{\ell=3}^{k} a_\ell \mu_\ell$ and $\mu_n = \mathbb{E} \sum_{\ell=3}^{k} a_\ell X(H_\ell; G_n) = \sum_{\ell=3}^{k} a_\ell |\mathcal{H}_\ell| p_{n,\ell}$ where $p_{n,\ell} = \mathbb{P}(H_\ell \subset G_n) = (\lambda/n)^\ell$. If $F \in \mathcal{H}_\ell$, we set $p_F = p_{n,\ell}$. We write,

$$\begin{aligned}\mu f(X+1) - X f(X) &= (\mu - \mu_n) f(X+1) + \sum_{F \in \mathcal{H}} a_F p_F \left(f(X+1) - f(X_F+1) \right) \\ &\quad + \sum_{F \in \mathcal{H}} a_F (p_F - Y_F) f(X_F + 1) + \sum_{F \in \mathcal{H}} a_F Y_F \left(f(X_F+1) - f(X) \right).\end{aligned}$$

The first term goes to 0 by proposition 2.1. As in the proof of theorem 2.14, for the second term, we use the identity $X - X_F = \sum_{F' : F' \cap F \neq \emptyset} Y_{F'}$ and for $F \in \mathcal{H}$,

$$|\{F' \in \mathcal{H}_\ell : F' \cap F \neq \emptyset\}| \leq k(n-1)_{\ell-1} = 2k\ell n^{-1} |\mathcal{H}_\ell|.$$

Then, by theorem 2.10(iii), $|f(x+1) - f(x)| \leq 1$ and

$$\begin{aligned}\mathbb{E} \sum_{F \in \mathcal{H}} a_F p_F \left(f(X+1) - f(X_F+1) \right) &\leq \sum_{F \in \mathcal{H}} a_F p_F \sum_{F' : F' \cap F \neq \emptyset} \mathbb{P}(F' \subset G_n) \\ &\leq \sum_{F \in \mathcal{H}} a_F p_F \left(\sum_{\ell=3}^{k} p_{n,\ell} 2k\ell n^{-1} |\mathcal{H}_\ell| \right) \\ &\leq 2k^2 \mu_n \left(\sum_{\ell=3}^{k} \mathbb{E} X(H_\ell; G_n) \right) n^{-1}.\end{aligned}$$

By proposition 2.1, the above expression goes to 0 as n goes to infinity. The remainder of the proof of theorem 2.14 carries over here also. □

Exercise 2.16 (Subgraph count for Chung-Lu graphs) *Let $P \in \mathcal{P}(\mathbb{R}_+)$ and $\underline{\lambda}(n) \in \mathbb{R}_+^n$ an array satisfying (H'_p), $p \geq 1$. G_n be a Chung-Lu graph with distribution $\mathcal{G}(\underline{\lambda}_n)$.*

1. *Let $H \in \mathbb{G}[k]$ with m edges, c elements in its automorphism group and max degree p. Show that, as n goes to infinity,*
$$\mathbb{E}X(H; G_n) \sim c^{-1} n^{-\mathrm{exc}(H)} (\mathbb{E}\lambda)^{-m} \prod_{i=1}^{k} \mathbb{E}\lambda^{\deg(i;H)}.$$
2. *Assume now that H is a cycle of length $k \geq 3$ and $p = 2$. We set $\mu = \frac{(\mathbb{E}\lambda^2)^k}{2k(\mathbb{E}\lambda)^k}$. Show that $X(H; G_n)$ converges weakly to Poi_μ.*

2.3.2 Configuration model

Theorem 2.14 has a natural analog in the configuration model. Let $P \in \mathcal{P}(\mathbb{Z}_+)$ be a probability measure on integers and for $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$ an array of integers such that for any n, $\sum_{i=1}^n d_i(n)$ is even. We may then consider a random graph G_n with distribution $\mathcal{G}(\mathbf{d}_n)$.

Theorem 2.17 (Poisson asymptotic for cycles in configuration model) *For integer $k \geq 3$, let $H_k = ([k], \{\{1,2\}, \{2,3\}, \cdots \{k,1\}\})$ be a cycle of length k, let $H_1 = (\{1\}, \{\{1,1\}\})$ be a single loop and let $H_2 = (\{1,2\}, \{\{1,2\}, \{1,2\}\})$ be a single multi-edge. Let $G_n \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_2). Then for all $k \geq 1$,*
$$X(H_k; G_n) \xrightarrow{d} \mathrm{Poi}_{\mu_k},$$
with $\mu_k = \frac{(\mathbb{E}(D)_2)^k}{2k(\mathbb{E}D)^k}$ and D has distribution P.

Proof. The proof follows the same strategy than theorem 2.14. For ease of notation, we fix $k \geq 1$, set $\mu = \mu_k$, $H = H_k$ and write d_i in place of $d_i(n)$. As in the proof of proposition 2.4, we define $Y(H; G_n)$ as the number of times that $H \subset G_n$, for $i \in [n]$, let $S = \sum_{i=1}^n d_i$. If $Y_F = Y(F; G_n)$ and $p_n(F) = \mathbb{E}[Y_F]$, we define
$$\mu_n = \sum_{F \in \mathcal{H}} p_n(F),$$
where as in the proof of theorem 2.14, $\mathcal{H} = \{F :\ \text{multigraph with } V_F \subset [n] \text{ and } F \simeq H\}$. We have $\mathbb{E}X(H; G_n) = \mu_n$. Let $f = f_{\mu,A}$ be as in theorem 2.10 and
$$X_F = \sum_{F' \in \mathcal{H} : F \cap F' = \emptyset} Y_{F'}.$$

We write

$$\begin{aligned}\mu f(X+1)-Xf(X) &= (\mu-\mu_n)f(X+1)+\sum_{F\in\mathcal{H}}p_n(F)\left(f(X+1)-f(X_F+1)\right) \qquad (2.10)\\ &\quad+\sum_{F\in\mathcal{H}}(p_n(F)-Y_F)f(X_F+1)+\sum_{F\in\mathcal{H}}Y_F\left(f(X_F+1)-f(X)\right).\end{aligned}$$

As shown in the proof of proposition 2.4,

$$\mu_n=\frac{1}{2k}\sum_{\tau}\frac{\prod_{i=1}^k(d_{\tau(i)})_2}{(\!(S)\!)_k}\to\mu,$$

where the sum is over the set of injective maps from $[k]$ to $[n]$. Hence the first term (2.10) goes to 0.

The argument used in the proof of theorem 2.14 carries over here as well for the second and last term of (2.10) with minor changes. More precisely, by theorem 2.10(iii), $|f(x+1)-f(x)|\leq 1$, we write

$$\begin{aligned}\sum_{F\in\mathcal{H}}p_n(F)\mathbb{E}\left(f(X+1)-f(X_F+1)\right) &\leq \sum_{F\in\mathcal{H}}p_n(F)\mathbb{E}\sum_{F':F\cap F'\neq\emptyset}Y_{F'}\\ &= \frac{1}{(2k)^2}\sum_{*}\frac{\prod_{i=1}^k(d_{\tau(i)})_2(d_{\tau'(i)})_2}{(\!(S)\!)_k(\!(S)\!)_k},\end{aligned}$$

where the sum is over all pairs (τ,τ') of injective maps $[k]\to[n]$ such that the images of τ and τ' have a non empty intersection. We set

$$M(n)=\max(d_1,\cdots,d_n).$$

Since the image of such map (τ,τ') has cardinal at most $2k-1$, we have

$$\begin{aligned}\frac{1}{n^{2k}}\sum_{*}\prod_{i=1}^k(d_{\tau(i)})_2(d_{\tau'(i)})_2 &\leq \frac{M(n)^2}{n}\frac{1}{n^{2k-1}}\sum_{1\leq i_1,\cdots,i_{2k-1}\leq n}\prod_{\ell=1}^{2k-1}d_{i_\ell}(n)^2\\ &= \frac{M(n)^2}{n}\left(\frac{1}{n}\sum_{i=1}^n d_i^2\right)^{2k-1}\\ &= \frac{M(n)^2}{n}\left(\mathbb{E}D(n)^2\right)^{2k-1},\end{aligned}$$

where $D(n)$ has distribution $F_{\mathbf{d}_n}$. Now, from lemma 1.5, we have

$$M(n)^2=o(n). \qquad (2.11)$$

Moreover, from (2.1), $(\!(S)\!)_k(\!(S)\!)_k\sim n^{2k}(\mathbb{E}D)^{2k}$. It follows that the second term of (2.10) goes to 0.

We now turn to the last term of (2.10). Let E be the event that for all $F \in \mathcal{H}$, $Y_F \in \{0,1\}$. Note that if $Y_F \geq 2$, then $X(L;G_n) \geq 1$, where L is the multiset union of H and the edge $\{1,2\}$ (or the loop $\{1,1\}$ if $k=1$). The maximum degree of L is 4 and $\mathrm{exc}(L) = 1$. Then, if assumption (H_4) holds, we could apply corollary 2.6, and get, as $n \to \infty$,

$$\mathbb{P}(E^c) = \mathbb{P}(X(L;G_n) \geq 1) \leq \mathbb{E}X(H \cup H; G_n) \to 0. \tag{2.12}$$

With the sole assumption (H_2), the above equation (2.12) still holds. Indeed, if $m_i = \deg(i;L)$, from (2.4),

$$\mathbb{E}X(L;G_n) \leq \sum_{\tau} \frac{\prod_{i=1}^{k} (d_{\tau(i)}(n))_{m_i}}{(\!(S)\!)_{k+1}} \leq \frac{M(n)^2}{(\!(S)\!)_{k+1}} \left(\sum_{i=1}^{n} d_i(n)^2 \right)^k,$$

where the first sum is over all injective maps $\tau : [k] \to [n]$. Using (2.11) and (2.1), we deduce that (2.12) holds.

We have, by theorem 2.10(ii)-(iii), $|f(x)| \leq 1$ and $|f(x+1) - f(x)| \leq 1$,

$$\begin{aligned}
\mathbb{E} \sum_{F \in \mathcal{H}} Y_F \left(f(X_F + 1) - f(X) \right) &\leq 2\mathbb{P}(E^c) + \mathbb{E} \sum_{F \in \mathcal{H}} \sum_{F' \neq F, F' \cap F \neq \emptyset} Y_{F \cup F'} \\
&= 2\mathbb{P}(E^c) + \sum_{L} \mathbb{E}X(L;G_n),
\end{aligned}$$

where the sum is over all equivalence classes of graphs L such that $L \simeq F \cup F'$ with $F, F' \in \mathcal{H}$, $F' \neq F$ and $F' \cap F \neq \emptyset$. In the proof of theorem 2.14, we have seen all such L satisfies $\mathrm{exc}(L) \geq 1$. Fix such $L \in \widehat{\mathbb{G}}(k')$, it has k' vertices, $m' \geq k'+1$ edges and $m'_i = \deg(i;L)$, $\sum_i m'_i = 2m'$. Moreover, from (2.4),

$$\mathbb{E}X(L;G_n) \leq \mathbb{E}X(L';G_n) \leq \sum_{\tau} \frac{\prod_{i=1}^{k'} (d_{\tau(i)}(n))_{m'_i}}{(\!(S)\!)_{m'}} \leq \frac{M(n)^2}{(\!(S)\!)_{m'}} \left(\sum_{i=1}^{n} d_i(n)^2 \right)^{m'-1},$$

where the first sum is over all injective maps $\tau : [k'] \to [n]$. We may then conclude by a new application of lemma 1.4-1.5 that the above expression goes to 0. It follows that the fourth term of (2.10) goes to 0.

For the third term of (2.10), a new difficulty arises compared to the proof of theorem 2.14, X_F and Y_F are no longer independent. We should prove

$$\mathbb{E} \sum_{F \in \mathcal{H}} (p_n(F) - Y_F) f(X_F + 1) \to 0.$$

From (2.12), we find

$$\sum_{F \in \mathcal{H}} \sum_{k \geq 2} k \mathbb{P}(Y_F = k) \to 0 \text{ or equivalently } \sum_{F \in \mathcal{H}} (p_n(F) - \mathbb{P}(Y_F = 1)) \to 0.$$

By theorem 2.10(ii), $|f(x)| \le 1$. Hence, in order to prove that the third term goes to 0, it is sufficient to prove that

$$\mathbb{E} \sum_{F \in \mathcal{H}} p_n(F) \left(\mathbb{E} f(X_F + 1) - \mathbb{E}[f(X_F+1) | Y_F \ge 1] \right) \to 0. \tag{2.13}$$

We will use a coupling argument. Let σ be the uniform matching of $\Delta = \{(i,j) : i \in [n], 1 \le j \le d_i\}$ that matches the half-edges of G_n. Let $x \neq y \in \Delta$. The *switch* of σ at (x,y) is the matching σ' such that $\sigma'(x) = y$, $\sigma'(\sigma(x)) = \sigma(y)$ while $\sigma'(z) = \sigma(z)$ for all $z \notin \{x, y, \sigma(x), \sigma(y)\}$ (see figure 2.1). Note that, since σ is a uniform matching, the switch of σ at (x,y) is a random matching sampled uniformly among all matchings $m \in M(\Delta)$ such that $m(x) = y$.

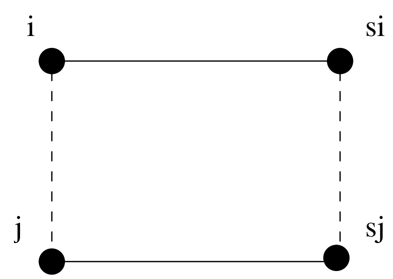

Figure 2.1: A switch : σ is plain and the switch of σ is dashed.

Similarly, let $\{\sigma(i,j) : i \in V_F, 1 \le j \le d_i\}$, where $V_F = \{i_1, \cdots, i_k\}$ is the vertex set of F (see figure 2.2). The law of G_n given $\{Y_F \ge 1\}$ is realized by taking independently, for $1 \le \ell \le k$, a distinct pair (j_ℓ, j'_ℓ) uniformly distributed on $\{1, \cdots, d_{i_\ell}\}$ and perform a switch of σ at $((i_1, j_1), (i_2, j'_2))$, then at $((i_2, j_2), (i_3, j'_3))$, and we continue up to $((i_k, j_k), (i_1, j'_1))$. (In this construction, we implicit assume that $i_\ell \ge 2$, otherwise, $Y_F = 0$). We define $\tilde{\sigma}$ as the corresponding matching and $\tilde{G}_n \in \widehat{\mathbb{G}}(\mathbf{d})$ is the multi-graph associated to $\tilde{\sigma}$. We set $\mathcal{H}_F = \{F' \in \mathcal{H} : F' \cap F = \emptyset\}$, $\tilde{Y}_{F'} = Y(F'; \tilde{G}_n)$ and

$$\tilde{X}_F = \sum_{F' \in \mathcal{H}_F} Y_{F'}.$$

Then, by theorem 2.10(ii), it follows that

$$\mathbb{E} \sum_{F \in \mathcal{H}} p_n(F) \left(\mathbb{E} f(X_F + 1) - \mathbb{E}[f(X_F+1) | Y_F \ge 1] \right) \le 2 \sum_{F \in \mathcal{H}} p_n(F) \mathbb{P}\left(X_F \neq \tilde{X}_F \right).$$

By construction, σ and $\tilde{\sigma}$ may only differ on the half-edges involved in the switches

$$\Delta_0 = \{(i_\ell, j_\ell), (i_\ell, j'_\ell), \sigma(i_\ell, j'_\ell), \sigma(i_\ell, j'_\ell)\}.$$

Also note that $\tilde{X}_F \ge X_F$ and the inequality is strict only if one of the switch, say (x,y), creates a new cycle $[n] \backslash V_F$ which contains the new edge formed by the half-edges $x' = \sigma(x)$ and $y' = \sigma(y)$. In such case, the half-edges x' and y' are part in G_n of a subgraph formed with half-edges in

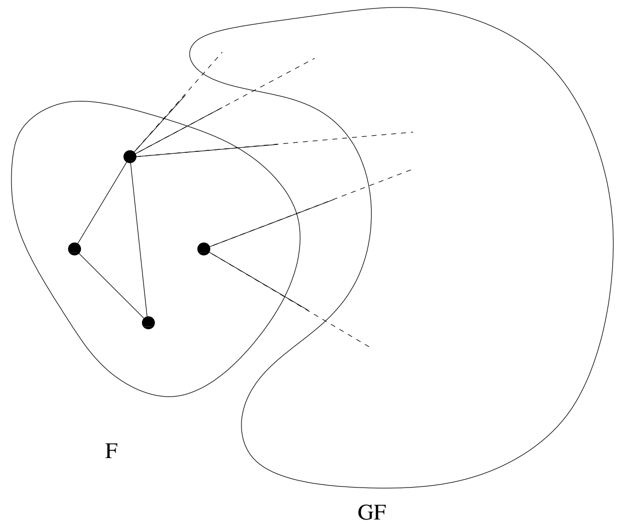

Figure 2.2: F and $G_n\backslash F$

$\Delta\backslash\Delta_0$ and isomorphic to a line of length k. From the union bound, this probability is bounded by

$$k\sum_{\tau}\frac{\prod_{\ell=1}^{k-2}(d_{\tau(\ell)})_2}{(\!(S-4k)\!)_{k-1}},$$

where the sum is over all injective maps $[k-2]\to[n]\backslash V_F$. The term k in front comes from the possible pairs (x,y) involved in the switch. The term $S-4k$ comes from the fact the half-edges in $\Delta\backslash\Delta_0$ are uniformly matched and $|\Delta_0|\le 4k$. The above is bounded by

$$2k\sum_{i_1\le\cdots\le i_{k-2}}\frac{\prod_{\ell=1}^{k-2}(d_{\tau(\ell)})_2}{(\!(S-2k)\!)_{k-1}}=\frac{2kn^{k-2}}{(\!(S-2k)\!)_{k-1}}\left(\frac{1}{n}\sum_{i=1}^{n}(d_i)_2\right)^{k-2}.$$

From (2.1), $(\!(S-2k)\!)_{k-1}\sim(\mathbb{E}D)^{k-1}n^{k-1}$. Using (H_p), we deduce that the above term is bounded by c/n for some constant $c=c(k)$ independent of F. This concludes the proof of (2.13). We may then conclude by applying corollary 2.11. □

Again, there is a multivariate version of the previous theorem.

Theorem 2.18 (Poisson asymptotic for joint cycles in configuration model) *For integers $k\ge 1$, let $H_1=(\{1\},\{\{1,1\}\})$ be a single loop, $H_2=(\{1,2\},\{\{1,2\},\{1,2\}\})$ be a single multi-edge and for $3\le\ell\le k$, let $H_\ell=([\ell],\{\{1,2\},\{2,3\},\cdots\{\ell,1\}\})$ be a cycle of length ℓ. Let $G_n\overset{d}{\sim}\widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_2). Then for any $(a_1,\cdots,a_k)\in\{0,1\}^k$,*

$$\sum_{\ell=1}^{k}a_\ell X(H_\ell;G_n)\xrightarrow{d}\mathrm{Poi}_{\sum_{\ell=1}^{k}a_\ell\mu_\ell},$$

with $\ell \geq 1$, $\mu_\ell = \frac{(\mathbb{E}(D)_2)^\ell}{2l(\mathbb{E}D)^\ell}$ *and* D *has distribution* P.

Proof. The proof is an extension of theorem 2.17 and follows the same strategy than theorem 2.15. With the notation of the proof of theorem 2.15, we have

$$X = \sum_{F \in \mathcal{H}} a_F Y_F,$$

and

$$\begin{aligned} \mu f(X+1) - Xf(X) &= (\mu - \mu_n) f(X+1) + \sum_{F \in \mathcal{H}} a_F p_n(F) \left(f(X+1) - f(X_F+1) \right) \\ &\quad + \sum_{F \in \mathcal{H}} a_F (p_n(F) - Y_F) f(X_F+1) + \sum_{F \in \mathcal{H}} a_F Y_F \left(f(X_F+1) - f(X) \right). \end{aligned}$$

The first, second and last term are treated as in the proof of theorem 2.15. For the third term, the argument used in theorem 2.17 works. We leave the details to the reader. □

2.4 Graphs with given degree sequence

Theorem 2.18 and its variants have important consequences on the labeled graphs with given degree sequence. Recall that a degree sequence $(d_1, \cdots, d_n)$ is graphic is there exists a graph G in $\mathbb{G}(n)$ such that for all $i \in [n]$, $\deg(i; G) = d_i$. As usual, we consider P, a probability measure on $\mathbb{Z}_+$.

Lemma 2.19 (Asymptotic graphic sequence) *Let* $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$ *be an array of integers such that for any* n, $\sum_{i=1}^n d_i(n)$ *is even and* (H_2) *holds. Then, for all* n *large enough,* $(d_1(n), \cdots, d_n(n))$ *is graphic.*

Proof. Let G_n be a random multigraph with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$. We have

$$\mathbb{P}(G_n \in \mathbb{G}(\mathbf{d})) = \mathbb{P}(X(H_1; G_n) + X(H_2; G_n) = 0).$$

Then from theorem 2.18,

$$\lim_n \mathbb{P}(G_n \in \mathbb{G}(\mathbf{d})) = e^{-\frac{\mathbb{E}(D)_2}{2\mathbb{E}D} - \frac{(\mathbb{E}(D)_2)^2}{4(\mathbb{E}D)^2}} > 0. \tag{2.14}$$

It implies in particular that $\mathbb{G}(\mathbf{d})$ is not empty and hence $\mathbf{d}_n$ is a graphic for all n large enough. □

Lemma 2.19 is a nice instance of the *probabilistic method* : we have used random variables to deduce the existence of a deterministic object. We refer to Alon and Spencer (2008) for a beautiful account of this method. The next theorem implies that the configuration model is a powerful tool to analyze the probability measure $\mathcal{G}(\mathbf{d}_n)$. The original proof of the next result can be found in (Janson, 2009, theorem 1.1).

Theorem 2.20 (Contiguity of $\widehat{\mathcal{G}}(\mathbf{d}_n)$ and $\mathcal{G}(\mathbf{d}_n)$) *Let $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$ be an array of integers such that for any n, $\sum_{i=1}^n d_i(n)$ is even and (H_2) holds. For $n \in \mathbb{N}$, let A_n be a subset of $\hat{\mathbb{G}}(n)$. We denote by $\hat{G}_n$ a random multigraph with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$ and, if $\mathbf{d}_n$ is graphic, by G_n a random graph with distribution $\mathcal{G}(\mathbf{d}_n)$. We assume that*

$$\lim_{n\to\infty} \mathbb{P}(\hat{G}_n \in A_n) = 1.$$

Then

$$\lim_{n\to\infty} \mathbb{P}(G_n \in A_n) = 1.$$

Proof. By theorem 2.18, $\liminf_n \mathbb{P}(X(H_1;\hat{G}_n) + X(H_2;\hat{G}_n) = 0) > 0$. Hence

$$\lim_n \mathbb{P}(\hat{G}_n \in A_n | \hat{G}_n \in \mathbb{G}(n)) = 1.$$

Now, by lemma 1.6, the distribution of $\hat{G}_n$ given $\{\hat{G}_n \in \mathbb{G}(n)\}$ has the same distribution than G_n. The statement follows. □

In the sequel, we will use repeatedly theorem 2.20. For example, it implies that the statement in probability of Corollary 2.6 holds with $\mathcal{G}(\mathbf{d}_n)$ replaced by $\widehat{\mathcal{G}}(\mathbf{d}_n)$ provided that (H_2) holds.

There is also an important combinatoric consequence of the above argument in terms of counting the cardinality of $\mathbb{G}(\mathbf{d})$, the set of graphs on $[n]$ with degree sequence $\mathbf{d}$.

Theorem 2.21 (Asymptotic number of graphs with given degree sequence) *Let $\mathbf{d}_n = (d_1(n), \cdots, d_n(n))$ be an array of integers such that for any n, $S_n = \sum_{i=1}^n d_i(n)$ is even and (H_2) holds. Then, as n goes to infinity,*

$$|\mathbb{G}(\mathbf{d}_n)| \sim \sqrt{2} e^{-\frac{\mathbb{E}(D)_2}{2\mathbb{E}D} - \frac{(\mathbb{E}(D)_2)^2}{4(\mathbb{E}D)^2}} \frac{\left(S_n e^{-1}\right)^{\frac{S_n}{2}}}{\prod_{i=1}^n (d_i)!}.$$

For d-regular graph, the above theorem specializes to a nice formula.

Corollary 2.22 (Asymptotic number of regular graphs) *Let $d \geq 2$. For integer n, let $\mathbb{G}(n,d)$ denote the (possibly empty) set of d-regular graphs on $[n]$. Then for dn even and n going to infinity,*

$$|\mathbb{G}(n,d)| \sim \sqrt{2} e^{-\frac{(d^2-1)}{4}} \left(\frac{d^{d/2}}{e^{d/2} d!}\right)^n n^{dn/2}.$$

Proof of theorem 2.21. For $n = 2m-1$ odd, let $n!! = n(n-2)\cdots 1 = \frac{(2m)!}{2^m m!}$. We consider the configuration model $\widehat{\mathcal{G}}(\mathbf{d}_n)$. Let $\Delta = \{(i,j) : i \in [n], 1 \leq j \leq d_i\}$ be the set of half-edges. For each matching σ of Δ, we denote by $G(\sigma)$ the d-regular multigraph on $[n]$ associated to σ. The number of possible matchings of Δ is

$$(S_n - 1)!! = \frac{(S_n)!}{2^{\frac{S_n}{2}} \left(\frac{S_n}{2}\right)!}.$$

By lemma 1.6, each graph in $\mathbb{G}(\mathbf{d})$ can be obtained by $\prod_{i=1}^n (d_i)!$ different matchings. We thus get

$$|\mathbb{G}(\mathbf{d})| = \frac{1}{\prod_{i=1}^n (d_i)!} \sum_{\sigma} \mathbf{1}(G(\sigma) \in \mathbb{G}(n))) = \frac{(S_n)!}{2^{\frac{S_n}{2}} \left(\frac{S_n}{2}\right)! \prod_{i=1}^n (d_i)!} \mathbb{P}(G_n \in \mathbb{G}(n)),$$

where the sum is over all matchings of Δ and G_n is a random multigraph with distribution $\widehat{\mathcal{G}}(n, d)$. Now, we use the identity $\mathbb{P}(G_n \in \mathbb{G}(n)) = \mathbb{P}(X(H_1; G_n) + X(H_2; G) = 0)$. It remains to apply (2.14) and use Stirling's formula, $n! \sim_n \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$. $\square$

Chapter 3

Local weak convergence

3.1 Weak convergence in metric spaces

In this paragraph, we recall some facts on weak convergence in metric spaces. For proofs and details on the weak convergence, we refer the reader to Chapter 1 in Billingsley (1999). Let S be a metric space endowed with its Borel σ-algebra, $\mathcal{S}$.

Theorem 3.1 (Characterization of measures) *Probability measures P and Q on $(S, \mathcal{S})$ coincide if and only if for all bounded continuous functions $f : S \mapsto \mathbb{R}$, $\int f dP = \int f dQ$.*

The proof of this theorem will be included in the forthcoming theorem 3.2.

A sequence of probability measures $(P_n)_{n\in\mathbb{N}}$ on S converges weakly to a probability measure P if for every bounded continuous function f, $\int f dP_n$ converges to $\int f dP$. This convergence is usually denoted by $P_n \rightsquigarrow P$. With a slight abuse of notation, if X_n is a random variable with law P_n and X with law P, we shall also write $X_n \stackrel{d}{\to} X$.

Theorem 3.2 (Portemanteau theorem) *The following conditions are equivalent.*

(i) $P_n \rightsquigarrow P$.

(ii) $\int f dP_n \to \int f dP$ *for all bounded, uniformly continuous functions* f.

(iii) $\limsup P_n(F) \leq P(F)$ *for all closed sets* F.

(iv) $\liminf P_n(G) \geq P(G)$ *for all open sets* G.

(v) $\lim P_n(A) = P(A)$ *for all* $A \in \mathcal{S}$ *such that* $P(\partial A) = 0$.

Proof. Let $d(x,y)$ be the distance in S. $(i) \Rightarrow (ii)$ is trivial. For $(ii) \Rightarrow (iii)$, let $\varepsilon > 0$, F be a closed set, $F^\varepsilon = \{x \in S : d(x,F) \leq \varepsilon\}$, and

$$f(x) = \min(0, 1 - \varepsilon^{-1} d(x,F)).$$

This function is bounded and uniformly continuous because $|f(x) - f(y)| \leq \varepsilon^{-1} d(x,y)$. Moreover for every $x \in S$,

$$\mathbf{1}_F(x) \leq f(x) \leq \mathbf{1}_{F^\varepsilon}(x).$$

Indeed if $x \in F$, then $d(x,F) = 0$ and $f(x) = 1$, while if $x \notin F^\varepsilon$, $d(x,F) \geq \varepsilon$ and $f(x) = 0$. It follows that

$$P_n(F) \leq \int f dP_n \leq P_n(F^\varepsilon).$$

By assumption (ii), letting n tend to infinity, it implies that

$$\limsup P_n(F) \leq \int f dP_n \leq P(F^\varepsilon).$$

Since F is closed, as ε goes to 0, $\mathbf{1}_{F^\varepsilon \backslash F}(x)$ converges to 0 for all $x \in S$. Thus by the dominated convergence theorem, $\lim_{\varepsilon \downarrow 0} \int \mathbf{1}_{F^\varepsilon \backslash F} dP = 0$. It follows that $\lim_{\varepsilon \downarrow 0} P(F^\varepsilon) = P(F)$ and (iii) follows. The statements (iii) and (iv) are equivalent by complementation. To prove that (iii)&(iv) imply (v), let $\dot{A}$ and $\bar{A}$ denote the interior and closure of A. Assumption (iii) and (iv) imply

$$P(\dot{A}) \leq \liminf P_n(\dot{A}) \leq \liminf P_n(A) \leq \limsup P_n(A) \leq \limsup P_n(\bar{A}) \leq P(\bar{A}).$$

The extreme left hand and right hand side are equal because $P(\partial A) = 0$, and (v) follows. It remains to check that $(v) \Rightarrow (i)$. We may assume that $0 \leq f \leq 1$. Then from Fubini's theorem,

$$\int f dP = \int_0^1 P(\{x : f(x) > t\}) dt,$$

and similarly for P_n. Since f is continuous, $\partial\{x : f(x) > t\} \subset \{x : f(x) = t\}$. The probability measure on $[0,1]$, $Q = P f^{-1}$ has at most a countable number of atoms. Hence, from (v), for almost all $t \in [0,1]$,

$$\lim_n P_n(\{x : f(x) > t\}) = P(\{x : f(x) > t\}).$$

It follows, by dominated convergence that

$$\lim_n \int P_n(\{x : f(x) > t\}) dt = \int P(\{x : f(x) > t\}) dt.$$

and (i) follows. □

Let Π be a collection of probability measures of measure on S. We say that Π is *tight* if for all $\varepsilon > 0$ there exists a compact set K such that for all $P \in \Pi$, $P(K) > 1 - \varepsilon$. The collection Π is *relatively compact* if for every sequence of elements (P_n) in Π, there exists a subsequence (P_{n_k}) and a probability measure Q such that $P_{n_k} \rightsquigarrow Q$. Prohorov's theorem states that the two notions are equivalent in complete separable metric spaces.

Theorem 3.3 (Prohorov) *If Π is tight then it is relatively compact. If $(S, \mathcal{S})$ is complete and separable, the converse also holds : if Π is relatively compact then it is tight.*

The most difficult part in the theorem is the first statement. The second statement implies in particular that a single probability measure is always tight.

Proof of the second statement of theorem 3.3. Consider a increasing sequence of open sets G_n covering S. Then for all $\varepsilon > 0$ there exists an n such that for all $P \in \Pi$, $P(G_n) > 1 - \varepsilon$. Indeed otherwise, for some $\varepsilon > 0$ there would exist a sequence $(P_n) \in \Pi$ such that $P_n(G_n) \leq 1 - \varepsilon$. Since G_n is increasing, that for all n_0 and $n \geq n_0$, $P_n(G_{n_0}) \leq 1 - \varepsilon$. By relative compactness, there would exist a measure Q and a subsequence (n_k), such that $P_{n_k} \rightsquigarrow Q$. We deduce from theorem 3.2 that for all n_0, $Q(G_{n_0}) \leq 1 - \varepsilon$. Letting n_0 go to infinity, this leads to a contradiction since $G_n \uparrow S$ and $Q(S) = 1$.

Now, by separability, for every integer k, there exists a collection $B_{k1}, B_{k2}, \cdots$ of open balls of radius $1/k$ covering S. From what precedes, there exists n_k such that $P(\cup_{i \leq n_k} B_{ki}) > 1 - \varepsilon/2^k$ for all $P \in \Pi$. By completeness, the closure K of $\cap_{k \geq 1} \cup_{i \leq n_k} B_{ki}$ is a compact set. Finally from the union bound, $P(K) > 1 - \varepsilon$ for all $P \in \Pi$. $\square$

3.2 The space of rooted unlabeled networks

In the previous chapter, we have counted subgraphs in a random graph with a non-negative excess. A connected graph with excess -1 is a tree and we are now going to look at the subtrees of a random graph. From propositions 2.1, 2.4, the number of occurrences of a given tree in a random graph is of order of a magnitude its number of vertices. This motivates the introduction of rooted graphs.

Let Ω be a complete separable metric space with distance d_Ω. We shall consider networks (V, E, ω) with $\omega : V \cup E \to \Omega$.

A *rooted network* $G = (V, E, \omega, \emptyset)$ is the pair formed by a network (V, E, ω) and a distinguished vertex $\emptyset \in V$, called the root. A *rooted isomorphism* between two rooted networks is an isomorphism that takes the root of one to the root of the other. As for networks isomorphisms, we will also denote by "$\simeq$" the equivalence relation of rooted isomorphisms.

If $G = (V, E, \omega, \emptyset)$ is a rooted network, $[G]$ will denote the class of rooted graphs that are rooted isomorphic to G. With the terminology of graph theory, $[G]$ is an *unlabeled rooted network*.

Let $\mathbb{G}_*(\Omega)$ denote the set of all $[G]$, with G ranging over connected locally finite networks with mark space Ω. In other words, $\mathbb{G}_*(\Omega)$ is the set of rooted unlabeled connected locally finite networks with mark space Ω.

If $\Omega = \{1\}$, then we can identify, $\mathbb{G}_* := \mathbb{G}_*(\{1\})$ with the set of unlabeled locally finite rooted graphs. Similarly, if $\Omega = \mathbb{Z}_+ = \{0, 1, \cdots\}$, $\widehat{\mathbb{G}}_* := \mathbb{G}_*(\mathbb{Z}_+)$ is the set of rooted unlabeled connected locally finite multigraphs.

There is a natural metric on $\mathbb{G}_*$. First, let $G = (V, E, \omega)$ be a connected network. For any pair u, v in V, we define $d_G(u, v)$ as the infimum of the length of the paths from u to v. This is the distance induced by G on V. The ball of radius t and center u is

$$B_G(u,t) := \{v \in V : d_G(u,v) \le t\}.$$

For the rooted network $G = (V, E, \omega, \varnothing)$ and real $t > 0$, let $(G)_t$ denote the network whose vertex set is $B_G(\varnothing, t)$ and whose edge set consists of the edges of G that have both vertices in $B_G(\varnothing, t)$.

Consider two elements g_1 and g_2 in $\mathbb{G}_*(\Omega)$. There exists, for $i \in \{1, 2\}$, a network $G_i = (V_i, E_i, \omega_i, \varnothing_i)$ with $[G_i] = g_i$. Then, the distance between g_1 and g_2 is defined as $1/(1 + T)$, where

$$\begin{aligned} T \;=\; & \sup\{t > 0 : \text{ there exists a rooted isomorphism } \sigma \text{ from } (V_1, E_1, \varnothing_1)_t \text{ to } (V_2, E_2, \varnothing_2)_t \\ & \quad \text{and for all } v \in V_{(G_1)_t},\, e \in E_{(G_1)_t},\, d_\Omega(\omega_1(v), \omega_2(\sigma(v)) \le 1/t \;,\; d_\Omega(\omega_1(e), \omega_2(\sigma(e)) \le 1/t\big\}\,. \end{aligned}$$

Note that the value of T does not depend on the particular choice of the rooted network in the equivalence class. For the case of graphs $\mathbb{G}_*$ or multigraphs, $\widehat{\mathbb{G}}_*$ (or more generally for Ω discrete), the distance is equivalently defined as $1/(1 + T)$, where

$$T \;=\; \sup\{t > 0 : \text{ there exists a rooted isomorphism from } (G_1)_t \text{ to } (G_2)_t\}.$$

The next lemma follows from the mere definition but is essential.

Lemma 3.4 (Properties of $\mathbb{G}_*(\Omega)$) *The space $\mathbb{G}_*(\Omega)$ is separable and complete.*

Proof. We start with separable. Since Ω is separable, let $(x_n)_{n\ge1}$ be a dense collections of elements in Ω. For $n \ge 1$, consider the countable family $\mathcal{X}_n$ of rooted networks on $[n]$ rooted at 1 with mark space $(x_n)_{n\ge1}$. We define the countable family $\mathcal{X} = \cup_n \mathcal{X}_n$. Let $g \in \mathbb{G}_*(\Omega)$ with $G = (V, E, \omega, \varnothing)$ in the equivalence class of g, $[G] = g$. For any real $t > 0$, since G is locally finite, there exists an integer n such that $(G)_t$ has n vertices. Hence, for some $F \in \mathcal{X}_n \subset \mathcal{X}$, there exists a rooted isomorphism from $(V, E, \varnothing)_t$ to $(V_F, E_F, 1)_t$ which distorts the marks by a distance less than $1/t$. It follows that the distance between $[F]$ and $[G]$ is less than $1/(t + 1)$.

We now turn to $\mathbb{G}_*(\Omega)$ complete. Let $(g_n)_{n\ge1}$ be a Cauchy sequence in $\mathbb{G}_*(\Omega)$. We consider a sequence $(G_n)_{n\ge1}$ of elements in their equivalence class: $[G_n] = g_n$. We may assume that $V_{G_n} = V_n = \{1, \cdots, K_n\}$ and G_n rooted at 1. We set $G_n = (V_n, E_n, \omega_n, 1)$ and $H_n = (V_n, E_n, 1)$. By assumption, there is an increasing sequence $(n_t)_{t\in\mathbb{N}}$, such that for all $n \ge n_t$, $m \ge 0$, the distance between G_n and G_{n+m} is less than $1/(t + 1)$. In particular, for all $m \ge 0$, $(H_{n_t})_t$ and $(H_{n_t+m})_t$ are rooted isomorphic and the corresponding marks in G_{n_t} and G_{n_t+m} are within distance $1/t$. Let N_t be the number of vertices in $(H_{n_t})_t$, and assume for example that $\lim N_t = \infty$. We may then define iteratively a graph $H = (V, E, 1)$ with $V = \mathbb{N}$ rooted at 1 such that for all $t \ge 1$, $(H)_t \simeq (H_{n_t})_t$. It follows that $([H_n])_{n\ge1}$ converges to $[H]$ in $\mathbb{G}_*$.

Now, by construction, there is a rooted isomorphism σ_t from $(H)_t$ to $(H_{n_t})_t$ such that for any $v \in V$, $e \in E$, and t large enough, $v \in V_t$, $e \in E_t$ and the marks $\omega_{n_t}(\sigma_t(v))$, $\omega_{n_t}(\sigma_t(e))$ are Cauchy sequences. Since Ω is complete, they converge to say $\omega(v)$ and $\omega(e)$. This defines a limit network $G = (V, E, \omega, 1)$ and (g_n) converges to $[G]$ in $\mathbb{G}_*(\Omega)$. □

The next elementary lemma may be useful to prove tightness of sequence of probability measures in $\widehat{\mathbb{G}}_*$. For a finite rooted multigraph, $G = (V, E, \o)$ we set $|G| = |V| + |E|$ (beware that $|E|$ is here the cardinal of a multiset).

Lemma 3.5 (Criterion of compactness) *Let $h : \mathbb{N} \to \mathbb{N}$ be an increasing function. The set*

$$\mathcal{K} = \{g \in \widehat{\mathbb{G}}_* : \textit{if } [G] = g, \textit{ for all } t \geq 0 \ |(G)_t| \leq h(t)\}.$$

is compact.

Proof. For each $t \geq 1$, there is a finite number of equivalence classes of rooted multigraphs $F_{t,1}, \cdots, F_{t,n_t}$ such that $|F| \leq h(t)$. Therefore, the collection $A_{1,1}, \cdots, A_{t,n_t}$ where $A_{t,k} = \{[G] \in \widehat{\mathbb{G}}_* : (G)_t \simeq F_{t,k}\}$ is a finite covering of $\mathcal{K}$ of radius $1/(1+t)$. □

3.3 Converging graph sequences

In the above section, we have described a natural metric space for rooted connected networks. However, our prime interest in the preceding chapters was on networks not on rooted network. There is a way to lift the above setting to the case of unrooted and not necessarily connected networks. This is called the *local weak convergence*, a notion that was introduced and developed in Benjamini and Schramm (2001), Aldous and Steele (2004), Aldous and Lyons (2007). The word "local" stems for the fact that the metric is defined through a root, the term "weak" from the choice of a random root.

For ease of notation, we fix the mark space Ω and write $\mathbb{G}_*$ in place of $\mathbb{G}_*(\Omega)$. We introduce the Borel σ-algebra of $\mathbb{G}_*$ and define $\mathcal{P}(\mathbb{G}_*)$ as the set of probability measures on $\mathbb{G}_*$ and endow this space of measures with the topology of weak convergence. By lemma 3.4, $\mathbb{G}_*$ is a separable metric space (Polish space). It implies that $\mathcal{P}(\mathbb{G}_*)$ is also a Polish space. We are in the framework of the standard theory of weak convergence of probability measures, as in the preceding section 3.1.

To a finite network $G = (V, E, \omega)$, we can associate a probability measure $U(G)$ in $\mathcal{P}(\mathbb{G}_*)$ defined as the law of $[G(\o), \o]$, where $\o$ is a uniformly chosen vertex in V and, for $v \in V$, $G(v)$ denotes the sub-network of G spanned by the vertices in the connected component of v. In other words,

$$U(G) = \frac{1}{|V|} \sum_{v \in V} \delta_{[G(v),v]}.$$

where δ is the Dirac delta function.

Definition 3.6 (Converging graph sequence) *A sequence of finite networks* $(G_n)_{n\geq 1}$ *has random weak limit* $\rho \in \mathcal{P}(\mathbb{G}_*)$ *if* $U(G_n) \rightsquigarrow \rho$.

Not all probability measures $\rho \in \mathcal{P}(\mathbb{G}_*)$ can be random weak limits. Due to the uniform rooting, there should satisfy a form of stationarity. This is formalized by the notion of *unimodularity* which plays a crucial role in local weak convergence theory. Consider networks with two roots or distinguished vertices : $(G, ø, o)$ with $G = (V, E, \omega)$ and $ø, o \in V$. Then, the natural notion of equivalence classes is with respect to isomorphisms which preserves the two roots. Let $\mathbb{G}_{**}$ be the set of equivalence classes of locally finite connected networks with two roots. We endow $\mathbb{G}_{**}$ with the natural metric which generalizes directly the metric on $\mathbb{G}_*$. With a slight abuse of notation, if f is a function from $\mathbb{G}_{**}$ to $\mathbb{R}_+$ and (G, u, v) is in the equivalence class of $g \in \mathbb{G}_{**}$, $[G, u, v] = g$, we define $f(G, u, v) := f(g)$.

Definition 3.7 (Unimodularity) *A measure* $\rho \in \mathcal{P}(\mathbb{G}_*)$ *is unimodular if for all measurable non-negative functions* $f : \mathbb{G}_{**} \to \mathbb{R}_+$,

$$\mathbb{E}_\rho \sum_{v \in V_G} f(G, ø, v) = \mathbb{E}_\rho \sum_{v \in V_G} f(G, v, ø), \tag{3.1}$$

where under $\mathbb{P}_\rho$, $[G, ø]$ *has law* ρ.

Note that the fact the expectation could be infinite in the definition of unimodularity is not issue from Fubini-Tonnelli theorem. If $f(G, u, v)$ is thought as an amount of mass sent from u to v, the unimodularity is a *mass transport principle*.

Let G be a finite network. We notice that $U(G)$ is unimodular : indeed, if u and v are connected then $G(u) = G(v)$. It follows that

$$\begin{aligned} \mathbb{E}_{U(G)} \sum_{v \in V_{G(ø)}} f(G(ø), ø, v) &= \frac{1}{|V_G|} \sum_{u \in V_G} \sum_{v \in V_{G(u)}} f(G(u), u, v) \\ &= \frac{1}{|V_G|} \sum_{v \in V_G} \sum_{u \in V_{G(v)}} f(G(u), u, v) \\ &= \mathbb{E}_{U(G)} \sum_{v \in V_{G(ø)}} f(G(ø), v, ø). \end{aligned}$$

Lemma 3.8 (Random weak limits are unimodular) *Let* $\rho \in \mathcal{P}(\mathbb{G}_*)$. *Assume that there exists a sequence of finite networks* $(G_n)_{n\geq 1}$ *with random weak limit* ρ. *Then* ρ *is unimodular.*

Proof. We should prove that the set of unimodular measures is closed for the weak topology. Let ρ_n be a sequence of unimodular probability measures converging weakly to ρ. From Lusin's theorem, it is sufficient to check (3.1) for f continuous and such that both terms in (3.1) are finite. For $\tau > 0$, we define a function $f_\tau : \mathbb{G}_{**} \to \mathbb{R}_+$ by setting, with $g = [G, u, v]$, $f_\tau(g) = \tau \wedge f(g)$

if u and v are at distance less than τ in G and if there are less than τ vertices in $B_G(u,\tau)$. Otherwise, we set $f_\tau(G,u,v) = 0$.

Then, by construction, $[G,\varnothing] \mapsto \sum_{v\in V_G} f_\tau(G,\varnothing,v)$ is continuous and bounded by τ^2. The dominated convergence theorem implies that

$$\lim_{n\to\infty} \mathbb{E}_{\rho_n} \sum_{v\in V_G} f_\tau(G,\varnothing,v) = \mathbb{E}_\rho \sum_{v\in V_G} f_\tau(G,\varnothing,v)$$

and similarly for $\mathbb{E}_{\rho_n} \sum_{v\in V_G} f_\tau(G,v,\varnothing)$. Since ρ_n is unimodular, we get

$$\mathbb{E}_\rho \sum_{v\in V_G} f_\tau(G,\varnothing,v) = \mathbb{E}_\rho \sum_{v\in V_G} f_\tau(G,v,\varnothing),$$

It remains to let τ tend to infinity and apply the monotone convergence theorem. $\square$

We will see in the next chapters that surprisingly many functions are continuous with respect to the local weak convergence. The following criterion is quite convenient to prove unimodularity. It is called the involution invariance property.

Lemma 3.9 (Involution invariance) *Let $\rho \in \mathbb{P}(\mathbb{G}_*)$ and assume that* (3.1) *holds for all functions $f : \mathbb{G}_{**} \to \mathbb{R}_+$ such that $f(G,u,v) = 0$ unless $\{u,v\} \in E_G$. Then ρ is unimodular.*

Proof. It is sufficient to prove (3.1) holds for all functions such that $f(G,u,v) = 0$ unless $d_G(u,v) = \tau$ for some integer $\tau \geq 1$. Indeed any function can be written as a sum of such functions. We prove the property that (3.1) holds for all functions such that $f(G,u,v) = 0$ unless $d_G(u,v) = \tau$ by recursion on τ. The case $\tau = 1$ is the involution invariance. We now take a general $\tau \geq 2$. For integer $k \geq 1$, $\partial B_G(u,k) = B_G(u,k)\backslash B_G(u,k-1)$ is the set of vertices at distance k from $u \in V_G$. If $x \in \partial B_G(u,\tau)$, let $\pi(G,u,x) \geq 1$ be the number of geodesic paths from u to x. If $y \in \partial B_G(u,\tau-1)$, we denote by $\pi(G,u,x,y)$ the number of geodesic paths from u to x whose first visited vertex is y. By construction, if $x \in \partial B_G(u,\tau)$, we have the balance equation

$$\pi(G,u,x) = \sum_{y\in\partial B_G(u,\tau-1)} \pi(G,u,x,y). \tag{3.2}$$

Now consider a function such that $f(G,u,x) = 0$ unless $d_G(u,x) = \tau$ or equivalently $x \in \partial B_G(u,\tau)$. We define the function, for $y \in \partial B_G(u,\tau-1)$,

$$h(G,u,y) = \sum_{x\in\partial B_G(u,\tau)} f(G,u,x)\frac{\pi(G,u,x,y)}{\pi(G,u,x)}.$$

and $h(G,u,v) = 0$ if $v \notin \partial B_G(u,\tau-1)$. From (3.2), we find

$$\sum_{v\in V_G} h(G,u,v) = \sum_{y\in\partial B_G(u,\tau-1)} \sum_{x\in\partial B_G(u,\tau)} f(G,u,x)\frac{\pi(G,u,x,y)}{\pi(G,u,x)} = \sum_{v\in V_G} f(G,u,v).$$

This proves the recursion step. $\square$

3.4 Unimodular Galton-Watson trees

In the next sections we will be interested by proving the convergence of $U(G_n)$ where $(G_n)_{n\geq 1}$ is a sequence of graphs either sampled from the Erdős-Rényi law $\mathcal{G}(n, \lambda/n)$, from the configuration model $\widehat{\mathcal{G}}(\mathbf{d}_n)$, or from $\mathcal{G}(\mathbf{d}_n)$, uniform law on graphs with degree distribution $\mathbf{d}_n$. As we shall see, the unimodular limit will be supported on trees.

3.4.1 Galton-Watson trees

Let $\mathbb{N}^f = \cup_{k\geq 0}\mathbb{N}^k$, with the convention $\mathbb{N}^0 = \{\o\}$. For $k \geq 1$ and $\mathbf{i} = (i_1, \cdots, i_k) \in \mathbb{N}^k$, we call $(i_1, \cdots, i_{k-1}) \in \mathbb{N}^{k-1}$ the *ancestor* or *genitor* of $\mathbf{i}$. We consider a sequence $(N_\mathbf{i})$, $\mathbf{i} \in \mathbb{N}^f$, of integers. We define the set

$$V = \{\o\} \cup \{\mathbf{i} = (i_1, \cdots, i_k) \in \mathbb{N}^f : \text{ for all } 1 \leq \ell \leq k, 1 \leq i_\ell \leq N_{i_1,\cdots,i_{\ell-1}}\}. \tag{3.3}$$

Note that the ancestor of an element in V is in V. For $\mathbf{i} \in V$, we call the set $\{(\mathbf{i}, 1), \cdots, (\mathbf{i}, N_\mathbf{i})\}$, the set of *offsprings* of $\mathbf{i}$. Then, we define a rooted tree $T = (V, E, \o)$ by putting an edge between all vertices in V and their ancestors. In particular, if $\mathbf{i} \neq \o$, the degree of $\mathbf{i}$ in T is $\deg(\mathbf{i}; T) = N_\mathbf{i}+1$, and $\deg(\o; T) = N_\o$. The set of vertices in $V \cap \mathbb{N}^k$ is called the k-th generation vertices. The *descendants* of a given vertex $\mathbf{i} \in V$ are the vertices in

$$V_\mathbf{i} = V \cap \{(\mathbf{i}, \mathbf{j}) : \mathbf{j} \in \mathbb{N}^f\}.$$

Finally, we denote by $T_\mathbf{i}$ the subtree rooted at $\mathbf{i}$ of vertices in $V_\mathbf{i}$.

Let $P \in \mathcal{P}(\mathbb{Z}_+)$ be a probability distribution on $\mathbb{Z}_+$. If the sequence $(N_\mathbf{i})$, $\mathbf{i} \in \mathbb{N}^f$, is an i.i.d. sequence with distribution P, the random rooted tree T is called a *Galton-Watson tree* with *offspring* distribution $\widehat{P}$. We will denote by $\mathrm{GWT}(P)$ the probability distribution of $[T]$ in $\mathbb{G}_*$.

Now, assume further that P has a positive finite first moment. We define $\widehat{P}$ as the distribution on $\mathbb{Z}_+$, defined for $k \geq 1$ by

$$\widehat{P}(k-1) = \frac{kP(k)}{\sum_\ell \ell P(\ell)}. \tag{3.4}$$

Then, the GWT with *degree* distribution P is the random rooted tree T where $(N_\mathbf{i})$, $\mathbf{i} \in \mathbb{N}^f \backslash \{\o\}$, is an i.i.d. sequence with distribution $\widehat{P}$, independent of $N_\o$ with distribution P. We will then denote by $\mathrm{GWT}_*(P)$ the probability distribution of $[T]$.

It is interesting to note that P is a Poisson random variable Poi_λ with $\lambda > 0$, then $\widehat{P} = P$. Thus, for the Poisson distribution, GWT's with degree and offspring distribution are identical. See also figure 3.1 for the case of regular trees.

We will prove that $\mathrm{GWT}_*(P)$ is the random weak limit of some finite random graph sequence $(G_n)_{n\geq 1}$ defined in the previous chapters. In particular, by lemma 3.8, it will prove that $\mathrm{GWT}_*(P)$ is unimodular. Let us however give a direct proof of this statement.

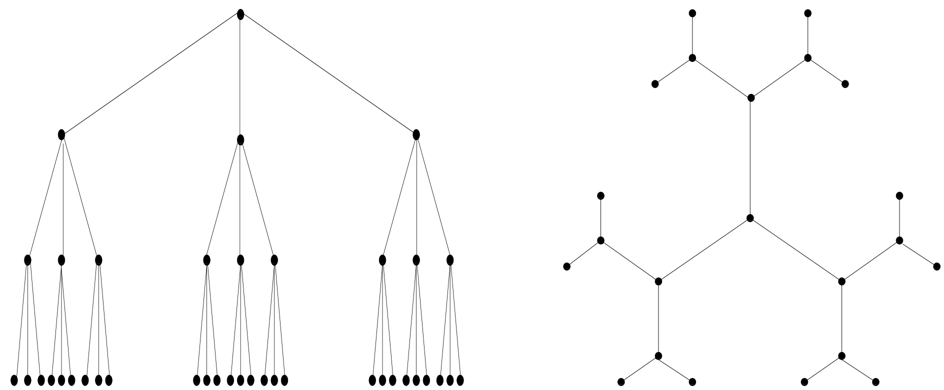

Figure 3.1: Left: representation of a 3-ary tree. Right: representation of a 3-regular tree.

Lemma 3.10 (Unimodular Galton-Watson trees) *If $P \in \mathcal{P}(\mathbb{Z}_+)$ has positive first moment, then* $\mathrm{GWT}_*(P)$ *is a unimodular measure in* $\mathcal{P}(\mathbb{G}_*)$.

Proof. We should prove that (3.1) holds. By lemma 3.9, we may restrict to functions $f : \mathbb{G}_{**} \to \mathbb{R}_+$ such that $f(G, u, v) = 0$ unless $\{u, v\} \in E_G$. If $(T_0, T_1, \cdots, T_k)$ are rooted trees, we denote by $R_u(T_1, \cdots, T_k)$ a tree T where $u \in V_T$ has k neighbors and the subtrees spanned by the neighbors of u are isomorphic to $T_1, \cdots, T_k$. Similarly $R_{u,v}(T_0, \cdots, T_k)$ is a tree T with $u, v \in V_T$ where u has $k+1$ neighbors, v with subtree isomorphic to T_0 and k others with subtrees isomorphic to $T_1, \cdots, T_k$ (see figure 3.2).

Now, let $T = (V, E)$ be a Galton-Watson tree with degree distribution P built from the sequence of random variables $(N_{\mathbf{i}})_{\mathbf{i} \in \mathbb{N}^f}$. We find

$$\mathbb{E} \sum_{i=1}^{N_\varnothing} f(T, \varnothing, i) = \sum_{k=1}^{\infty} P(k) \sum_{i=1}^{k} \mathbb{E}[f(T, \varnothing, i) | N_\varnothing = k] = \sum_{k=1}^{\infty} k P(k) \mathbb{E}[f(T, \varnothing, 1) | N_\varnothing = k].$$

Now, consider $(T_i), i \geq 1$, i.i.d. Galton-Watson trees with offspring distribution $\widehat{P}$. Then given $N_\varnothing = k$, $[T]$ and $[R_\varnothing(T_1, \cdots, T_k), \varnothing]$ have the same law. If N and $\widehat{N}$ are independent variables with law P and $\widehat{P}$, we get

$$\begin{aligned} \mathbb{E} \sum_{i=1}^{N_\varnothing} f(T, \varnothing, i) &= \mathbb{E} N \sum_{k=0}^{\infty} \widehat{P}(k) \mathbb{E} f(R_{\varnothing,1}(T_1, \cdots, T_{k+1}), \varnothing, 1) \\ &= \mathbb{E} N \mathbb{E} f(R_{\varnothing,1}(T_1, \cdots, T_{\widehat{N}+1}), \varnothing, 1). \end{aligned}$$

Now, up to a rooted isomorphism, $(R_\varnothing(T_2, \cdots, T_{\widehat{N}+1}), \varnothing)$ has same law than T_2. Define $S_{u,v}(T_1, T_2)$ as a tree where u and v are connected by an edge, and besides this edge u has a subtree isomorprhic to T_1 and v has a subtree isomorprhic to T_2 (see figure 3.2). Using the symmetry of $S_{u,v}$, we deduce that

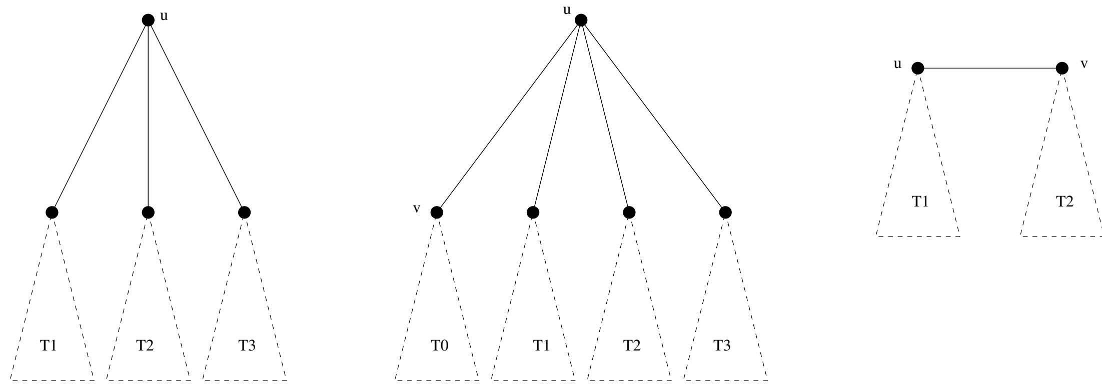

Figure 3.2: Left, $R_u(T_1, T_2, T_3)$. Center $R_{u,v}(T_0, \cdots, T_3)$. Right $S_{u,v}(T_1, T_2)$

$$\mathbb{E}\sum_{i=1}^{N_\varnothing} f(T, \varnothing, i) = \mathbb{E}N\mathbb{E}f(S_{\varnothing,1}(T_1, T_2), \varnothing, 1) = \mathbb{E}N\mathbb{E}f(S_{\varnothing,1}(T_1, T_2), 1, \varnothing).$$

Similarly, we perform the same computation for

$$\mathbb{E}\sum_{i=1}^{N_\varnothing} f(T, i, \varnothing) = \sum_{k=1}^{\infty} P(k) \sum_{i=1}^{k} \mathbb{E}[f(T, i, \varnothing)|N_\varnothing = k] = \sum_{k=1}^{\infty} kP(k)\mathbb{E}[f(T, 1, \varnothing)|N_\varnothing = k].$$

As above we find

$$\mathbb{E}\sum_{i=1}^{N_\varnothing} f(T, i, \varnothing) = \mathbb{E}N\mathbb{E}f(S_{\varnothing,1}(T_1, T_2), 1, \varnothing).$$

This proves that (3.1) holds. □

Exercise 3.11 *Let $P \in \mathcal{P}(\mathbb{Z}_+)$ with positive first moment. Prove that* GWT(P) *is unimodular if and only if P is a Poisson random variable. (Hint : use* (3.1) *with $f(G, u, v) = \mathbf{1}(\deg(G, u) = k)$).*

3.5 Convergence of random graphs

3.5.1 Erdős-Rényi graphs

Let G_n be an Erdős-Rényi graphs with distribution $\mathcal{G}(n, \lambda/n)$ with $\lambda > 0$ and $n \in \mathbb{N}$. We define the random probability measure on $\mathbb{G}_*$:

$$U(G_n) = \frac{1}{n}\sum_{i=1}^{n} \delta_{[G_n(i),i]}.$$

where δ is a Dirac mass. As already pointed, the measure $U(G_n)$ corresponds to the distribution of the random rooted graph $[G_n(ø), ø]$ where the root is drawn uniformly over the vertex set. Averaging over the randomness of the graph, we get for any event A in $\mathbb{G}_*$,

$$\mathbb{E}U(G_n)(A) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}([G_n(i), i] \in A) = \mathbb{P}([G_n(1), 1] \in A),$$

where we have used exchangeability. In other words, the measure $\mathbb{E}U(G_n)$ is simply the distribution of $[G_n(1), 1]$. The aim of this paragraph is to prove the following theorem.

Theorem 3.12 (Local convergence in Erdős-Rényi graph) *Let $\lambda > 0$ and, for integer $n \geq 1$, let G_n be an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$. Then, as n goes to infinity* $\mathbb{E}U(G_n) \rightsquigarrow \mathrm{GWT}(\mathrm{Poi}_\lambda)$.

This theorem should be compared with theorem 2.14 which asserts that there exists a Poisson number of cycles of finite length in $\mathcal{G}(n, \lambda/n)$. By exchangeability of the variable, it implies that the probability that i is in a cycle of fixed length k is of order $1/n$.

The proof of theorem 3.12 is based on an exploration of the connected component $G(v)$ of a graph $G = (V, E)$ that contains $v \in V$. This exploration is called the *breadth-first search*. Consider the total order in $\mathbb{N}^f$: for two elements $\mathbf{i} = (i_1, \cdots, i_n)$ and $\mathbf{j} = (j_1, \cdots, j_m)$ we set $\mathbf{i} < \mathbf{j}$ if $n < m$ or if $n = m$ and there exists k such that $(i_1, \cdots, i_k) = (j_1, \cdots, j_k)$ and $i_{k+1} < j_{k+1}$. We build an bijective map ϕ from $S \subset \mathbb{N}^f$ to the vertex set of $G(v)$. The set S will be of the type (3.3) and the map ϕ are defined iteratively and if $\mathbf{i} < \mathbf{j}$ are both in S then the value of $\phi(\mathbf{i})$ will be determined before the value of $\phi(\mathbf{j})$.

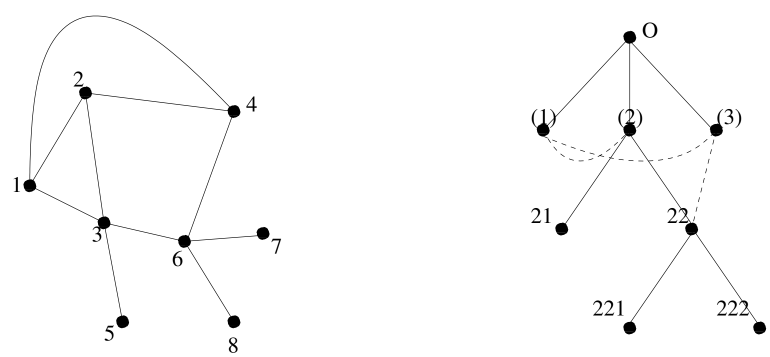

Figure 3.3: $\phi(ø) = 1$, $\phi(1) = 2$, $\phi(2) = 3$, $\phi(3) = 4$, $\phi((2,1)) = 5$, $\phi((2,2)) = 6$, $\phi((2,2,1)) = 7$, $\phi((2,2,2)) = 8$.

This exploration is iterative, at integer step t, a vertex may belong to the active set A_t, to the unexplored set U_t or to the connected set $C_t = V \backslash (A_t \cup U_t)$. We start with $A_0 = \{v\}$, $C_0 = \emptyset$, $U_0 = V \backslash \{v\}$ and $\phi(ø) = v$. For integer $t \geq 0$, if $A_t \neq \emptyset$, we define $v_{t+1} = \phi(\mathbf{i}_{t+1})$ as the vertex in A_t such that whose preimage by ϕ is minimal for the order on $\mathbb{N}^f$. Let $I_{t+1} = \{u \in$

$U_t : \{u, v_{t+1}\} \in E\}$ be the set of neighbors of v_{t+1} in U_t, we set

$$\left\{ \begin{array}{lcl} A_{t+1} & = & A_t \backslash \{v_{t+1}\} \cup I_{t+1} \\ U_{t+1} & = & U_t \backslash I_{t+1} \\ C_{t+1} & = & C_t \cup \{v_{t+1}\} \end{array} \right. \tag{3.5}$$

If $N_{\mathbf{i}_{t+1}} = |I_{t+1}|$ and $I_{t+1} = \{u_1, \cdots, u_{N_{\mathbf{i}_{t+1}}}\}$, we also set $\phi((\mathbf{i}_{t+1}, 1)) = u_1, \cdots, \phi((\mathbf{i}_{t+1}, N_{\mathbf{i}_{t+1}})) = u_{N_{\mathbf{i}_{t+1}}}$. If $A_t = \emptyset$, then the process stops. It follows by construction that

$$|G(v)| = \inf\{t \geq 1 : A_t = \emptyset\}.$$

For integer t, the image by ϕ of the vertices of generation t in S, $\phi(S \cap \mathbb{N}^t)$, are the set of vertices in G at distance t from v.

For ease of notation, we set $X_{t+1} = N_{\mathbf{i}_{t+1}} = |I_{t+1}|$ and $\tau = |G(v)|$. So that, for $t < \tau$,

$$|A_t| = 1 + \sum_{k=1}^{t} (X_k - 1) \ , \quad |U_t| = n - 1 - \sum_{k=1}^{t} X_k \ , \quad |C_t| = t. \tag{3.6}$$

Now, we consider the breadth-first search when $v = 1$ and the graph $G = G_n$ is an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$. We define the filtration

$$\mathcal{F}_t = \sigma((A_0, U_0, C_0), \cdots, (A_t, U_t, C_t)).$$

The hitting time $\tau = \inf\{t \geq 1 : A_t = \emptyset\}$ is a stopping time for this filtration. Notice that for any integer $t \geq 0$, given $\mathcal{F}_t$, if $\{t < \tau\} \in \mathcal{F}_t$ then X_{t+1} has distribution a binary random variable $\mathrm{Bin}(|U_t|, \lambda/n)$.

Lemma 3.13 (Convergence of exploration) *On an enlarged probability space, there exists a sequence $(X'_t)_{t \geq 1}$ of i.i.d.* Poi_λ *variables such that*

$$\mathbb{P}\left((X_1, \cdots, X_{t \wedge \tau}) \neq (X'_1, \cdots, X'_{t \wedge \tau})\right) \leq \frac{\lambda(\lambda+1)(t+1)^2}{n}.$$

Proof. The stopping property implies that $\{t < \tau\}$ is $\mathcal{F}_t$-measurable. We note also that from (3.6),

$$\mathbb{E}\left[(|A_t| - 1)\mathbf{1}_{t<\tau}\right] \leq \sum_{s=0}^{t-1} \mathbb{E}(X_{s+1} - 1)\mathbf{1}_{s<\tau} \leq \lambda t, \tag{3.7}$$

where we have used the fact that if $\{t < \tau\}$ holds then $\mathbb{E}(X_{t+1}|\mathcal{F}_t) = \lambda|U_t|/n \leq \lambda$.

Now, on an enlarged probability space, let ξ_{t+1} be, given $\mathcal{F}_t$, a binary variable $\mathrm{Bin}(n - |U_t|, \lambda/n)$ independent of X_t. Then $Y_{t+1} = X_{t+1} + \xi_{t+1}$ is a binary variable $\mathrm{Bin}(n, \lambda/n)$ and (Y_t)

is an i.i.d. sequence. Hence, from the union bound,

$$\begin{aligned}\mathbb{P}\left((X_1,\cdots,X_{t\wedge\tau})\neq(Y_1,\cdots,Y_{t\wedge\tau})\right) &\leq \mathbb{E}\sum_{s=1}^{t}\mathbf{1}_{s<\tau}\mathbb{P}\left(X_s\neq Y_s|\mathcal{F}_s\right)\\ &\leq \sum_{s=1}^{t}\mathbf{1}_{s<\tau}\mathbb{P}\left(\xi_s\neq 0|\mathcal{F}_s\right).\end{aligned}$$

If $s<\tau$, then $|U_t|\geq n-t-|A_t|$. It follows that

$$\mathbb{P}\left(\xi_s\neq 0|\mathcal{F}_s\right)=1-(1-\lambda/n)^{n-|U_s|}\leq\lambda(n-|U_s|)/n\leq\lambda\left(s+|A_s|\right)/n.$$

In particular, from (3.7), we get

$$\mathbb{P}\left((X_1,\cdots,X_{t\wedge\tau})\neq(Y_1,\cdots,Y_{t\wedge\tau})\right)\leq\sum_{s=1}^{t}\frac{\lambda(1+(\lambda+1)s)}{n}\leq\frac{\lambda t+\lambda(\lambda+1)t^2}{n}.$$

Then from (2.7), $d_{TV}(\mathcal{L}(Y_1,\cdots,Y_t),\mathrm{Poi}_\lambda^{\otimes t})\leq\lambda t/n$. We conclude by using the maximal coupling inequality. □

Lemma 3.14 (Asymptotically tree-like) *For integer $t\geq 0$, let $J_t=\{u\in A_t:\{u,v_{t+1}\}\in E\}$. We have*

$$\mathbb{P}\left(\exists 1\leq s\leq t\wedge\tau:|J_s|\neq 0\right)\leq\frac{\lambda^2t^2}{n}.$$

If $\{\forall 1\leq s\leq t:|J_s|=0\}$, the subgraph of G_n spanned by C_t is a tree.

Proof. Given $\mathcal{F}_t$, if $t<\tau$, $|J_t|$ is a binary variable $\mathrm{Bin}(|A_t|-1,\lambda/n)$. The union bound yields

$$\begin{aligned}\mathbb{P}\left(\exists 1\leq s\leq t\wedge\tau:|J_s|\neq 0\right) &\leq \mathbb{E}\sum_{s=1}^{t}\mathbf{1}_{s<\tau}\mathbb{P}\left(|J_s|\neq 0|\mathcal{F}_s\right)\\ &\leq \mathbb{E}\sum_{s=1}^{t}\mathbf{1}_{s<\tau}\left(1-\left(1-\frac{\lambda}{n}\right)^{|A_t|-1}\right)\\ &\leq \sum_{s=1}^{t}\frac{\lambda}{n}\mathbb{E}(|A_s|-1)\mathbf{1}_{s<\tau}.\end{aligned}$$

It remains to use (3.7) and the first statement follows.

To prove the second statement, we note that for all integer s, there cannot be an edge between an element of C_s and U_s. Therefore, if there is an edge between $u=\phi(\mathbf{i}_s)$ and $v=\phi(\mathbf{i}_{s'})$ with $s\leq s'$, then either $\mathbf{i}_s$ is the genitor of $\mathbf{i}_{s'}$, or $\mathbf{i}_s$ and $\mathbf{i}_{s'}$ are both active at time s. If $\{\forall 1\leq s\leq t:|J_s|=0\}$ holds the latter cannot happen. In particular, on this event, every vertex in $C_t \backslash \{1\}$ has a unique neighbor with a smaller index (its genitor). It follows that the graph spanned by C_t cannot have cycles. □

Proof of theorem 3.12. For ease of notation, we denote by $\overline{\rho}_n = \mathbb{E}U(G_n)$ the law of $[G_n(1), 1]$. With an abuse of notation, let us also write $(G_n, 1)$ instead of $(G_n(1), 1)$. Define $A = A_T = \{[G] \in \mathbb{G}_* : (G)_t \simeq T\}$ where T is a finite rooted tree of depth at most t. We first prove that $\overline{\rho}_n(A)$ converges to $\rho(A)$, where $\rho = \mathrm{GWT}(\mathrm{Poi}_\lambda)$. The number of vertices of T is equal to some integer m. Let $\mathcal{K}$ be the set of elements $[G]$ of $\mathbb{G}_*$ such the number of vertices in $(G)_t$ is less or equal than m. With the notation of lemma 3.14, if for all $1 \leq s \leq m \wedge \tau$, $|J_s| = 0$ and $[G_n, 1] \in \mathcal{K}$ then $(G_n, 1)_t$ is a tree. Moreover, from lemma 3.13, , if $[G_n, 1] \in \mathcal{K}$, there is a coupling such that the offsprings of the vertices of $(G_n, 1)_t$ are equal to independent Poisson variables on an event of probability at least $1 - \lambda(\lambda + 1)(m + 1)^2/n$. We deduce that

$$|\mathbb{P}((G_n, 1)_t \simeq T) - \rho(A)| = |\mathbb{P}((G_n, 1)_t \simeq T; G_n \in \mathcal{K}) - \rho(A)| \leq \frac{\lambda(\lambda + 1)(m + 1)^2 + \lambda^2 m^2}{n}.$$

Letting n tend to infinity, we obtain for any finite rooted tree T,

$$\lim_n \overline{\rho}_n(A_T) = \rho(A_T).$$

We are going to check that theorem 3.2(ii) holds. Let f be a bounded uniformly continuous function and $\varepsilon > 0$. By assumption there exists t such that $|f((G)_t) - f(G)| \leq \varepsilon$ for all $G \in \mathbb{G}_*$. Also there exists a finite collection of trees $\mathcal{S}$ such that

$$\sum_{T \in \mathcal{S}} \rho(A_T) > 1 - \varepsilon.$$

From what precedes, it follows that for n large enough, $\sum_{T \in \mathcal{S}} \overline{\rho}_n(A_T) > 1 - 2\varepsilon$ and

$$\left| \int f d\overline{\rho}_n - \int f d\rho \right| \leq \varepsilon(1 + 3\|f\|_\infty) + \sum_{T \in \mathcal{S}} f(T) \left| \overline{\rho}_n(A_T) - \rho(A_T) \right|.$$

Letting n tend to infinity and then ε goes to zero, we deduce the statement. □

3.5.2 Configuration model

Let $(\mathbf{d}_n) \in \mathbb{Z}_+^n$ be a vector of integers with even sum. We consider G_n a random multi-graph with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$. Again, we define the random probability measure on $\widehat{\mathbb{G}}_*$:

$$U(G_n) = \frac{1}{n} \sum_{i=1}^{n} \delta_{[G_n(i), i]}.$$

The measure $\mathbb{E}U(G_n)$ is the law of $[G_n(\o), \o]$ where $\o$ is an independent and uniform on $[n]$, law with respect to the randomness of G_n and $\o$.

Theorem 3.15 (Local convergence in configuration model) *Let $G_n \stackrel{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ with $\mathbf{d}_n$ satisfying (H_2), then as n goes to infinity, $\mathbb{E}U(G_n) \rightsquigarrow \mathrm{GWT}_*(P)$.*

As for Erdős-Rényi graphs, the proof is based on an exploration of the connected component $G(v)$ of a multigraph $G = (V, E)$ that contains $v \in V$. We will also build an bijective map ϕ from $S \subset \mathbb{N}^f$ to the vertex set of $G(v)$. The value of ϕ is defined iteratively and if $\mathbf{i} < \mathbf{j}$ are in S then the value of $\phi(\mathbf{i})$ will be determined before the value of $\phi(\mathbf{j})$. However, we change slightly the exploration procedure to be more adapted to the configuration model.

Let $\mathbf{d} = (d_v)_{v\in V}$ be a sequence of integers with $\sum_{v\in V} d_v$ even. We consider the set $\Delta = \{(v, j) : v \in V, 1 \leq j \leq d_v\}$ and we call $\Delta_v = \{(v, j) : 1 \leq j \leq d_v\}$ the set of half-edges with endpoint v. As in the configuration model, to a matching say σ on Δ, we associate the multigraph $G = G(\sigma) \in \widehat{\mathbb{G}}(\mathbf{d})$ where the half-edges are matched to form the edges of G.

The exploration is on the set of half-edges Δ and it is iterative. At integer step t, we partition Δ in 3 sets, an half-edge may belong to the active set A_t, to the unexplored set U_t or to the connected set $C_t = \Delta\backslash(A_t \cup U_t)$. At stage t, a vertex with an half-edge in $C_t \cup A_t$ will have a pre-image by ϕ in $\mathbb{N}^f$. We start with $v \in V$, $A_0 = \Delta_v$, $C_0 = \emptyset$ and $U_0 = \Delta\backslash\Delta_v$. Finally we set $\phi(\varnothing) = v$.

For integer $t \geq 0$, if $A_t \neq \emptyset$, we define $e_{t+1} = (\phi(\mathbf{i}_t), j_t)$ as the half-edge in A_t such that $\mathbf{i}_t$ is minimal and $(\phi(\mathbf{i}_t), k) \notin A_t$ for $k = 1, \cdots, j_t - 1$. Let $I_{t+1} = (\Delta_{v_{t+1}}\backslash\{\sigma(e_{t+1})\}) \cap U_t$ where v_{t+1} is the vertex such that $\sigma(e_{t+1}) \in \Delta_{v_{t+1}}$. I_{t+1} is the set of new half-edges and our partition of Δ is updated as

$$\begin{cases} A_{t+1} &= A_t\backslash\{e_{t+1}, \sigma(e_{t+1})\} \bigcup I_{t+1} \\ U_{t+1} &= U_t\backslash\left(I_{t+1} \cup \{\sigma(e_{t+1})\}\right) \\ C_{t+1} &= C_t \cup \{e_{t+1}, \sigma(e_{t+1})\}. \end{cases} \tag{3.8}$$

If $\sigma(e_{t+1}) \notin A_t$, we also set $\phi((\mathbf{i}_t, j_t)) = v_{t+1}$. Finally, if $A_t = \emptyset$, then the process stops.

We notice that the elements in C_t are the half-edges for which we know by step t their matched half-edge. It implies that $\sigma(e_{t+1}) \in A_t \cup U_t$. Moreover, for any vertex u, we cannot have simultaneously $\Delta_u \cap U_t \neq \emptyset$ and $\Delta_u \cap A_t \neq \emptyset$. With a slight abuse, we may thus write $u \in U_t$ or $u \in A_t$ if, respectively, $\Delta_u \cap U_t \neq \emptyset$ or $\Delta_u \cap A_t \neq \emptyset$. Now, if $v_{t+1} \in U_t$, then $I_{t+1} = \Delta_{v_{t+1}}\backslash\{\sigma(e_{t+1})\}$, otherwise $v_{t+1} \in A_t$ and $I_{t+1} = \emptyset$. Again, for integer k, the image by ϕ of the vertices of generation k in S, $\phi(S \cap \mathbb{N}^k)$, are the set of vertices in G at distance k from v.

For ease of notation, we set $X_0 = d_v$, $X_{t+1} = |I_{t+1}|$,

$$\varepsilon_{t+1} = \mathbf{1}_{v_{t+1}\in A_t} = \mathbf{1}_{\sigma(e_{t+1})\in A_t}$$

and

$$\tau = \inf\{t : A_t = \emptyset\}.$$

We get

$$|A_t| = d_v + \sum_{k=1}^{t}(X_k - 1 - \varepsilon_k) , \quad |U_t| = |\Delta| - d_v - \sum_{k=1}^{t}(X_k + 1 - \varepsilon_k) , \quad |C_t| = 2t.$$

Setting for $t > \tau$, $\varepsilon_t = 0$, we have by construction

$$|G(v)| = 1 + \tau - \sum_{t \geq 1} \varepsilon_t.$$

As in the statement of theorem 3.15, consider a random multi-graph G_n with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$. For $t \in \mathbb{N}$, we consider the filtration $\mathcal{F}_t = \sigma((A_0, U_0, C_0), \cdots, (A_t, U_t, C_t))$. The hitting time τ is a stopping time for this filtration. Also, the matching σ being uniformly distributed, given $\mathcal{F}_t$, if $\{t < \tau\}$, $\sigma(e_{t+1})$ is uniformly distributed on $U_t \cup A_t \backslash e_{t+1}$. It follows that for $u \in [n]$,

$$\begin{aligned}
\mathbb{P}(v_{t+1} = u | \mathcal{F}_t) &= \frac{|\Delta_u \cap (U_t \cup A_t \backslash \{e_{t+1}\})|}{|U_t| + |A_t| - 1} \\
&= \frac{\mathbf{1}_{u \in U_t} d_u}{|U_t| + |A_t| - 1} + \frac{\mathbf{1}_{u \in A_t}(|\Delta_u \cap A_t| - \mathbf{1}_{e_{t+1} \in \Delta_u})}{|U_t| + |A_t| - 1}.
\end{aligned}$$

If $\sigma(e_{t+1}) \in U_t$, then $X_{t+1} = d_{v_{t+1}} - 1$ otherwise, $\sigma(e_{t+1}) \in A_t$ and $X_{t+1} = 0$. We recall also that $|U_t| + |A_t| = |\Delta| - |C_t| = |\Delta| - 2t$. We get, for $k \geq 1$

$$\mathbb{P}(X_{t+1} = k | \mathcal{F}_t) = \begin{cases} \sum_{u \in U_t} \frac{\mathbf{1}_{d_u = k+1}(k+1)}{|\Delta| - 2t - 1} & k \geq 1, \\ \sum_{u \in U_t} \frac{\mathbf{1}_{d_u = 1}}{|\Delta| - 2t - 1} + \frac{|A_t| - 1}{|\Delta| - 2t - 1} & k = 0. \end{cases} \tag{3.9}$$

For integer t, the variable $|A_t|$ depends on the size n of graph and from the initial condition v. The next lemma implies that under $\mathbb{P}$, the sequence of random variables $|A_t|$ is tight in n when $v = \o$ is uniformly distributed on $[n]$.

Lemma 3.16 (Tightness of active set) *Under the assumption of theorem 3.15, consider the exploration process on the rooted graph $(G_n(\o), \o)$. There exists a constant $c > 0$ such that, for each integer $t \geq 0$, $\mathbb{E}|A_{t \wedge \tau}| \leq c(t+1)$.*

Proof. Let us use write $\mathbf{d}$ instead of $\mathbf{d}_n$. We order the sequence set $\mathbf{d} = (d_1, \cdots d_n)$ in non-decreasing order, we get a permutation π of $[n]$ such that $\mathrm{d}_{\pi(1)} \geq d_{\pi(2)} \cdots \geq d_{\pi(n)}$. Let n_0 be the number of non-null degrees. From assumption (H_0), $P(0) < 1$ and for all n large enough, $n_0 \geq 2$. We may then define the set

$$\hat{\Pi} = \left\{ \pi(i) : 1 \leq i \leq \frac{n_0}{2} \right\}.$$

This is the subset of vertices with the $n_0/2$ larger degrees. We denote by $\hat{\Delta} = \cup_{i \in \hat{\Pi}} \Delta_i$ and $Q_{\mathbf{d}}$ be the distribution on $\mathbb{N}$,

$$Q_{\mathbf{d}}(\{k\}) = \frac{k+1}{|\hat{\Delta}|} \sum_{i \in \hat{\Pi}} \mathbf{1}_{d_i = k+1}, \quad \text{for } k \geq 0.$$

We note that

$$\frac{|\Delta|}{2} \leq |\hat{\Delta}| \leq |\Delta| - \frac{n_0}{2}.$$

We first define a sequence $(Y_t)_{t\geq 1}$ of i.i.d. variables with distribution $Q_{\mathbf{d}}$, such that for all $1 \leq t \leq n_0/4 - 1$,

$$X_{t\wedge\tau} \leq Y_{t\wedge\tau}. \tag{3.10}$$

For $t \geq 0$, this is done explicitly by setting $Y_{t+1} = d_{u_{t+1}} - 1$ for some random $u_{t+1} \in \hat{\Pi}$ such that $\mathbb{P}(u_{t+1} = u|\mathcal{F}_t) = d_u/|\hat{\Delta}|$. We order decreasingly the half-edges from 1 to Δ :

$$(\pi(1), 1) \succ (\pi(1), 2) \cdots \succ (\pi(1), d_{\pi(1)}) \succ (\pi(2), 1) \cdots \succ (\pi(n), d_{\pi(n)}).$$

In particular, $\hat{\Delta}$ is the set of $|\hat{\Delta}|$ largest half-edge of Δ. We recall that $|U_t \cup A_t| = |\Delta| - 2t$ and that $\sigma(e_{t+1})$ is uniformly distributed on $U_t \cup A_t \backslash e_{t+1}$. Now, let $1 \leq t \leq \tau \wedge (n_0/4 - 1)$, if $\sigma(e_{t+1})$ is the k-th largest half-edge of $U_t \cup (A_t \backslash e_{t+1})$ and $k \leq |\hat{\Delta}|$ then we define u_{t+1} as the vertex such that the k-th largest half-edge of Δ is in $\Delta_{u_{t+1}}$. Otherwise, $d_{v_{t+1}}$ is less or equal to any degrees in $\hat{\Pi}$ and we define u_{t+1} as the vertex such that the N-th largest half-edge of $\hat{\Delta}$ is in $\Delta_{u_{t+1}}$, where N is an independent variable uniformly distributed in $\hat{\Delta}$. Since $1 \leq t \leq n_0/4 - 1$, we have $|U_t \cup A_t \backslash e_{t+1}| = |\Delta| - 2t - 1 \geq |\Delta| - n_0/2 \geq |\hat{\Delta}|$. It follows that $\mathbb{P}(Y_{t+1} \in \cdot|\mathcal{F}_t) = Q_{\mathbf{d}}$ and $X_t \leq Y_t$. We deduce that (3.10) holds for $1 \leq t \leq n_0/4 - 1$.

It yields that for $1 \leq t \leq n_0/4 - 1$,

$$\sum_{i=1}^{t\wedge\tau} X_i \leq \sum_{i=1}^{t\wedge\tau} Y_i. \tag{3.11}$$

Now, the inequality, $|\Delta|/2 \leq |\hat{\Delta}|$ gives

$$\mathbb{E}[Y] \leq \sum_{i\in\hat{\Pi}} \frac{d_i(d_i-1)}{|\hat{\Delta}|} \leq 2\sum_{i=1}^{n} \frac{d_i(d_i-1)}{|\Delta|}$$

Let D be a variable with law P. By lemma 1.4 we deduce,

$$\limsup_{n\to\infty} \mathbb{E}[Y] \quad \leq \quad 2\frac{\mathbb{E}D(D-1)}{\mathbb{E}D},$$

and

$$\lim_{n\to\infty} \frac{n_0}{n} = \mathbb{P}(D \geq 1).$$

In particular, for n large enough, $t \leq n_0/4 - 1$. Similarly, we have $X_0 = d_{\o}$ and by lemma 1.4, we find

$$\lim_{n\to\infty} \mathbb{E}[X_0] = \mathbb{E}D.$$

Finally, using (3.11), we take the expectation of $|A_t| = X_0 + \sum_{k=1}^{t}(X_k - 1 - \varepsilon_k)$ and the claim follows. $\square$

We extend the sequence $(X_0, \cdots, X_\tau)$ for $t \geq \tau + 1$, by setting for all $s \geq 1$, $X_{\tau+s} = Y_s$ for some iid sequence $(Y_t)_{t\geq 1}$ with distribution $\widehat{P}$.

Lemma 3.17 (Convergence of exploration) *Under the assumption of theorem 3.15, consider the exploration process on the rooted graph $(G_n(\emptyset), \emptyset)$. The variable $(X_0, X_1, \cdots, X_t)$ converges in distribution to $P \otimes \widehat{P}^{\otimes t}$.*

Proof. Since $X_0 = d_\emptyset$, X_0 converges in distribution to P. Note also that $|A_t| + 2t$ half-edges are not in U_t. It follows by (3.9) that, if $\{t < \tau\}$ holds, for any $k \geq 0$,

$$\left| \mathbb{P}(X_{t+1} = k | \mathcal{F}_t) - \frac{k+1}{|\Delta| - 2t - 1} \sum_{i=1}^{n} \mathbf{1}_{d_i = k+1} \right| \leq \frac{k+1}{|\Delta| - 2t - 1} \left(2t + |A_t|\right).$$

By lemma 1.4 implies that $|\Delta|/n$ converges to $\mathbb{E}D$, where D has law P. Hence, for any $a > 0$, we get on the event $\{|A_t| \leq a\}$,

$$\lim_{n \to \infty} \mathbb{P}(X_{t+1} = k | \mathcal{F}_t) = \widehat{P}(\{k\}).$$

However, by lemma 3.16, for each $t \geq 1$, $\mathbb{P}(|A_{t \wedge \tau}| \geq a) \leq c(t+1)/a$. Hence the probability that there exists $1 \leq s \leq t \wedge \tau$ such that $|A_s| > a$ is bounded above by $ct(t+1)/a$.

Letting n tend to infinity and then a to infinity, it implies that $(X_0, X_1, \cdots, X_t)$ converges weakly to $P \otimes \widehat{P}^{\otimes t}$. □

We introduce a variable that counts the number of times that two elements in the active sets are matched by step t :

$$E_t = \sum_{k=1}^{t} \varepsilon_k.$$

Lemma 3.18 (Asymptotically tree-like) *Under the assumption of theorem 3.15, consider the exploration process on the rooted graph $(G_n(\emptyset), \emptyset)$. For every integer $t \geq 0$, we have*

$$\lim_n \mathbb{P}\left(E_{t \wedge \tau} \neq 0\right) = 0.$$

If $t \leq \tau$ and $E_t = 0$, the subgraph of G_n spanned by the vertices with all their half-edges in C_t is a tree.

Proof. We start with the second statement. To every vertex u with an half-edge in $C_t \cup A_t$, there is an element $\mathbf{i}$ in $\mathbb{N}^f$ such that $\phi(\mathbf{i}) = u$. We may thus order these vertices by the order through ϕ^{-1} in $\mathbb{N}^f$. Every such vertex is adjacent to its genitor. By construction if $E_t = 0$ or equivalently if for all $1 \leq s \leq t$, $\varepsilon_s = 0$, then every vertex with an half-edge in $C_t \cup A_t$ has a unique adjacent vertex with a smaller index (and it is its ancestor). It follows easily that there cannot be a cycle in the subgraph spanned by these vertices.

If $E_{t\wedge\tau} \neq 0$, there exists an integer $s \leq t \wedge \tau$ such that $\sigma(e_s) \in A_{s-1}$. It follows from the union bond and the fact that $\{s < \tau\} \in \mathcal{F}_s$,

$$\begin{aligned}
\mathbb{P}(\exists 1 \leq s \leq t \wedge \tau : \sigma(e_s) \in A_{s-1}) &\leq \mathbb{E}\left[\sum_{s\geq 0} \mathbf{1}_{s<t\wedge\tau} \mathbb{P}(v_{s+1} \in A_s | \mathcal{F}_s)\right] \\
&\leq \mathbb{E}\sum_{s=0}^{t-1} \frac{(|A_s|-1)_+}{|\Delta| - 2s - 1}.
\end{aligned}$$

From lemma 3.16, for each $t \geq 0$, $\mathbb{E}|A_t| \leq c(t+1)$. Also, by lemma 1.4, $|\Delta|/n$ converges to $\mathbb{E}D$, where D has law P. The conclusion of the first statement follows. □

Proof of theorem 3.15. The proof follows the argument of the proof of theorem 3.12. For ease of notation, we write $(G_n, \emptyset)$ in place of $(G_n(\emptyset), \emptyset)$. We denote by $\overline{\rho}_n$, the law of $[G_n, \emptyset]$ and $\rho = \mathrm{GWT}_*(P)$. Define $A = \{[G] \in \mathbb{G}_* : (G)_t \simeq T\}$ where T is a finite rooted tree of depth at most t. From theorem 3.2, it is sufficient to prove that for any integer $t \geq 1$ and any such rooted tree T, $\overline{\rho}_n(A)$ converges to $\rho(A)$.

The number of vertices of T is equal to some integer m. Let $\mathcal{K}$ be the set of elements of $\mathbb{G}_*$ such the number of vertices in $(G)_t$ is less or equal than m. From lemma 3.18, if $E_{m\wedge\tau} = 1$ and $[G_n, \emptyset] \in \mathcal{K}$ then $(G_n, \emptyset)_t$ is a tree. Moreover, by lemma 3.17, if $[G_n, \emptyset] \in \mathcal{K}$, the number of offsprings of vertices different from $\emptyset$ in $(G_n, \emptyset)_t$ converges in distribution to independent variables with distribution $\widehat{P}$. The number of offsprings of root vertex $\emptyset$ converges to an independent variable with distribution P. We deduce that

$$\lim_n |\mathbb{P}((G_n, \emptyset)_t \simeq T) - \rho(A)| = \lim_n |\mathbb{P}((G_n, \emptyset) \simeq T; [G_n, \emptyset] \in \mathcal{K}) - \rho(A)| = 0.$$

The conclusion follows. □

Exercise 3.19 *Let G_n be a Chung-Lu graph with distribution $\mathcal{G}(n, \underline{\lambda}_n)$ with $\underline{\lambda}_n$ satisfying (H_2'). By extending the proof of theorem 3.12, show that $\mathbb{E}U(G_n)$ converges weakly to* $\mathrm{GWT}_*(Q)$ *where $Q(k) = \int \mathrm{Poi}_\lambda(k) P(d\lambda)$.*

3.6 Concentration and convergence of random graphs

3.6.1 Bounded difference inequality

Let $\mathcal{X}_1 \cdots \mathcal{X}_n$ be metric spaces and let F be a measurable function on $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ and P a product measure on $\mathcal{X}$. There is very powerful tool to bound the deviation of F from its mean when F is Lipschitz for a weighted Hamming distance, i.e. for every x and y in $\mathcal{X}$,

$$\sum_{k=1}^n a_k \mathbf{1}_{x_k \neq y_k} \leq F(x) - F(y) \leq \sum_{k=1}^n b_k \mathbf{1}_{x_k \neq y_k}. \tag{3.12}$$

for some $a = (a_1, \cdots, a_n) \in \mathbb{R}_-^n$, $b = (b_1, \cdots, b_n) \in \mathbb{R}_+^n$. We denote by $\|y\|_2 = \sqrt{\sum_i y_i^2}$, the usual Euclidean norm.

Theorem 3.20 (Azuma-Hoeffding's inequality) *Let F be as above, then*

$$P\left(F - \int F dP \geq t\right) \leq \exp\left(\frac{-2t^2}{\|b-a\|_2^2}\right).$$

This type of result is called a concentration inequality. It has found numerous applications in mathematics over the last decades. For more on concentration inequalities, we refer to Ledoux (2001). As a corollary, we deduce the Hoeffding's inequality.

Corollary 3.21 (Hoeffding's inequality) *Let $(X_k)_{1\leq k\leq n}$ be an independent sequence of real random variables such that for all integer k, $X_k \in [a_k, b_k]$. Then,*

$$\mathbb{P}\left(\sum_{k=1}^n X_k - \mathbb{E}X_k \geq t\right) \leq \exp\left(\frac{-t^2}{2\sum_{k=1}^n (b_k - a_k)^2}\right). \tag{3.13}$$

The proof of theorem 3.20 will be based on a lemma due to Hoeffding.

Lemma 3.22 *Let X be real random variable in $[a, b]$ such that $\mathbb{E}X = 0$. Then, for all $\lambda \geq 0$,*

$$\mathbb{E}e^{\lambda X} \leq e^{\frac{\lambda^2(b-a)^2}{8}}.$$

Proof. By the convexity of the exponential,

$$e^{\lambda X} \leq \frac{b-X}{b-a}e^{\lambda a} + \frac{X-a}{b-a}e^{\lambda b}.$$

Taking expectation, we obtain, with $p = -a/(b-a)$,

$$\begin{aligned} \mathbb{E}e^{\lambda X} &\leq \frac{b}{b-a}e^{\lambda a} - \frac{a}{b-a}e^{\lambda b} \\ &= \left(1 - p + pe^{\lambda(b-a)}\right) e^{-p\lambda(b-a)} \\ &= e^{\varphi(\lambda(b-a))}, \end{aligned}$$

where $\varphi(x) = -px + \ln(1 - p + pe^x)$. The derivatives of φ are

$$\varphi'(x) = -p + \frac{pe^x}{(1-p)e^{-x} + p} \quad \text{and} \quad \varphi''(x) = \frac{p(1-p)}{((1-p)e^{-x} + p)^2} \leq \frac{1}{4}.$$

Since $\varphi(0) = \varphi'(0) = 0$, we deduce from Taylor expansion that

$$\varphi(x) \leq \varphi(0) + x\varphi'(0) + \frac{x^2}{2}\|\varphi''\|_\infty \leq \frac{x^2}{8}.$$

□

Proof of theorem 3.20. Let $(X_1, \cdots, X_n)$ be a random variable on $\mathcal{X}$ with distribution P. We shall prove that

$$\mathbb{P}(F(X_1, \cdots, X_n) - \mathbb{E}F(X_1, \cdots, X_n) \geq t) \leq \exp\left(\frac{-t^2}{2\|b-a\|_2^2}\right).$$

For integer $1 \leq k \leq n$, let $\mathcal{F}_k = \sigma(X_1, \cdots, X_k)$, $Z_0 = \mathbb{E}F(X_1, \cdots, X_n)$, $Z_k = \mathbb{E}[F(X_1, \cdots, X_n)|\mathcal{F}_k]$, $Z_n = F(X_1, \cdots, X_n)$. We also define $Y_k = Z_k - Z_{k-1}$, so that $\mathbb{E}[Y_k|\mathcal{F}_{k-1}] = 0$. Finally, let $(X'_1, \cdots, X'_n)$ be an independent copy of $(X_1, \cdots, X_n)$. If $\mathbb{E}'$ denote the expectation over $(X'_1, \cdots, X'_n)$, we have

$$Z_k = \mathbb{E}'F(X_1, \cdots, X_k, X'_{k+1}, \cdots, X'_n).$$

It follows by (3.12)

$$Y_k = \mathbb{E}'F(X_1, \cdots, X_k, X'_{k+1}, \cdots, X'_n) - \mathbb{E}'F(X_1, \cdots, X_{k-1}, X'_k, \cdots, X'_n) \in [a_k, b_k].$$

Since $\mathbb{E}[Y_k|\mathcal{F}_{k-1}] = 0$, we may apply Lemma 3.22: for every $\lambda \geq 0$,

$$\mathbb{E}[e^{\lambda Y_k}|\mathcal{F}_{k-1}] \leq e^{\frac{\lambda^2(b_k-a_k)^2}{8}}.$$

This estimates does not depend on $\mathcal{F}_{k-1}$, it follows that

$$\mathbb{E}e^{\lambda(Z_n - Z_0)} = \mathbb{E}[e^{\lambda\sum_{k=1}^n Y_k}] \leq e^{\frac{\lambda^2\|b-a\|_2^2}{8}}.$$

From Chernov bound, for every $\lambda \geq 0$,

$$\mathbb{P}(F(X_1, \cdots, X_n) - \mathbb{E}F(X_1, \cdots, X_n) \geq t) \leq \exp\left(-\lambda t + \frac{\lambda^2\|b-a\|_2^2}{8}\right).$$

Optimizing over the choice of λ, we choose $\lambda = 4t/\|b-a\|_2^2$. □

3.6.2 Almost sure convergence of Erdős-Rényi random graphs

Let G_n be an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$ with $\lambda > 0$ and $n \in \mathbb{N}$. As above, we consider the random probability measure on $\mathbb{G}_*$:

$$U(G_n) = \frac{1}{n}\sum_{i=1}^n \delta_{[G_n(i),i]},$$

where δ is the Dirac mass. The measure $U(G_n)$ corresponds to the distribution of the random rooted graph $[G_n(\varnothing), \varnothing]$ where the root is drawn uniformly over the vertex set.

Theorem 3.23 (Almost sure local weak convergence of Erdős-Rényi graphs) *Let $\lambda > 0$ and for integer $n \geq 1$, $G_n \overset{d}{\sim} \mathcal{G}(n, \lambda/n)$ built on a common probability space. As n goes to infinity, a.s. $U(G_n) \rightsquigarrow \mathrm{GWT}(\mathrm{Poi}_\lambda)$.*

Proof. Define $\rho_n = U(G_n)$, $A = \{[G] \in \mathbb{G}_* : (G)_t \simeq H\}$ where H is a finite rooted graph of diameter at most t. From theorem 3.12, it is sufficient to check that $|\rho_n(A) - \mathbb{E}\rho_n(A)|$ converges a.s. to 0. For $1 \leq k \leq n$, let $Z_k = \{1 \leq i \leq k : \{i,k\} \in E_n\}$, where E_n is the edge set of G_n. The vector $(Z_1, \cdots, Z_n)$ is an independent vector and for some function F (depending on n) :

$$\rho_n(A) = \sum_{i=1}^n \delta_{[G_n,i]}(A) = \sum_{i=1}^n \mathbf{1}_{(G_n,i)_t \simeq H} = F(Z_1, \cdots, Z_n).$$

If $\mathcal{X}_k$ is the set of subsets of $[k]$, F is a function from $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ to $\mathbb{N}$. We cannot apply directly theorem 3.20 since the function F is Lipschitz with bad constants in (3.12). We shall reduce our set $\mathcal{X}$ to obtain better Lipshitz constants. This makes the proof a little cumbersome.

Let $M = \max_{1 \leq i \leq n} |Z_i|$. For each $1 \leq i \leq n$, the variable $|Z_i|$ is a Binary random variable $\mathrm{Bin}(i-1, \lambda/n)$. For $\theta \geq 0$, hence

$$\mathbb{E}e^{\theta |Z_i|} = \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n}e^{\theta}\right)^{i-1} \leq e^{\lambda(e^\theta - 1)}.$$

From Chernov bound, we get

$$\begin{aligned} \mathbb{P}(M \geq \log n) &\leq n\mathbb{P}(|Z_1| \geq \log n) \\ &\leq ne^{-\theta \log n} e^{\lambda(e^\theta - 1)}. \end{aligned} \tag{3.14}$$

We define $\mathbb{E}_n$, $\mathbb{P}_n$, as the conditional expectation and probabilities given $\{M < \log n\}$. Since $0 \leq \rho_n(A) \leq 1$, we find easily

$$|\mathbb{E}_n F - \mathbb{E}F| \leq 2\frac{\mathbb{P}(M \geq \log n)}{1 - \mathbb{P}(M \geq \log n)}.$$

Choosing any $\theta > 1$ in (3.14) yields to

$$\lim_{n\to\infty} |\mathbb{E}_n F - \mathbb{E}F| = 0. \tag{3.15}$$

Let $c = \sum_{s=0}^{t-1} d^s$, where d is the maximal degree of H and take n sufficiently large such that $\log n \geq c$. We define $\tilde{\mathcal{X}}_k$ as the set of subsets of $[k]$ of cardinal less than $\log n$, and $\tilde{\mathcal{X}} = \tilde{\mathcal{X}}_1 \times \cdots \times \tilde{\mathcal{X}}_n$. As a function on $\tilde{\mathcal{X}}$, F satisfies (3.12) with $-a_k = b_k = 2c \log n$. Indeed, assume that $x_k = y_k$ for all but one coordinate, say i. Let G be the graph with edge set $x_1 \cup \cdots \cup x_{i-1} \cup x_{i+1} \cup \cdots \cup x_n$. To affect the value of $F(x) - F(y)$ an edge must be of the type $\{i,j\}$ where $1 \leq j \leq i$ satisfies for some $v \in [n]$, $j \in B_G(v,t)$ and $(G,v)_t$ is isomorphic to a subgraph of H. Also, since the maximal degree in H is d, for this vertex j there is at most $c(\log n)$ vertices $v \in [n]$ with $j \in B_G(v,t)$ and $(G,v)_t$ isomorphic to a subgraph of H. Since the $|x_i| \leq \log n$, we deduce $|F(x) - F(y)| \leq 2c(\log n)^2$.

Given $\{M \leq \log n\}$, the vector $(Z_1, \cdots, Z_n)$ is still independent, we deduce from theorem 3.20 that

$$\mathbb{P}_n\left(|F - \mathbb{E}_n F| \geq s\right) \leq 2\exp\left(\frac{-ns^2}{8c^2(\log n)^4}\right).$$

So finally, we use the inequality

$$\mathbb{P}\left(|F - \mathbb{E}_n F| \geq s\right) \quad \leq \quad \mathbb{P}_n\left(|F - \mathbb{E}_n F| \geq s\right) + \mathbb{P}\left(M \leq \log n\right).$$

The conclusion follows from (3.14) with $\theta = 3$, equation (3.15) and Borel-Cantelli lemma. □

A near consequence of theorem 3.23 and proposition 2.1 is the following.

Corollary 3.24 (Almost sure convergence of subtree counts) *Under the assumptions of 3.23, let T be a tree with m edges and c elements in its automorphism group. Then, as n goes to infinity $X(T;G_n)/n$ converges a.s. to $c^{-1}\lambda^m$.*

Proof (sketch). Let H be a finite graph. From proposition 2.1, it is sufficient to check that $|X(H;G_n)/n - \mathbb{E}X(H;G_n)/n|$ converges a.s. to 0. Define the continuous function $f(G,\varnothing) = \sum_{F\subset G} \mathbf{1}_{\varnothing\in V_F}\mathbf{1}_{F\simeq H}$. We have

$$nU(G_n)(f) = \sum_{i=1}^{n}\sum_{F\simeq H} \mathbf{1}_{i\in V_F}\mathbf{1}_{F\subset G_n} = |V_H|X(H;G_n).$$

Note that we cannot apply directly theorem 3.23 since f is not bounded. To overcome this difficulty, it is in fact simpler to prove directly that a.s. $X(H;G_n)$ converges. We skip the details, but it is possible to compute the 4-th moment of $X(H;G_n)$. It gives that

$$\mathbb{E}(X(H;G_n) - \mathbb{E}X(H;G_n))^4 \leq \frac{c'}{n^2}.$$

In particular, $X(H;G_n) - \mathbb{E}X(H;G_n)$ converges a.s. to 0. □

Remark 3.25 (Concentration for graph functionals) *In the proof of theorem 3.23, we have checked the following inequality. Assume that L is a map from $\mathbb{G}(n)$ to $\mathbb{R}$ such that for some $\delta, c > 0$ and any $G = ([n], E) \in \mathbb{G}(n)$ with degree bounded by δ and $e \in E$, we have*

$$|L(G) - L(G-e)| \leq c,$$

where $G - e = ([n], E\backslash\{e\})$. Then, if $G \overset{d}{\sim} \mathcal{G}(n,p)$, for any $\theta > 0$ and $t > 0$, we have

$$\mathbb{P}\left(|L(G) - \mu| \geq t\right) \leq ne^{-\theta\delta}e^{np(e^{\theta}-1)} + 2\exp\left(\frac{-t^2}{8c^2\delta^2}\right),$$

where $\mu = \mathbb{E}(L(G)|M \leq \delta)$ and M was defined in the proof of theorem 3.23. This concentration inequality is certainly not optimal but it will be useful in a few applications.

3.6.3 Concentration inequality on uniform matchings

We start with an alternative statement of Azuma-Hoeffding's inequality.

Theorem 3.26 (Azuma-Hoeffding's inequality, second form) *Let $Z_0, \cdots, Z_n$ be a real martingale with respect to a filtration $\mathcal{F}_0, \cdots, \mathcal{F}_n$. Assume that for any integer $1 \le k \le n$, almost surely $Z_k - Z_{k-1} \in [a_k, b_k]$, then*

$$\mathbb{P}(Z_n - Z_0 \ge t) \le \exp\left(\frac{-2t^2}{\|b-a\|_2^2}\right).$$

Proof. Setting $Y_k = Z_{k+1} - Z_k$, the proof is contained in the proof of theorem 3.20. □

From this form of Azuma-Hoeffding's inequality, we are able to derive a concentration inequality on matchings. Let Δ be a finite set with even cardinal. We say that two matchings σ, σ' on Δ differ from *at most a switch* if there exists a subset J, with $|J| \le 4$, such that $\sigma(k) = \sigma'(k)$ for all $k \in \Delta \backslash J$. Note that if $|\Delta|$ is even and σ, σ' differ from at most a switch then either $\sigma = \sigma'$ (corresponding to $J = \emptyset$) or there exist $i \ne j$ such that $\sigma(i) \ne j$, $\sigma'(j) = i$ and $\sigma'(\sigma(j)) = \sigma(i)$ (corresponding to $|J| = 4$, see figure 2.1).

The next corollary is stated in (Wormald, 1999, theorem 2.19).

Corollary 3.27 (Concentration on uniform matchings) *Let Δ be a finite set with even cardinal and F be a real function on matchings of Δ such that*

$$|F(m') - F(m)| \le c,$$

if m, m' differ from at most a switch. Then, if σ is a uniformly drawn matching of Δ,

$$\mathbb{P}(F(\sigma) - \mathbb{E}F(\sigma) \ge t) \le \exp\left(\frac{-t^2}{|\Delta|c^2}\right).$$

Proof. Without loss of generality, we assume that $\Delta = \{1, \cdots, n\}$, with $n = |\Delta|$. We may identify a matching of Δ as the set of $n/2$ matched pairs. We order these $n/2$ pairs by the index of their smallest element. We then define $\mathcal{F}_0$ as the trivial σ-algebra and for $1 \le k \le n/2$, we define $\mathcal{F}_k$ as the σ-algebra generated by the first k pairs of matched elements of σ. We set $Z_k = \mathbb{E}[F(\sigma)|\mathcal{F}_k]$, so that $Z_0 = \mathbb{E}F(\sigma)$, $Z_{n/2-1} = F(\sigma)$. By construction, Z_k is a Doob martingale.

Let $\mathbb{M}(\Delta)$ be the set of matchings of Δ. For $1 \le k \le n/2$, an element σ of $\mathbb{M}(\Delta)$ can be uniquely decomposed into $(\sigma^-_{k-1}, \sigma^+_k)$ where $\sigma^-_{k-1} \in \mathbb{M}(\Delta_{k-1})$ is the restriction of σ to the $k-1$ smallest pairs and $\sigma^+_k \in \mathbb{M}(\Delta \backslash \Delta_{k-1})$ is the restriction of σ to $\Delta \backslash \Delta_{k-1}$.

If v_k is the smallest element of $\Delta \backslash \Delta_{k-1}$, we set $w_k = \sigma(v_k) \in \Delta \backslash \Delta_{k-1}$, so that $\Delta_k = \Delta_{k-1} \cup \{v_k, w_k\}$. Now, for $w \in \Delta \backslash (\Delta_{k-1} \cup \{v_k\})$, let $\mathbb{M}_w$ denote the set of matchings of $\Delta \backslash \Delta_{k-1}$

such that $m(v_k) = w$. Then for any $w, w' \in \Delta\backslash(\Delta_{k-1} \cup \{v_k\})$, each $m \in \mathbb{M}_w$ corresponds to a unique $m' \in \mathbb{M}_{w'}$ through the switch $\{\{v_k, w\}, \{w', z\}\} \to \{\{v_k, w'\}, \{w, z\}\}$, where $m(w') = z$. This gives a bijection between $\mathbb{M}_w$ and $\mathbb{M}_{w'}$, and we set $N_k = |\mathbb{M}_w|$. By assumption, we deduce that for any w, w',

$$\left| \sum_{m \in \mathbb{M}_w} F(\sigma_k^-, m) - \sum_{m \in \mathbb{M}_{w'}} F(\sigma_k^-, m) \right| \leq c.$$

Applying the above inequality to w_k, we deduce that

$$\left| \frac{1}{N_k} \sum_{m \in \mathbb{M}_{w_k}} F(\sigma_k^-, m) - \frac{1}{n-2k+1} \sum_{w \in \Delta\backslash(\Delta_{k-1} \cup \{v_k\})} \frac{1}{N_k} \sum_{m \in \mathbb{M}_w} F(\sigma_k^-, m) \right| = |Z_k - Z_{k-1}| \leq c.$$

We may then apply theorem 3.26. □

3.6.4 Almost sure convergence in the configuration model

For integer n, let $\mathbf{d}_n$ be an array of variables satisfying assumption (H_2). Consider a sequence $(G_n)_{n \in \mathbb{N}}$ of random multigraphs with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$. As usual, we define the random probability measure on $\widehat{\mathbb{G}}_*$:

$$U(G_n) = \frac{1}{n} \sum_{i=1}^{n} \delta_{[G_n(i), i]}.$$

Theorem 3.28 (Almost sure LWC in configuration model) *Let $(\mathbf{d}_n)_{n \geq 1}$ be an array satisfying (H_p) for some $p > 2$. Consider a sequence random multigraph $G_n \stackrel{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ built on a common probability space. Then as n goes to infinity, almost surely $U(G_n) \rightsquigarrow \mathrm{GWT}_*(P)$.*

Proof. Define $\rho_n = U(G_n)$ and $A = \{[G] \in \mathbb{G}_* : (G)_t \simeq H\}$ where H is a finite rooted graph of depth at most t. By theorem 3.15, it is sufficient to check that $\rho_n(A) - \mathbb{E}\rho_n(A)$ converges a.s. to 0. We write

$$n\rho_n(A) = \sum_{i=1}^{n} \mathbf{1}((G_n(i), i)_t \simeq T) = F(\sigma),$$

where F is a function on matchings of $\Delta = \{(i, j) : 1 \leq i \leq n, 1 \leq j \leq d_i\}$ and σ is uniformly drawn matching on Δ.

Let $M = \max_{i \in [n]} d_i(n)$ and d be the maximal degree of H and $c = \sum_{s=0}^{t-1} d^s$. If two matchings m, m' of Δ differ by at most a switch then $|F(m) - F(m')| \leq 4cM$. Indeed, a switch changes the status 4 edges and, arguing as in the proof of theorem 3.23, the addition or the removal of an edge can modify for at most cM vertices the value of $\mathbf{1}((G_n(i), i)_t \simeq H)$. From corollary 3.27, we get

$$\mathbb{P}\left(|F(\sigma) - \mathbb{E}F(\sigma)| > nt\right) \leq 2 \exp\left(\frac{-n^2 t^2}{16 |\Delta| c^2 M^2} \right). \tag{3.16}$$

By lemma 1.5, $M = o(n^{1/p})$. From Borel Cantelli lemma, we deduce that $F(\sigma) - \mathbb{E}F(\sigma)$ converges a.s. to 0. $\square$

Corollary 3.29 (Almost sure LWC in graphs given degree sequences) *Let $(\mathbf{d}_n)_{n\geq 1}$ be an array satisfying (H_p) for some $p > 2$. Consider a sequence random multigraph $G_n \overset{d}{\sim} \mathcal{G}(\mathbf{d}_n)$ built on a common probability space. Then as n goes to infinity, almost surely $U(G_n) \rightsquigarrow \mathrm{GWT}_*(P)$.*

Proof. Let $\hat{G}_n \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ build from the random matching σ. With the notation of the proof of theorem 3.28,

$$\mathbb{P}(|\rho_n(A) - \mathbb{E}\rho_n(A)| \geq t) \leq \frac{\mathbb{P}\left(|F(\sigma) - \mathbb{E}F(\sigma)| > nt\right)}{\mathbb{P}(\hat{G}_n \text{ is a graph})}.$$

It remains to apply (2.14), lemma 1.6 and (3.16). $\square$

A consequence of theorem 3.28 and proposition 2.4 is the following. It can be proved along the line of corollary 3.24.

Corollary 3.30 (Almost sure convergence of subtree counts) *Let $1 \leq k \leq n$, T be a tree with k vertices and maximal degree bounded by $p \geq 2$. Assume that T has c elements in its automorphism groups. Let $(\mathbf{d}_n)_{n\geq 1}$ be an array satisfying (H_{4p}) and consider a sequence random multigraph $G_n \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ built on a common probability space. Then a.s.*

$$\lim_n \frac{X(T; G_n)}{n} = c^{-1}(\mathbb{E}D)^{-k-1} \prod_{i=1}^{k} \mathbb{E}\left[(D)_{\deg(i;T)}\right],$$

where D has distribution P.

Remark 3.31 (Concentration for graph functionals) *The proof of theorem 3.28 contains the following concentration inequality. Let $\mathbf{d} = (d_1, \cdots, d_n)$ be integer vector with $S = \sum_{i=1}^n d_i$ even. Assume that L is a map from $\widehat{\mathbb{G}}(\mathbf{d})$ to $\mathbb{R}$ such that for some $c > 0$ and any $G, G' \in \widehat{\mathbb{G}}(\mathbf{d})$ which differ by a single switch of edges, we have*

$$|L(G) - L(G')| \leq c,$$

Then, if $G \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d})$, for any $t > 0$,

$$\mathbb{P}\left(|L(G) - \mathbb{E}L(G)| \geq t\right) \leq 2\exp\left(\frac{-t^2}{c^2 S}\right).$$

If moreover $\mathbf{d}$ is graphic, then the same bound holds for $G \overset{d}{\sim} \mathcal{G}(\mathbf{d})$ by replacing $\mathbb{E}L(G)$ by $\mathbb{E}L(\hat{G})$ and the constant 2 in front of the exponential by $2/\mathbb{P}(\hat{G}$ is a graph$)$ where $\hat{G} \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d})$.

Chapter 4

The giant connected component

In this chapter, we will study the size of the connected components of our random graphs. In the first two sections, we shall start with some classical results on Galton-Watson trees and random walks.

4.1 Growth of Galton-Watson trees

A GWT can be an infinite or a finite tree. Consider a GWT with offspring distribution P, and let $Z_n = |V \cap \mathbb{N}^n|$ be the total number of n-th generation vertices, we have

$$Z_0 = 1 \quad \text{and} \quad Z_{n+1} = \sum_{\mathbf{i} \in V \cap \mathbb{N}^n} N_{\mathbf{i}},$$

with the usual convention that the sum over an empty set is 0. We denote by $(X_{n,1}, \cdots, X_{n,Z_n})$ the number of offsprings of n-th generation vertices, we get

$$Z_0 = 1 \quad \text{and} \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}. \tag{4.1}$$

The collection $(X_{n,i})$ is an i.i.d. array of random variables with distribution P. The process $(Z_n), n \in \mathbb{N}$, is called a *Galton-Watson branching process*. It represents the evolution with generations of the size of a population. There are Z_n individual of generation n and all individuals give birth independently of a random number of children with common distribution P. It is clear that the state 0 is an absorbing state of the process $(Z_n), n \in \mathbb{N}$. The *probability of extinction* ρ is defined as

$$\rho = \mathbb{P}(\exists\, n \geq 1 : Z_n = 0) = \mathbb{P}\left(\sum_{n \geq 0} Z_n < \infty \right).$$

The probability of extinction is the probability that the GWT is finite. We define the generating function, for $z \in [0,1]$,

$$\varphi(z) = \mathbb{E}[z^X] = \sum_{k \geq 0} P(\{k\}) z^k,$$

where X has distribution P.

Theorem 4.1 (Extinction probability for GWT**)** *For a GWT with offspring distribution P,*

(i) If $\mathbb{E}X < 1$, then $\rho = 1$.

(ii) If $\mathbb{E}X > 1$, then ρ is the unique fixed point in $(0,1)$ of $x = \varphi(x)$.

(iii) If $\mathbb{E}X = 1$ and $\mathbb{P}(X = 1) < 1$ then $\rho = 1$.

For a GWT with degree distribution P, we still denote by ρ the probability of extinction, i.e. the probability that the tree is finite. Let $\widehat{X}$ be a random variable with distribution $\widehat{P}$ and $\widehat{\varphi}(z) = \mathbb{E}[z^{\widehat{X}}] = \sum_{k \geq 0} \widehat{P}(\{k\}) z^k$ be the generating function of $\widehat{P}$. With the above notation for P, we find

$$\widehat{\varphi}(z) = \frac{\varphi'(z)}{\varphi'(1)} \quad \text{and} \quad \mathbb{E}\widehat{X} = \frac{\mathbb{E}[X(X-1)]}{\mathbb{E}[X]}.$$

Corollary 4.2 (Extinction probability for GWT_***)** *For a GWT with degree distribution P and $0 < \sum_\ell \ell P(\ell) < \infty$,*

(i) If $\mathbb{E}[X(X-2)] < 0$, then $\rho = 1$.

(ii) If $\mathbb{E}[X(X-2)] > 0$, then $\rho = \varphi(\widehat{\rho})$ where $\widehat{\rho}$ is the unique fixed point in $(0,1)$ of $x = \widehat{\varphi}(x)$.

(iii) If $\mathbb{E}[X(X-2)] = 0$ and $\mathbb{P}(X = 2) < 1$ then $\rho = 1$.

Corollary 4.3 (Extinction probability for Poisson-GWT) *If the offspring distribution is* Poi_λ *for some $\lambda > 0$. Then if $\lambda \leq 1$, $\rho = 1$, while if $\lambda > 1$, ρ is the unique solution in $(0,1)$ of the equation*

$$x = e^{\lambda(x-1)}. \tag{4.2}$$

Proof of theorem 4.1. We define the moment generating function of Z_n, $\varphi_n(x) = \mathbb{E}[x^{Z_n}]$. From (4.1), it follows that

$$\varphi_0(x) = x \quad \text{and} \quad \varphi_{n+1}(x) = \sum_k \mathbb{P}(Z_n = k)\mathbb{E}\left[\prod_{i=1}^k x^{X_{n,i}}\right] = \varphi_n(\varphi(x)).$$

We deduce that $\varphi_n = \varphi \circ \cdots \circ \varphi$ is the n-th composition of φ. The event $\{Z_n = 0\}$ is non-decreasing in n. It follows that

$$\rho = \lim_n \mathbb{P}(Z_n = 0) = \lim_n \varphi_n(0).$$

Now $\rho_n = \varphi_n(0)$ satisfies $\rho_0 = 0$, $\rho_{n+1} = \varphi(\rho_n)$ and $\lim_n \rho_n = \rho$. We deduce that ρ is the smallest solution in $[0,1]$ of the equation $x = \varphi(x)$.

Since φ is convex, the derivative of $f(x) = \varphi(x) - x$, $f'(x) = \varphi'(x) - 1$ is non-decreasing, $f'(1) = \mathbb{E}X - 1$. If $\mathbb{E}X < 1$, f is decreasing and the unique fixed point of φ is $\rho = 1$. If $\mathbb{E}X > 1$, f there is a second fixed point in $(0,1)$. This proves $(i) - (ii)$.

For (iii), we notice that if $\mathbb{E}X = 1$, then Z_n is a non-negative mean one martingale with respect to the filtration $\mathcal{F}_n = \sigma(Z_0, Z_1, \cdots, Z_n)$. Let $\mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$, from Doob's martingale convergence theorem, there exists a $\mathcal{F}_\infty$-measurable random variable Z, such that a.s. $\lim_n Z_n = Z$ and $Z_n = \mathbb{E}[Z|\mathcal{F}_n]$. Let $A = \{Z = 0\}$, since $Z_n = 0$ implies $Z = 0$, we have $\rho = \mathbb{P}(A)$. Similarly, $\mathbb{E}[\mathbf{1}_A|\mathcal{F}_n]$ is a bounded martingale and from Doob's martingale convergence theorem, a.s. $\lim_n \mathbb{E}[\mathbf{1}_A|\mathcal{F}_n] = \mathbf{1}_A$ (Levy's 0-1 law).

Now, we notice that $\mathbb{P}(A|\mathcal{F}_n) \geq \mathbb{P}(X_{n,1} = \cdots = X_{n,Z_n} = 0) = P(\{0\})^{Z_n} > 0$. From what precedes Z_n converges a.s. to Z and we deduce that a.s.

$$\mathbf{1}_A = \lim_n \mathbb{E}[\mathbf{1}_A|\mathcal{F}_n] \geq P(\{0\})^Z > 0.$$

It follows that a.s. $\mathbf{1}_A = 1$. □

Proof of corollary 4.2. Let T be a $\mathrm{GWT}_*(P)$, for $1 \leq i \leq N_\varnothing$, let T_i be the rooted subtree of T on the vertex set $V_i = V \cap \{\mathbf{i} \in \mathbb{N}^f : i_1 = i\}$. Then $T_1 \cdots, T_{N_\varnothing}$ are i.i.d. $\mathrm{GWT}(\widehat{P})$, independent of $N_\varnothing$. The event $\{T \text{ is finite}\}$ is equal to the event that all subtrees are finite, hence,

$$\rho = \sum_{k \geq 0} \mathbb{P}(N_\varnothing = k)\rho^k = \varphi(\rho).$$

To conclude, we apply theorem 4.1. □

Corollary 4.4 (Growth of GWT) *With the above notation, let $\mu = \mathbb{E}X$ and $\widehat{\mu} = \mathbb{E}\widehat{X} = \mathbb{E}[X(X-1)]/\mathbb{E}[X]$.*

(i) For a GWT with offspring or degree distribution P, there exists a random variable W such that a.s.

$$\lim_n \frac{Z_n}{\mu^n} = W.$$

(ii) For a GWT with degree distribution P, there exists a random variable W such that a.s.

$$\lim_n \frac{Z_n}{\mu \widehat{\mu}^{n-1}} = W.$$

Moreover, conditioned on non-extinction, W is positive. Finally, if $\int x^p dP < \infty$ for some $p > 1$ in case (i) or $p > 2$ in case (ii) then $\mathbb{E}W = 1$.

Proof. We note that for (i) and (ii), Z_n/μ^n and $Z_n/(\mu\widehat{\mu}^{n-1})$ are non-negative martingale with mean 1 with respect to their natural filtration. The statement follows then from the martingale convergence theorem. □

We conclude this section with the continuity of the extinction probability as a function of the offspring distribution. For a probability measure $P \in \mathcal{P}(\mathbb{Z}_+)$, we define $\rho(P) \in [0,1]$ as the smallest solution of $\varphi(x) = x$ where φ is the generating function of P.

Lemma 4.5 (Continuity of extinction probability) *The map $P \mapsto \rho(P)$ from $\mathcal{P}(\mathbb{Z}_+)$ to $[0,1]$ is continuous for the weak convergence at any $P \neq \delta_1$.*

Proof. Take $P \neq \delta_1$. Fix a sequence of probability measures P_n with $P_n \rightsquigarrow P$. Setting $\rho_n = \rho(P_n)$, $\rho = \rho(P)$ we should prove that $\rho_n \to \rho$. We denote by φ_n and φ the generating functions of P_n and P. For any $\varepsilon > 0$, we have the uniform convergence

$$\max_{x \in [0,1-\varepsilon]} |\varphi_n(x) - \varphi(x)| \to 0. \tag{4.3}$$

We first prove that $\liminf_n \rho_n \geq \rho$. Consider a subsequence of $\rho_{n'}$ converging to $\rho' \in [0,1]$. If $\rho' < 1$ then for some $\varepsilon > 0$ and all n' large enough $\rho_{n'} \in [0, 1-\varepsilon]$. Hence using (4.3), we find that

$$0 = \varphi_{n'}(\rho_{n'}) - \rho_{n'} = \varphi(\rho_{n'}) - \rho_{n'} + o(1) = \varphi(\rho') - \rho' + o(1).$$

In particular $\varphi(\rho') = \rho'$ and $\rho' = \rho < 1$ since there is at most one solution in $[0,1)$ of $\varphi(x) = x$. Indeed, since $P \neq \delta_1$, φ is strictly convex.

To conclude of the proof of the lemma, it remains to check that $\limsup_n \rho_n \leq \rho$. We may assume that $\rho < 1$ otherwise there is nothing to prove. Fix any $x \in (\rho, 1)$, the function φ being strictly convex $\varphi(x) - x < 0$. From (4.3), we deduce that for all n large enough, $\varphi_n(x) - x < 0$. In particular $\rho_n < x$. Since x may be arbitrarily close to ρ, we get $\limsup_n \rho_n \leq \rho$. □

4.2 Random walks and branching processes

We consider a Galton-Watson Branching process $(Z_n)_{n \geq 0}$ with offspring distribution P:

$$Z_0 = 1 \quad \text{and} \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i},$$

where $(X_{n,i}), (n,i) \in \mathbb{N}^2$, is an i.i.d. array of random variables with distribution P. When the process reaches 0, we pay attention to the total population size

$$\tau = \sum_{n \geq 0} Z_n.$$

We will interpret τ has the time that a random walk hits 0. Informally, imagine that we reveal one by one, for each individual, its number of offsprings. For integer $t \geq 0$, we define A_t as the set of active individuals, i.e. the set of individuals whose parent has been revealed but whose offsprings are still unknown. At time 0, there is one ancestor individual in A_0. For integer $t \geq 0$, if $A_t \neq \emptyset$, we pick an individual in A_t. We remove this individual from A_t, add its offsprings and we get A_{t+1}. The process stops when A_t is empty for the first time.

More formally, an individual is defined as a couple $v = (n, i), n \geq 1, 1 \leq i \leq Z_n$, where n is its generation, and i its index within its generation. The individual v has $X_v = X_{n,i}$ offsprings. Now since Z_{n+1} is the sum of the number of offsprings of generation n individuals, we may define the set of offsprings of $(n, 1)$ as $I_{(n,1)} = \{(n+1, 1), \cdots, (n+1, X_{n,1})\}$, of $(n, 2)$ as $I_{(n,2)} = \{(n+1, X_{n,1}+1), \cdots, (n+1, X_{n,1}+X_{n,2})\}$ and up to $I_{(n,Z_n)} = \{(n+1, \sum_{k=1}^{n-1} X_{n,k} + 1), \cdots, (n+1, Z_{n+1})\}$. We set $A_0 = \{(0, 1)\}$. For integer $t \geq 0$, if $A_t \neq \emptyset$, we define v_{t+1} as the oldest individual in A_t (i.e. the smallest individual in lexicographic order) and set

$$A_{t+1} = A_t \backslash \{v_{t+1}\} \cup I_{v_{t+1}}.$$

Notice that $|I_{v_{t+1}}|$ is independent from A_t. In particular, if $S_t = |A_t|$ and $X_{t+1} = X_{v_{t+1}}$, we have $S_0 = 1$ and

$$S_{t+1} = S_t - 1 + X_{t+1},$$

and (X_t) is an i.i.d. sequence with distribution P. (S_t) is nothing else that a random walk with i.i.d. increment $(X_t - 1)$. Moreover

$$\tau = \inf\{t \geq 1 : S_t = 0\}.$$

Therefore, hitting time properties on random walks translate into properties on the the total population size in Galton-Watson branching processes.

4.3 Hitting time for random walks

Let P be a probability measure on $\mathbb{R}$ and let X, $(X_n), n \in \mathbb{N}$, be a sequence of i.i.d random variables with distribution P. For integer $t \geq 1$, let $S_t = x + \sum_{i=1}^{t} X_i$ be a simple random walk starting at $S_0 = x > 0$. (S_t) is a Markov chain and we denote by $\mathbb{P}^x$ is distribution given $S_0 = x$. We define

$$\tau = \inf\{t \geq 1 : S_t \leq 0\}.$$

We assume that $\mathbb{E}|X| < \infty$. It follows easily from the law of large numbers that if $\mathbb{E}X < 0$ then τ is a.s. finite while if $\mathbb{E}X > 0$, the event $\{\tau = \infty\}$ has positive probability under $\mathbb{P}^x$, $x \geq 0$. Recall that if the characteristic function $\varphi(\theta) = \mathbb{E}e^{\theta X}$ is differentiable in a neighborhood of 0 then

$$\varphi'(0) = \mathbb{E}X.$$

In particular if $\mathbb{E}X < 0$, there exists $\theta > 0$ such that $\varphi(\theta) < 1$. Similarly, if $\mathbb{E}X > 0$, there exists $\theta < 0$ such that $\varphi(\theta) < 1$.

Theorem 4.6 (Hitting time estimates) *Let X be a real random variable and $(S_t)_{t\geq 0}$ be as above.*

(i) If $\mathbb{E}X < 0$, let $\theta > 0$ in the domain of φ such that $\varphi(\theta) \leq 1$. Then $\mathbb{P}^x(\tau \geq t) \leq e^{\theta x}\varphi(\theta)^t$.

(ii) If $\mathbb{E}X > 0$, let $\theta < 0$ in the domain of φ such that $\varphi(\theta) \leq 1$. Then $\mathbb{P}^x(\tau < \infty) \leq e^{\theta x}$.

Proof. Assume first that $\mathbb{E}X < 0$. $M_t = e^{\theta S_t}/\varphi(\theta)^t$ is non-negative martingale with mean $M_0 = e^{\theta x}$ with respect to the filtration $\mathcal{F}_t = \sigma(S_0, \cdots, S_t)$. From Doob's optional stopping time theorem, we have

$$\mathbb{E}^x[M_\tau] \leq \mathbb{E}^x[\varphi(\theta)^{-\tau}] = e^{\theta x}.$$

Then, since $0 < \varphi(\theta) \leq 1$, from Markov inequality,

$$\mathbb{P}(\tau \geq t) = \mathbb{P}(\varphi(\theta)^{-\tau} \geq \varphi(\theta)^{-t}) \leq e^{\theta x}\varphi(\theta)^t.$$

Assume now that $\mathbb{E}X > 0$. Let (M_t) be as above and $t \geq 1$ be a fix integer. From Doob's optional stopping time theorem, we have

$$\mathbb{E}^x[M_{\tau\wedge t}] = e^{\theta x}.$$

Now, we notice that $M_{\tau\wedge t} \geq \mathbf{1}_{\tau\leq t}M_\tau$. In particular, since $M_\tau \geq \varphi(\theta)^{-\tau} \geq 1$, we get

$$\mathbb{P}(\tau \leq t) \leq e^{\theta x}.$$

The above inequality holding for all $t \geq 1$, we deduce statement (ii). □

Corollary 4.7 (Hitting time for Binary variables) *Let $\lambda > 0$, n be an integer, and $\alpha = \lambda - 1 - \log\lambda > 0$. We assume that $X = Y - 1$ where Y is a binary random variable* $\mathrm{Bin}(n, \lambda/n)$. *Then*

(i) If $\lambda < 1$, then $\mathbb{P}^x(\tau \geq t) \leq \lambda^{-x}e^{-\alpha t}$.

(ii) If $\lambda > 1$, then $\mathbb{P}^x(\tau < \infty) \leq \lambda^{-x}$.

Proof. From the inequality for all real z, $(1+z) \leq e^z$, we get for real θ,

$$\mathbb{E}e^{\theta Y_1} = \left(1 - \frac{\lambda}{n} + \frac{\lambda}{n}e^\theta\right)^n \leq e^{\lambda(e^\theta - 1)}.$$

The left hand side if the characteristic function of a Poisson random variable. We get

$$\varphi(\theta) \leq e^{\lambda e^\theta - \lambda - \theta}.$$

We then minimize the exponential over θ, it gives $\theta = -\log\lambda$ and $\varphi(\theta) \leq e^{-\alpha}$. We may now apply theorem 4.8. □

The next lemma is a follows direcetly from Chernov bound.

Lemma 4.8 (Chernov bound) *Let X be a real random variable and $(S_t)_{t\geq 0}$ be as above and $\varphi(\theta)=\mathbb{E}e^{\theta X}$. Then for any $x>0$ and integer $t\geq 0$,*

(i) If $\theta>0$ is in the domain of φ then $\mathbb{P}^0(S_t-\mathbb{E}S_t\geq x)\leq e^{-\theta x}\varphi(\theta)^t e^{-t\theta\mathbb{E}X}$.

(ii) If $\theta<0$ is in the domain of φ then $\mathbb{P}^0(S_t-\mathbb{E}S_t\leq -x)\leq e^{\theta x}\varphi(\theta)^t e^{-t\theta\mathbb{E}X}$.

Corollary 4.9 (Chernov bound for Binary variables) *Let $\lambda>0$, n be an integer, and $\gamma(x)=(x+1)\log(1+x)-x\geq 0$. We assume that X is a binary random variable* $\mathrm{Bin}(n,\lambda/n)$, *and let $(S_t)_{t\geq 0}$ be as above. Let $x>0$, then*

(i) $\mathbb{P}^0(S_t-t\lambda\geq \lambda t x)\leq e^{-\lambda t\gamma(x)}$.

(ii) $\mathbb{P}^0(S_t-t\lambda\leq -\lambda t x)\leq e^{-\lambda t\gamma(x)}$.

Theorem 4.10 (Hitting time for heavy-tailed variables) *Let X be real random variable and $(S_t)_{t\geq 0}$ be as above. If $\mathbb{E}X<0$ and $\mathbb{E}|X|^\alpha<\infty$ for some $\alpha\geq 1$, then for any $x\geq 0$, $\mathbb{E}^x[\tau^\alpha]<\infty$.*

Lemma 4.11 *If Y is a real random variable such that $\mathbb{E}Y<0$ and $\mathbb{E}|Y|^\alpha<\infty$ for some $\alpha\geq 1$. There exists a constant $x_0>0$ such that for all $x\geq x_0$, $\mathbb{E}|x+Y|^\alpha\leq x^\alpha$.*

Proof. It is sufficient to prove that,

$$\lim_{x\to\infty}\mathbb{E}\left[x\left(|1+x^{-1}Y|^\alpha-1\right)\right]=\alpha\mathbb{E}Y<0.$$

Let $n=[\alpha]\geq 1$ be the integer part of α, and $r=\alpha-n\in[0,1)$. For all $y\leq 0$, $(1+y)^u\leq 1+y^u\wedge y$, and

$$(1+y)^\alpha=(1+y)^n(1+y)^u\leq 1+y+\sum_{k=1}^n\binom{n}{k}(y^k+y^{k+u}).$$

For all $x>0$, $|1+x^{-1}Y|^\alpha\leq(1+|x^{-1}Y|)^\alpha$, we get

$$x\left(\left|1+\frac{Y}{x}\right|^\alpha-1\right)\leq |Y|+\sum_{k=1}^n\binom{n}{k}\left(\frac{|Y|^k}{x^{k-1}}+\frac{|Y|^{k+u}}{x^{k+u-1}}\right).$$

The conclusion follows by dominated convergence. □

Proof of theorem 4.10. We set $S_0=x$ and $\mu=\mathbb{E}X$. There exists $L>0$, such that $\mathbb{E}X\mathbf{1}_{X\geq -L}<0$. The hitting time of the negative half plane of the random walk with increments $(X_t\mathbf{1}_{X_t\geq -L})_t$ is larger than the hitting time of the original random walk with increments $(X_t)_t$. It is thus sufficient to prove the theorem for a random variable X with support in $[-L,\infty)$ for some $L>0$. Then since $S_{\tau-1}>0$, we note that $S_\tau\leq -L$. We introduce the random variables

$$Y_t=1-\frac{2X_t}{|\mu|}\quad\text{and}\quad M_t=\sum_{s=1}^t Y_s=t-\frac{2(S_t-x)}{|\mu|}.$$

we write

$$\begin{aligned}
\tau^\alpha &= \left(\sum_{t=1}^{\tau}(1-\frac{2}{|\mu|}X_t)+\frac{2}{|\mu|}\sum_{t=1}^{\tau}X_t\right)^\alpha \\
&\leq 2^{\alpha-1}\left(|M_\tau|^\alpha+\left(\frac{2|S_\tau-x|}{|\mu|}\right)^\alpha\right) \\
&\leq 2^{\alpha-1}|M_\tau|^\alpha+2^{\alpha-1}\left(\frac{2(L+x)}{|\mu|}\right)^\alpha .
\end{aligned}$$

It is thus sufficient to prove that $\mathbb{E}|M_\tau|^\alpha<\infty$. For t integer, let $Z_t=|M_{t\wedge\tau}|^\alpha$, then $(Z_t)_t$ converges a.s. to $|M_\tau|^\alpha$ and

$$\mathbb{E}Z_t\leq\mathbb{E}\left(\sum_{s=1}^{t}|Y_s|\right)^\alpha\leq t^{\alpha-1}\sum_{s=1}^{t}\mathbb{E}\,|Y_s|^\alpha\leq t^\alpha\mathbb{E}\,|Y_1|^\alpha<\infty.$$

Now, since $\{\tau\geq t+1\}$ is $\mathcal{F}_t$-measurable,

$$\begin{aligned}
\mathbb{E}[Z_{t+1}-Z_t] &= \mathbb{E}[(Z_{t+1}-Z_t)\mathbf{1}_{\tau\geq t+1}] \\
&= \mathbb{E}\left[\mathbb{E}\left[|M_{t+1}|^\alpha-|M_t|^\alpha \;\middle|\; \mathcal{F}_t\right]\mathbf{1}_{\tau\geq t+1}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[|M_t+Y_{t+1}|^\alpha-|M_t|^\alpha \;\middle|\; \mathcal{F}_t\right]\mathbf{1}_{\tau\geq t+1}\right]
\end{aligned}$$

By construction, for all $1\leq t<\tau$, $S_t>0$, and in particular, $M_t=t-2(S_t-x)/|\mu|>t$. We may then apply lemma 4.11, we get that for all $t\geq x_0$, $\mathbb{E}[Z_{t+1}-Z_t]\leq 0$. We have proved that

$$\sup_{t\geq 1}\mathbb{E}Z_t\leq\sup_{1\leq t\leq x_0}\mathbb{E}Z_t<\infty.$$

We conclude by Fatou's lemma: $\mathbb{E}|M_\tau|^\alpha\leq\liminf_t\mathbb{E}Z_t<\infty$. □

Remark 4.12 *Let $(P_n)_n$ be a sequence of probability measures on $\mathbb{R}$. We assume that under $\mathbb{P}_n$, $(X_t)_{t\geq 1}$ is an i.i.d. sequence with distribution P_n. We consider the random walk $S_t=x+\sum_{s=1}^{t}X_s$ started at $x>0$. We assume that for some $\mu<0$, for all n, $\mathbb{E}_nX=\int xdP_n\leq\mu$, and that the random variable $|X|^\alpha$ is uniformly integrable over $(P_n)_n$. Then the proof of theorem 4.10 actually shows that there exists a constant $C>0$, such that for all n, $\mathbb{E}_n\tau^\alpha<C$.*

4.4 Emergence of the giant component

We now take interest to existence of a giant connected component in a random graph. To be more precise, let $G=(V,E)$ be a locally finite graph. For $v\in V$, we define $G(v)$ as the connected component of the graph G that contains the vertex v. If V is finite, we may take interest to the size of the largest component: $\max_{v\in V}|G(v)|$. If G is an Erdős-Rényi random graph, there is a celebrated phase transition for the size of the largest component.

Theorem 4.13 (Giant component in Erdős-Rényi graph) *Let $\lambda > 0$, $\alpha = \lambda - 1 - \log \lambda > 0$, and let G_n be a sequence of Erdős-Rényi graphs with distribution $\mathcal{G}(n, \lambda/n)$ built on a common probability space.*

(i) If $0 < \lambda < 1$, then for any $c > 1/\alpha$,

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{v\in[n]} |G_n(v)| \geq c\log n \right) = 0.$$

(ii) If $\lambda > 1$, then a.s.

$$\lim_{n\to\infty} \frac{\max_{v\in[n]} |G_n(v)|}{n} = 1 - \rho,$$

where ρ is given by (4.2). Moreover there exists $c > 0$ such that a.s. for all n large enough the second largest connected component is larger that $c \log n$.

This theorem is consistent with theorem 3.12. Indeed, $(G_n(1), 1)$ converges in distribution to $\mathrm{GWT}(\mathrm{Poi}_\lambda)$. Since the event $\{|G_n(1)| \leq t\}$ is measurable with respect to $(G(1), 1)_t$, we deduce that

$$\lim_n \mathbb{P}(|G_n(1)| \leq t) = \mathbb{P}(\tau \leq t),$$

where τ is the total population of a Galton-Watson branching process with offspring distribution Poi_λ. We deduce that

$$\lim_{t\to\infty} \lim_n \mathbb{P}(|G_n(1)| \leq t) = \rho.$$

In the proof of theorem 4.13, we shall see that if $0 < \lambda < 1$, a.s.

$$\limsup_{n\to\infty} \frac{\max_{v\in[n]} |G_n(v)|}{\log n} \leq \frac{2}{\alpha}.$$

Similarly, if G is a graph with given degree sequence, there is a phase transition for the size of the largest component. The probability of extinction of a Galton-Watson with degree distribution P is a scalar ρ given by corollary 4.2(ii):

$$\rho = \varphi(\widehat{\rho}) \quad \text{with } \widehat{\rho} \text{ smallest solution of} \quad \widehat{\varphi}(z) = z. \tag{4.4}$$

Theorem 4.14 (Giant component in configuration model) *Let $(\mathbf{d}_n)_{n\geq 1}$ be an array satisfying (H_2). Consider a sequence random multigraph $G_n \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ built on a common probability space. Let D be a random variables with distribution P.*

(i) If $\mathbb{E}D(D-2) < 0$ and $(H_{1+\alpha})$ holds for some $\alpha > 1$ then for any $c > 1/\alpha$,

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{v\in[n]} |G_n(v)| \geq n^c \right) = 0.$$

(ii) If $\mathbb{E}D(D-2) > 0$*, then a.s.*

$$\lim_{n\to\infty} \frac{\max_{v\in[n]} |G_n(v)|}{n} = 1 - \rho,$$

where ρ *is given by (4.4). Moreover there exists* $c > 0$ *such that a.s. for all* n *large enough the second largest connected component is larger that* $c \log n$*.*

The statement of theorem 4.14(i) could not be much improved. Indeed, notice that the maximum degree of a graph is a lower bound on the size of the largest connected component. However, if $\beta > \alpha$ and $\mathbb{P}(D \geq t) \sim t^{-1-\beta}$ then $\mathbb{E}D^{1+\alpha} < \infty$ and the maximum degree in G_n will typically be of order $n^{1/(1+\beta)}$.

Using corollary 2.20, we will find that

Corollary 4.15 (Giant component in configuration model) *Let* $(\mathbf{d}_n)_{n\geq 1}$ *be an array satisfying* (H_2)*. Consider a sequence random multigraph* $G_n \stackrel{d}{\sim} \mathcal{G}(\mathbf{d}_n)$ *built on a common probability space. Then the conclusion of theorem 4.14 also holds for* G_n*.*

In the next two sections, we give a proof of theorems 4.13, 4.14. It will be based on the correspondence between random walk and branching processes. For example, for the proof of theorem 4.13, we will explore the connected component $G(v)$ as in (3.5). With the notation of section 4.2, we define $X_t = |I_t|$ and $S_t = |A_t|$. So that

$$S_t = 1 + \sum_{k=1}^{t} (X_k - 1) \ , \quad |U_t| = n - 1 - \sum_{k=1}^{t} X_k$$

and

$$|G(v)| = \tau = \inf\{t \geq 1 : S_t = 0\}.$$

We will have to deal with a non-homogeneous random walk.

4.5 Erdős-Rényi graph : proof of theorem 4.13

4.5.1 Proof of theorem 4.13(i)

Step one : coupling from above. Let $G = G_n$ is an Erdős-Rényi graph with distribution $\mathcal{G}(n, \lambda/n)$ and $0 < \lambda < 1$. We consider the exploration procedure (3.5) started from $v \in [n]$. We introduce the filtration $\mathcal{F}_t = \sigma((A_0, U_0, C_0), \cdots, (A_t, U_t, C_t))$. The hitting time τ is a stopping time for this filtration. Also, for integer $t \geq 0$, given $\mathcal{F}_t$, if $\{t < \tau\}$, X_{t+1} has distribution a binary random variable $\mathrm{Bin}(|U_t|, \lambda/n)$. In particular, if ξ_{t+1} is given $\mathcal{F}_t$, a binary variable $\mathrm{Bin}(n - |U_t|, \lambda/n)$ independent of X_t. Then $Y_{t+1} = X_{t+1} + \xi_{t+1}$ is a binary variable $\mathrm{Bin}(n, \lambda/n)$. In particular,

$$\sum_{i=1}^{t\wedge\tau} X_i \leq \sum_{i=1}^{t\wedge\tau} Y_i.$$

It follows that

$$\tau \leq \tau_+ = \inf \left\{ t \geq 1 : 1 + \sum_{i=1}^{t} (Y_i - 1) = 0 \right\}. \tag{4.5}$$

Step two : fast extinction. Now, from corollary 4.7, we deduce that

$$\mathbb{P}(\tau_+ \geq t) \leq \lambda^{-1} e^{-\alpha t}. \tag{4.6}$$

Let $c > 1/\alpha$. It follows that, for $v \in [n]$,

$$\mathbb{P}(\tau \geq c \log n) = \mathbb{P}(|G(v)| \geq c \log n) \leq \lambda^{-1} n^{-\alpha c}.$$

The union bond yields to

$$\mathbb{P}\left(\max_{v \in [n]} |G(v)| \geq c \log n \right) \leq \lambda^{-1} n^{1-\alpha c}.$$

We obtain theorem 4.13(i).

4.5.2 Proof of theorem 4.13(ii)

Step one : coupling from below. This time we shall try to lower bound X_t. We assume that $\lambda > 1$. Let $1/2 < \beta < 1$, we define the stopping time

$$\tau_\beta = \tau \wedge \inf \left\{ t \geq 1 : \sum_{i=1}^{t} X_i \geq 2n^\beta \right\}.$$

Also, for integer $t \geq 0$, given $\mathcal{F}_t$, if $\{t < \tau_\beta\}$, X_{t+1} has distribution a binary random variable $\mathrm{Bin}(|U_t|, \lambda/n)$ and $|U_t| \geq n - 2n^\beta$. In particular, on the event if $\{t < \tau_\beta\}$, we may define $Z_{t+1} = \sum_u \mathbf{1}_{\{v_{t+1}, u\} \in E}$, where the sum is over the first $m = n - \lceil 2n^\beta \rceil$ elements of U_t in lexicographic order. By construction, given $\mathcal{F}_t$, Z_{t+1} is a binary variable $\mathrm{Bin}(m, \lambda/n)$ and $X_{t+1} \geq Z_{t+1}$. In particular,

$$\sum_{i=1}^{t \wedge \tau_\beta} Z_i \leq \sum_{i=1}^{t \wedge \tau_\beta} X_i. \tag{4.7}$$

Step two : fast extinction or long survival. For ease of notation for any positive real, we set $A_t(v) = A_{\lfloor t \rfloor}(v)$ where we write $A_t(v)$ in place of A_t to explicit the dependence of the starting point in the exploration procedure. We are first going to prove with probability tending to 1, for all vertices v, either $|G(v)| \leq c_1 \log n$ or $|A_{n^\beta}(v)| \geq c_2 n^\beta$, where c_1, is a positive constants that will be chosen later and any $0 < c_2 < 1 \wedge (\lambda - 1)$. Note in particular that this implies that for all vertices either $|G(v)| \leq c_1 \log n$ or $|G(v)| \geq c_2 n^\beta$. The complement of this event is contained in the event

$$\Omega_n = \left\{ \exists v \in [n] : A_{c_1 \log n}(v) \neq \emptyset \text{ and } \exists\, c_1 \log n \leq t \leq n^\beta : |A_t(v)| \leq c_2 t \right\}.$$

From the union bond, its probability is upper bounded by

$$\begin{aligned}\mathbb{P}(\Omega_n) &\leq n\,\mathbb{P}\left(A_{c_1\log n}\neq\emptyset \text{ and } \exists\, c_1\log n\leq t\leq n^\beta : |A_t|\leq c_2 t\right)\\ &\leq n\,\mathbb{P}\left(A_{c_1\log n}\neq\emptyset \text{ and } \exists\, c_1\log n\leq t\leq n^\beta : |A_{t\wedge\tau_\beta}|\leq c_2 (t\wedge\tau_\beta)\right). \qquad (4.8)\end{aligned}$$

Indeed, if for some integer t, $\sum_{i=1}^t X_i \geq 2n^\beta$ then for all $t\leq s\leq n^\beta$, $|A_s|\geq 1+2n^\beta - s > s$ (recall that $|A_t| = 1-t+\sum_{i=1}^t X_i$). We may thus use (4.7),

$$\mathbb{P}\left(A_{c_1\log n}\neq\emptyset \text{ and } \exists\, c_1\log n\leq t\leq n^\beta : |A_{t\wedge\tau_\beta}|\leq c_2 (t\wedge\tau_\beta)\right) \leq \sum_{t=\lceil c_1\log n\rceil}^{\infty}\mathbb{P}\left(\sum_{i=1}^t Z_i \leq (1+c_2)t\right).$$

We define $\lambda' = \mathbb{E}Z_1 = \frac{m\lambda}{n} = \lambda(1-\lceil 2n^\beta\rceil/n)$, then for all n large enough, $\lambda'-1$ is larger than c_2. It follows

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^t Z_i \leq (1+c_2)t\right) &= \mathbb{P}\left(\sum_{i=1}^t (Z_i-\lambda') \leq -t(\lambda'-1-c_2)\right)\\ &\leq e^{-\lambda' t\gamma\left(\frac{\lambda'-1-c_2}{\lambda'}\right)},\end{aligned}$$

where we have applied corollary 4.9. From (4.8), it follows easily that

$$\mathbb{P}(\Omega_n)\leq \frac{n^{1-c_1\lambda'\gamma\left(\frac{\lambda'-1-c_2}{\lambda'}\right)}}{1-n^{1-c_1\lambda'\gamma\left(\frac{\lambda'-1-c_2}{\lambda'}\right)}}.$$

Now as n goes to infinity, λ' converges to λ. Thus, if we pick some $c_1 > 1/(\lambda\gamma(\lambda-1-c_2)/\lambda))$, we have proven that with probability tending to 1, for all vertices v, either $|G(v)|\leq c_1\log n$ or $|A_{n^\beta}(v)|\geq c_2 n^\beta$.

More generally, for any $a>0$, the constant c_1 can be taken large enough so that Ω_n has probability $O(n^{-a})$.

Step three : at most one giant component. Assume that Ω_n^c holds and that there are two vertices u,v such that $|G(u)|\geq c_1\log n$ and $|G(v)|\geq c_1\log n$. Then, either the exploration processes will intersect by step n^β and $G(v)=G(u)$ or they have disjoint active sets $A_t(u), A_s(v)$, for all $0\leq s,t\leq n^\beta$ and $A_{n^\beta}(u)$, $A_{n^\beta}(v)$ have cardinal at least $c_2 n^\beta$. In such case, given $(A_{n^\beta}(u), C_{n^\beta}(u), A_{n^\beta}(v), C_{n^\beta}(v))$, the probability that there is no edge between $A_{n^\beta}(u)$ and $A_{n^\beta}(v)$ is

$$\left(1-\frac{\lambda}{n}\right)^{|A_{n^\beta}(u)||A_{n^\beta}(v)|}\leq\left(1-\frac{\lambda}{n}\right)^{c_2^2 n^{2\beta}}\leq \exp\left(-\lambda c_2^2 n^{2\beta-1}\right).$$

Hence, since $1/2<\beta<1$, we deduce that $G(u)=G(v)$ with probability tending to 1. Thus the probability that there is at least two components of size at least $c_1\log n$ is upper bounded by

$$\mathbb{P}(\Omega_n) + n^2\exp\left(-\frac{\lambda c_2^2 n^{2\beta-1}}{2}\right),$$

which goes to 0.

We will call the largest connected component of the graph, the *giant component* of the graph. We have however not checked yet that there exists with high probability a component of size at least n^β.

Step four : expected size of the giant component. Let n be an integer large enough such that $c_1 \log n \geq 2n^\beta$ and $\tau_- = \inf\{t \geq 1 : 1 + \sum_{s=1}^t (Z_s - 1) = 0\}$. From (4.7), we note also that $\mathbb{P}(|G(v)| \geq c_1 \log n) \geq \mathbb{P}(\tau_- \geq c_1 \log n)$. Also in section 4.2, we have seen that τ_- has the same distribution than the total population in a branching process with offspring distribution $Z_1 = \mathrm{Bin}(m, \lambda/n)$. If $\rho_- > 0$ is the probability of extinction of this branching process, it follows that

$$\mathbb{P}(|G(v)| \geq c_1 \log n) \geq 1 - \rho_-.$$

Similarly, from (4.5),

$$\mathbb{P}(|G(v)| \geq c_1 \log n) \leq \mathbb{P}(\tau_+ \geq c_1 \log n) = 1 - \rho_+ - \mathbb{P}(c_1 \log n \leq \tau_+ < \infty),$$

where ρ_+ is the probability of extinction of a branching process with offspring distribution $Y = \mathrm{Bin}(n, \lambda/n)$. Remark that if $\tau_+ = t$ then

$$1 + \sum_{s=1}^t (Y_s - \lambda) = -t(\lambda - 1).$$

Hence, by corollary 4.9,

$$\begin{aligned}
\mathbb{P}(c_1 \log n \leq \tau_+ < \infty) &\leq \sum_{t=c_1 \log n}^{\infty} \mathbb{P}\left(\sum_{s=1}^t (Y_s - \lambda) \leq -t(\lambda - 1)\right) \\
&\leq \sum_{t=\lceil c_1 \log n \rceil}^{\infty} e^{-\lambda t \gamma\left(\frac{\lambda-1}{\lambda}\right)} \\
&\leq \frac{n^{1 - c_1 \lambda \gamma\left(\frac{\lambda-1}{\lambda}\right)}}{1 - n^{1 - c_1 \lambda \gamma\left(\frac{\lambda-1}{\lambda}\right)}}.
\end{aligned}$$

For our choice of c_1, the above expression goes to 0.

Recall that from (2.7), the binary random variables $\mathrm{Bin}(n, \lambda/n)$ and $\mathrm{Bin}(m, \lambda/n)$, $m = n - 2n^\beta$, converge weakly to a Poisson random variable as n goes to infinity. Hence, by lemma 4.5, as n goes to infinity, ρ_- and ρ_+ converge to ρ, where ρ is given by (4.2). It yields that for any v,

$$\lim_n \mathbb{P}(|G(v)| \geq c_1 \log n) = 1 - \rho.$$

In particular the expected size of the giant component is equivalent to $(1 - \rho)n$.

Step five : a.s. size of the giant component. Now, it remains to improve this convergence. Let $I_v = \mathbf{1}_{|G(v)|\leq c_1 \log n}$ and $L_n = \sum_{v=1}^n \mathbf{1}_{|G(v)|\leq c_1 \log n}$, we have already proved that

$$\lim_{n\to\infty} \frac{\mathbb{E}L_n}{n} = 1-\rho.$$

The proof of theorem 4.13(ii) will be complete if we prove that a.s.

$$\lim_{n\to\infty} \frac{L_n - \mathbb{E}L_n}{n} = 0. \tag{4.9}$$

We may use a concentration inequality. Note that removing an edge $e = \{u,v\}$ of the graph G cannot decrease the function $L_n(G)$. Moreover, if $G' = G - e$ is the graph where the edge has been removed, we find

$$L_n(G') - L_n(G) \leq |G'(u)|\mathbf{1}_{|G'(u)|\leq c_1 \log n} + |G'(v)|\mathbf{1}_{|G'(v)|\leq c_1 \log n} \leq 2c_1 \log n.$$

We may thus use remark 3.25 for $\delta = \log n$, $\theta = 3$ and $c = 2c_1 \log n$. Statement (4.9) follows from Borel-Cantelli lemma. $\square$

4.6 Configuration Model : : proof of theorem 4.14

4.6.1 Proof of theorem 4.14(i)

Step one : coupling from above. Let $G = G_n$ with distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$. We consider now the exploration procedure (3.8) starting from $v \in [n]$. We set $\tau = \inf\{t \geq 1 : |A_t| = 0\}$, for $0 \leq t \leq \tau - 1$, $\varepsilon_{t+1} = \mathbf{1}_{v_{t+1}\in A_t}$ and $\varepsilon_t = 0$ for $t > \tau$. Again, we define $X_t = |I_t|$ and $S_t = |A_t|$. So that

$$S_t = d_v + \sum_{k=1}^t (X_k - 1 - \varepsilon_k) \ , \quad |U_t| = |\Delta| - d_v - \sum_{k=1}^t (X_k + 1 - \varepsilon_k),$$

and

$$|G(v)| = 1 + \tau - \sum_{t\geq 1} \varepsilon_t.$$

We also set

$$E_t = \sum_{k=1}^t \varepsilon_k.$$

We consider the filtration $\mathcal{F}_t = \sigma((A_0, U_0, C_0), \cdots, (A_t, U_t, C_t))$. The hitting time τ is a stopping time for this filtration. We recall also that $|U_t| + |A_t| = |\Delta| - |C_t| = |\Delta| - 2t$ and from (3.9), for every $k \geq 1$

$$\mathbb{P}(X_{t+1} = k|\mathcal{F}_t) = \begin{cases} \sum_{u\in U_t} \frac{\mathbf{1}_{d_u=k+1}(k+1)}{|\Delta|-2t-1} & k \geq 1, \\ \sum_{u\in U_t} \frac{\mathbf{1}_{d_u=1}}{|\Delta|-2t-1} + \frac{|A_t|-1}{|\Delta|-2t-1} & k = 0, \end{cases} \tag{4.10}$$

Hence, as for Erdős-Rényi random graphs, we have to deal with a non-homogeneous random walk. We will rely on coupling techniques. The argument will be slightly more involved. Let $\alpha < \beta < 1$ be a real number that we will chose later on. We order the sequence set $\mathbf{d} = (d_1, \cdots d_n)$ in non-decreasing order, we get a permutation π of $[n]$ such that $\mathrm{d}_{\pi(1)} \geq d_{\pi(2)} \cdots \geq d_{\pi(n)}$. Let n_0 be the number of vertices with degree different from 0. We then define the set

$$\Pi_+ = \{\pi(i) : 1 \leq i \leq n_0 - n_0^\beta\}. \tag{4.11}$$

This is the subset of vertices with the $n_0 - n_0^\beta$ larger degrees. We denote by $\Delta_+ = \cup_{i \in \Pi_+} \Delta_i$ and Q_+ denote the distribution on integers,

$$Q_+(k) = \frac{k+1}{|\Delta_+|} \sum_{i \in \Pi_+} \mathbf{1}_{d_i = k+1}, \qquad \text{for } k \geq 0.$$

We first define a sequence $(Y_t)_{t \geq 1}$ of i.i.d. variables with distribution Q_+, such that for all $1 \leq t \leq n_0^\beta$,

$$X_{t \wedge \tau} \leq Y_{t \wedge \tau}.$$

This is done explicitly by setting $Y_{t+1} = d_{u_{t+1}} - 1$ for some random $u_{t+1} \in \Pi_+$ such that $\mathbb{P}(u_{t+1} = u | \mathcal{F}_t) = d_u / |\Delta_+|$. We order decreasingly the half-edges from 1 to Δ, by setting

$$(\pi(1), 1) \succ (\pi(1), 2) \cdots \succ (\pi(1), d_{\pi(1)}) \succ (\pi(2), 1) \cdots \succ (\pi(n), d_{\pi(n)}).$$

In particular, Δ_+ is the set of $|\Delta_+|$ largest half-edge of Δ. We notice that $|\Delta_+| \leq |\Delta| - n_0^\beta$ and recall that $|U_t \cup A_t| = |\Delta| - 2t$. Now, let $1 \leq t \leq \tau \wedge n_0^\beta/2$, if $\sigma(e_{t+1})$ is the k-th largest half-edge of $U_t \cup A_t$ and $k \leq |\Delta_+|$ then we define u_{t+1} as the vertex such that the k-th largest half-edge of Δ. Otherwise, $d_{v_{t+1}}$ is smaller or equal than any degrees in Π_+ and we define u_{t+1} as the vertex such that the N-th largest half-edge of Δ_+ where N is an independent variable uniformly distributed in Δ_+. It follows easily that $\mathbb{P}_{\mathbf{d}}(Y_{t+1} \in \cdot | \mathcal{F}_t) = Q_+$ and $X_t \leq Y_t$.

From what precedes, for $1 \leq t \leq n_0^\beta/2$,

$$\sum_{i=1}^{t \wedge \tau} X_i \leq \sum_{i=1}^{t \wedge \tau} Y_i.$$

We set

$$\tau_+ = \inf \left\{ t \geq 1 : 1 + \sum_{i=1}^{t} (Y_i - 1) = 0 \right\}.$$

It follows that for all $1 \leq t \leq n_0^\beta/2$,

$$\{\tau \geq t\} \subset \{\tau_+ \geq t\}. \tag{4.12}$$

Step two : fast extinction. Now, in Π_+, we have removed the n_0^β smallest positive degrees. By assumption (H_0),

$$\lim_n \frac{n_0}{n} = \mathbb{P}(D \geq 1) > 0. \tag{4.13}$$

Also, there exists $\tau \geq 1$ such that $q = \mathbb{P}(1 \leq D \leq \tau) > 0$. Hence, by assumption (H_0), for all n large enough,

$$\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(1 \leq d_i \leq \tau) > q/2.$$

In particular, for all $i \in \{1, \cdots, n\} \backslash \Pi_+$ and n large enough, $d_i \leq \tau$. We deduce that

$$|\Delta_+| \geq |\Delta| - \sum_{i=1}^{n} d_i \mathbf{1}(i \notin \Pi_+) \geq \sum_{i=1}^{n} d_i - \tau n_0^\beta.$$

By assumption (H_1), it yields

$$\lim_{n\to\infty} \frac{|\Delta_+|}{n} = \mathbb{E}D.$$

From assumption (H_2) and the definition of Q_+, we have proved that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[Y] &= \lim_{n\to\infty} \sum_{i\in\Pi_+} \frac{d_i(d_i-1)}{|\Delta_+|} = \frac{\mathbb{E}D(D-1)}{\mathbb{E}D} < 1, \\
\lim_{n\to\infty} \mathbb{E}[Y^\alpha] &= \lim_{n\to\infty} \sum_{i\in\Pi_+} \frac{d_i(d_i-1)^\alpha}{|\Delta_+|} = \frac{\mathbb{E}D(D-1)^\alpha}{\mathbb{E}D}.
\end{aligned}$$

We use the inequality $\mathbb{E}[|Y-1|^\alpha] \leq 2^{\alpha-1}(\mathbb{E}Y^\alpha + 1)$. Then, Markov inequality implies that for any $c > 2^{\alpha-1}(\frac{\mathbb{E}D(D-1)^\alpha}{\mathbb{E}D} + 1)$, for all n large enough and $t \geq 1$, $\mathbb{P}[|Y-1| \geq t] < ct^{-\alpha}$. This implies for all $1 < \alpha' < \alpha$ that the sequence of distributions $\mathbb{P}(|Y-1|^{\alpha'} \in \cdot)$ are uniformly integrable in $\mathbf{d} = (d_1, \cdots, d_n)$, $n \in \mathbb{N}$.

We may thus apply theorem 4.10 to the variables $(Y-1)$ and the scalar α' (see remark 4.12). We get from (4.12) that for some constant $c_1 > 0$, there exists n_1 such that for all $n \geq n_1$, and $1 \leq t \leq n^\beta \mathbb{P}(D > 0)/4$,

$$\mathbb{P}(\tau \geq t) \leq c_1 t^{-\alpha'}.$$

Now, let $1/\alpha < c < 1$. We could have chosen β and α' such that $1/\alpha' < c$ and $c < \beta < 1$, then for all n all large enough $n^c \leq n^\beta \mathbb{P}(D > 0)/2$. We may thus apply the above inequality to $t = n^c$, from the union bond, for all n large enough,

$$\mathbb{P}\left(\max_{v\in[n]} |G(v)| \geq n^c\right) \leq c_1 n^{1-\alpha' c}.$$

We obtain theorem 4.13(i).

4.6.2 Proof of theorem 4.14(ii)

Step one : coupling from below. Let $1/2 < \beta < 1$, we define again the stopping time

$$\tau_\beta = \tau \wedge \inf\left\{ t \geq 1 : d_v + \sum_{i=1}^t X_i \geq 4n^\beta \right\}.$$

We may assume that n is large enough so that $|\Delta| \geq n_0 > 4n^\beta$. As abocve, we consider the ordering $\prec$ on the set Δ. We define Δ_- as the $|\Delta| - \lceil 4n^\beta \rceil$ smaller terms of Δ. We set

$$\Pi_- = \{i \in [n] : \exists 1 \leq j \leq d_i, (i,j) \in \Delta_-\}.$$

If $|\Pi_-| = m$, then Π_- is the subset of vertices with the m smaller degrees.

We introduce the distribution on integers,

$$Q_-(k) = \frac{1}{|\Delta_-|} \sum_{(i,j)\in\Delta_-} \mathbf{1}_{d_i = k+1} = \frac{1}{|\Delta_-|} \sum_{i \in \Pi_-} |\Delta_i \cap \Delta_-| \mathbf{1}_{d_i=k+1}, \quad \text{for } k \geq 0.$$

We define two independent sequences $(W_t)_{t\geq 1}, (\zeta_t)_{t \geq 1}$ of i.i.d. variables with distribution Q_- and Bernoulli

$$\mathbb{P}(\zeta_t = 1) = 1 - \mathbb{P}(\zeta_t = 0) = \frac{|\Delta_-|}{|\Delta|},$$

such that for all integer $t \geq 1$,

$$X_{t\wedge\tau_\beta} \geq W_{t\wedge\tau_\beta} \zeta_{t\wedge\tau_\beta}. \tag{4.14}$$

This is done explicitly by setting $W_{t+1} = (d_{u_{t+1}} - 1)\mathbf{1}_{u_{t+1}}$ for some $u_{t+1} \in \Pi_-$ such that

$$\mathbb{P}(u_{t+1} = u | \mathcal{F}_t) = \frac{|\Delta_u \cap \Delta_-|}{|\Delta_-|}.$$

Let $0 \leq t < \tau_\beta$, we first notice that $|U_t| = |\Delta| - d_v - \sum_{i=1}^t X_i \geq |\Delta| - \lceil 4n^\beta \rceil = |\Delta_-|$. Now if $\sigma(e_{t+1})$ is the k-th smallest half-edge of U_t and $k \leq |\Delta_-|$ then we define u_{t+1} as the vertex such that the k-th smallest half-edge of Δ_- is in $\Delta_{u_{t+1}}$. This event $\{k \leq |\Delta_-|\}$ happens with probability $|\Delta_-|/(|\Delta| - t - E_t) \geq |\Delta_-|/|\Delta|$. Conditioned on this event, we set $\zeta_{t+1} = 1$ with probability $(|\Delta| - t - E_t)/|\Delta|$ and $\zeta_{t+1} = 0$ otherwise. On the contrary if $k \geq |\Delta_-|$, then we choose $\mathbb{P}(u_{t+1} = u|\mathcal{F}_t) = |\Delta_u \cap \Delta_-|/|\Delta_-|$ independently of X_{t+1} and we set $\zeta_{t+1} = 0$. It follows easily that $\mathbb{P}(W_{t+1} = k, \zeta_{t+1} = 1 | \mathcal{F}_t) = Q_-(\{k\})|\Delta_-|/|\Delta|$ and $X_t \geq W_t \zeta_t$.

We have

$$\frac{1}{|\Delta_-|}\sum_{i=1}^n d_i(d_i-1) - \frac{1}{|\Delta_-|}\sum_{(i,j)\in\Delta\backslash\Delta_-}(d_i - 1) \leq \mathbb{E}W_1 \leq \frac{1}{|\Delta_-|}\sum_{i=1}^n d_i(d_i-1).$$

By (4.13), for n large enough, if $1 \leq i \leq \lceil 4n^\beta \rceil$, we have $d_{\pi(i)} \geq 1$. it follows that

$$\frac{1}{|\Delta|}\sum_{i=1}^n d_i(d_i-1) - \frac{1}{|\Delta|}\sum_{i=1}^{\lceil 4n^\beta \rceil} d_{\pi(i)}(d_{\pi(i)}-1) \leq \mathbb{E}[W_1\zeta_1] \leq \frac{1}{|\Delta|}\sum_{i=1}^n d_i(d_i-1).$$

Then if P has support included in $[0,\kappa]$, we have $\frac{1}{|\Delta|}\sum_{i=1}^{n^\beta} d_{\pi(i)}(d_{\pi(i)}-1) \leq \frac{4n^\beta\kappa^2}{|\Delta|}$ converges to 0. Otherwise the support is infinite and, for all $\kappa > 0$, the event $\{d_{\pi(\lceil 4n^\beta\rceil)} > \kappa\}$ holds for n large enough (indeed by assumption (H_0), a positive fraction of degrees is larger that κ). Then, from assumption (H_2), for all $\varepsilon > 0$, there exists $\kappa > 0$ such that $\mathbb{E}D(D-1)\mathbf{1}_{D>\kappa} \leq \varepsilon$. In particular, $\limsup_n \frac{1}{n}\sum_{i=1}^{n^\beta} d_{\pi(i)}(d_{\pi(i)}-1) \leq \limsup_n \frac{1}{n}\sum_{i=1}^{n} d_i(d_i-1)\mathbf{1}_{d_i>\kappa} \leq \varepsilon$. This last bound holding for all $\varepsilon > 0$, we deduce that

$$\lim_n \frac{1}{n}\sum_{i=1}^{n^\beta} d_{\pi(i)}(d_{\pi(i)}-1) = 0.$$

We have thus checked that for all κ large enough,

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}[W_1\zeta_1] &= \frac{\mathbb{E}D(D-1)}{\mathbb{E}D} > 1,\\
\lim_{n\to\infty}\mathbb{E}[W_1\zeta_1\mathbf{1}_{\{W_1\leq\kappa\}}] &= \frac{\mathbb{E}[D(D-1)\mathbf{1}_{\{D<\kappa\}}]}{\mathbb{E}D} > 1.
\end{aligned}$$

For ease of notation, we set

$$Z_t = W_t\zeta_t\mathbf{1}_{\{W_t\leq\kappa\}}.$$

and define Q'_- as the distribution of Z. We may assume that n is large enough to guarantee that $\mathbb{E}[Z] > 1$.

Step two : fast extinction or long survival. As for Erdős-Rényi graphs, we are first going to prove with probability tending to 1, for all vertices v, either $|G(v)| \leq c_1\log n$ or $|A_{n^\beta}(v)| \geq c_2 n^\beta$, where c_1, is a positive constants that will be chosen later and $c_2 = 1 \wedge \frac{\mathbb{E}[Z]-1}{2}$. We may upper bound the probability of the complement of this event by (4.8). Arguing as for Erdős-Rényi graphs, we get

$$\begin{aligned}
\mathbb{P}\left(A_{c_1\log n} \neq \emptyset; |A_{t\wedge\tau_\beta}| \leq c_2 t\wedge\tau_\beta\right) &\leq \mathbb{P}\left(\sum_{i=1}^{t} Z_i \leq (1+c_2)t\right)\\
&= \mathbb{P}\left(\sum_{i=1}^{t}(W_i - \mathbb{E}[Z]) \leq -t\frac{\mathbb{E}[Z]-1}{2}\right)\\
&\leq \exp\left(-\frac{t(\mathbb{E}[Z]-1)^2}{8\kappa^2}\right).
\end{aligned}$$

Where we have applied Hoeffding's inequality (3.13). From (4.8), it follows easily that

$$\mathbb{P}\left(\exists v\in[n] : A_{c_1\log n}(v) \neq \emptyset \text{ and } \exists\, c_1\log n \leq t \leq n^\beta : |A_t(v)| \leq c_2 t\right) \leq \frac{n^{1-c_1(\mathbb{E}[Z]-1)^2/(8\kappa^2)}}{1-e^{-(\mathbb{E}[Z]-1)^2/(8\kappa^2)}}.$$

Now as n goes to infinity, $\mathbb{E}[Z]$ converges to $\lambda = \frac{\mathbb{E}[D(D-1)\mathbf{1}_{\{D<\kappa\}}]}{\mathbb{E}D} > 1$. Thus, if we chose some $c_1 > (8\kappa^2)/(\lambda-1)^2$, we have proven that with probability tending to 1, for all vertices v, either $|G(v)| \leq c_1\log n$ or $|A_{n^\beta}(v)| \geq c_2 n^\beta$.

Step three : at most one giant component. Assume there are two vertices u, v such that $|G(u)| \geq c_1 \log n$ and $|G(v)| \geq c_1 \log n$. Then, either the exploration processes will intersect by step n^β or they have two disjoint active sets $A_{n^\beta}(u)$, $A_{n^\beta}(v)$ of cardinal at least $c_2 n^\beta$.

Indeed, assume that this event holds. We order the half-edges of $A_{n^\beta}(u)$ by lexicographic order. We pick the smallest half-edge of $A_{n^\beta}(u)$, say e_1, the probability that e_1 is not matched to an element of $A_{n^\beta}(v)$ is $1 - |A_{n^\beta}(v)|/(|\Delta| - n^\beta - E_{n^\beta}(v)) \leq 1 - c_2 n^\beta/|\Delta|$. Then, let e_2 be the smallest half-edge of $A_{n^\beta}(u) \backslash \{e_1, \sigma(e_1)\}$. Then given e_1 is not matched to an element of $A_{n^\beta}(v)$, the probability that e_2 is not matched to an element of $A_{n^\beta}(v)$ is $1 - |A_{n^\beta}(v)|/(|\Delta| - n^\beta - E_{n^\beta}(v) - 2) \leq 1 - c_2 n^\beta/|\Delta|$. We may continue this process for at least $c_2 n^\beta/2$ steps. We get that the probability that there is no matching between $A_{n^\beta}(u)$ and $A_{n^\beta}(v)$ is upper bounded by

$$\left(1 - \frac{c_2 n^\beta}{|\Delta|}\right)^{\frac{c_2 n^\beta}{2}} \leq \exp\left(-\frac{c_2^2 n^{2\beta}}{2|\Delta|}\right).$$

Hence, since $1/2 < \beta < 1$ and $\lim_n |\Delta|/n = \mathbb{E}D$, we deduce that $G(u) = G(v)$ with probability tending to 1. Thus with probability tending to 1 there is at most a unique giant component of size at least n^β.

Step four : expected size of the giant component. We note also that by comparison with $(Z_t)_t$ that

$$\mathbb{P}(|G(v)| \geq c_1 \log n) \geq 1 - \rho_-(v),$$

where $\rho_-(v)$ is the probability of extinction of a branching process where the progenitor has d_v offsprings and all other genitors have offspring distribution Q'_-. We have $\rho_-(v) = \rho_-^{d_v}$, where ρ_- is the probability of extinction in a Galton-Watson process with offspring distribution Q'_-.

Similarly,

$$\mathbb{P}(|G(v)| \geq c_1 \log n) \leq 1 - \rho_+^{d_v} - \mathbb{P}(c_1 \log n < \tau_+ < \infty).$$

We argue as in the proof of theorem 4.13(ii). Since $Y \geq -1$, $\varphi(\theta) = \mathbb{E}e^{\theta Y}$ is well defined for all $\theta < 0$. We find, from Chernov bound, for any $\theta < 0$ and integer $x > 0$,

$$\begin{aligned}
\mathbb{P}(x < \tau_+ < \infty) &\leq \sum_{t=x}^{\infty} \mathbb{P}\left(\sum_{s=1}^{t} Y_s \leq t\right) \\
&\leq \sum_{t=x}^{\infty} \varphi(\theta)^t e^{-t\theta}.
\end{aligned}$$

Moreover, for any $\varepsilon > 0$, for all $\theta \in (a_\varepsilon, 0]$ close enough to 0, $\varphi(\theta) \leq 1 + \theta(\mathbb{E}Y - \varepsilon)$. Choosing $0 < \varepsilon < \mathbb{E}Y - 1$, for some $\theta < 0$, we get

$$\varphi(\theta)^t e^{-t\theta} \leq (1 + \theta(\mathbb{E}Y - \varepsilon))^t e^{-t\theta} \leq e^{t\theta(\mathbb{E}Y - \varepsilon - 1)}.$$

In particular $\mathbb{P}(c_1 \log n < \tau_+ < \infty)$ decreases polynomially to 0.

Now, Q_+ converges weakly to $\widehat{P}$ and Q'_- converges weakly to the distribution $\widehat{P}'$ on $\{0, \cdots, \kappa\}$, defined by $\widehat{P}'(\{k\}) = \widehat{P}(\{k\})$ for $1 \leq k \leq \kappa$ and $\widehat{P}'(\{0\}) = \widehat{P}(\{0\}) + \widehat{P}([\kappa+1, \infty))$.

We note finally that for any integer d_v and $x, y \in [0,1]$,

$$|x^{d_v} - y^{d_v}| \leq |x - y|.$$

Hence, letting n tend to infinity and then κ, using lemma 4.5, we have checked that

$$\lim_{n\to\infty} \max_{v\in[n]} \left| \mathbb{P}(|G(v)| < c_1 \log n) - \widehat{\rho}^{d_v} \right| = 0.$$

Summing over all n and using (H_0), it yields to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{v\in[n]} \mathbb{P}(|G(v)| \geq c_1 \log n) = 1 - \rho.$$

Step five : a.s. size of the giant component. Now, it remains to improve the convergence. The concentration argument used in the proof of theorem 4.13 works in this case also. It suffices to replace remark 3.25 by remark 3.31. □

4.7 Application to network epidemics

4.7.1 A simple SIR dynamic

Network epidemics gives an insightful application to the emergence of a giant component in a graph. Let $G = ([n], E)$ be a finite graph on $[n]$. The propagation of an epidemic in the graph is classically modeled as follows. Each vertex has a state either (S)usceptible, (I)nfected or (R)esilient. The state of the network at discrete time $t \in \mathbb{N}$ is $X_t = (S_t, I_t, R_t)$, where S_t, I_t and R_t is the set of vertices in state S, I or R at time t. The evolution is as follows : any vertex in state I at time $t \in \mathbb{N}$ becomes R at time $t+1$ and each of its neighbors in G in state S becomes I with probability $p \in (0,1)$ independently. To keep the model simple, we assume that at time $t = 0$, a single vertex, say $1 \in [n]$, is infected : $X_0 = ([n]\backslash\{1\}, \{1\}, \emptyset)$.

More formally, let $(\xi_{\{i,j\}})_{i,j\in[n]}$ be a collection of i.i.d. random variable with Bernoulli distribution $\mathbb{P}(\xi_e = 1) = 1 - \mathbb{P}(\xi_e = 0) = p$. The process $(X_t)_{t\in\mathbb{N}}$ is a Markov chain on the set of partitions of $[n]$ in 3 sets : $X_{t+1} = (S_{t+1}, I_{t+1}, R_{t+1})$, with

$$I_{t+1} = \bigcup_{v\in I_t} \{u \in S_t : \{u,v\} \in E, \xi_{\{u,v\}} = 1\}, \qquad S_{t+1} = S_t \backslash I_{t+1}, \qquad R_{t+1} = R_t \cup I_t.$$

This defines a Markov chain because each random variable ξ_e is used at most once. Recall that an *absorbing state* of a Markov chain is a state such that $\mathbb{P}(X_1 = x | X_0 = x) = 1$. Here, the absorbing states are the states $x = (s, \emptyset, r)$ with $s \cap r = \emptyset$, $s \cup r = [n]$. From Kolmogorov $0-1$

law, the probability that $\mathbb{P}((X_t)$ reaches an absorbing state $|X_0 = x) \in \{0,1\}$. For any state $x = (s,i,r)$ the probability $\mathbb{P}(X_1$ is an absorbing state $|X_0 = x) > 0$. We deduce that with probability one, the chain $(X_t)_{t\in\mathbb{N}}$ reaches an absorbing state (without invoking Kolmogorov $0-1$ law, we could also notice that $\mathbb{P}(X_n$ is an absorbing state $|X_0 = x) = 1$).

Let $\tau = \inf\{t \geq 1 : I_t = \emptyset\}$ be the almost surely finite time the chain reaches an absorbing state. With our choice of initial condition, the set R_τ is the set of vertices that have been infected at some time before the epidemic stops. This pair (τ, R_τ) is random and the basic question in network epidemics is to analyze it. We denote by H_τ the subgraph of G spanned by the vertices in R_τ. We also define the percolation graph $G^p = (V, E^p)$ as the subgraph of G defined by $e = \{u,v\} \in E^p$ if and only if $e \in E$ and $\xi_e = 1$.

Assume for a moment that G is a tree. Remark then that with our choice of initial condition, for integer $t \geq 1$, R_t is the set of vertices at distance $t-1$ from 1 in G^p and I_t is the set of vertices at distance exactly t from 1. In particular H_τ is the connected component of G^p that contains 1.

More generally, even if G is not necessarily a tree, H_τ is also the connected component of G^p that contains 1. Indeed, if $v \in H_\tau$ then it has been infected at some time k. Let $i_k = v$ and i_{k-1} be a vertex that has infected v : $\{i_{k-1}, i_k\} \in E$ and $\xi_{\{i_{k-1},i_k\}} = 1$. By recursion, there exists a sequence $i_0, i_1, \cdots, i_k$ such that $i_0 = 1$, $i_k = v$ which is a path in G^p. The reciprocal goes along the same line.

4.7.2 Dynamic on the Erdős-Rényi graph

Now, assume that $G = G_n = K_n$ is the complete graph on n vertices. Then G_n^p has distribution $\mathcal{G}(n,p)$. More generally, if G_n is a random graph with distribution $\mathcal{G}(n,\lambda/n)$, independent of $(\xi_{\{i,j\}})$ then, G_n^p is a random graph with distribution $\mathcal{G}(n,\lambda p/n)$. In particular, we may apply theorem 3.12 : $(G_n^p, 1) = (H_\tau, 1)$ converges to a $\mathrm{GWT}(\mathrm{Poi}_{\lambda p})$. If $\lambda p < 1$, then $|R_\tau|$ converges to the total population in a Galton-Watson branching process with $\mathrm{Poi}_{\lambda p}$ offspring distribution whose tail distribution is sub-exponential as shown in corollary 4.9. Also, from equation (4.6),

$$\mathbb{P}(|R_\tau| \geq t) \leq (p\lambda)^{-1} e^{-\alpha t},$$

with $\alpha = p\lambda - 1 - \log(p\lambda)$.

Otherwise, $\lambda p > 1$ and by theorem 4.13(ii), there exists a giant component whose size is equivalent to $(1-\rho)n$, where ρ is given by (4.2) with λp replacing λ, and other connected component are of size $o(n)$. By exchangeability of the vertices, with probability $1-\rho$, vertex 1 belongs to the giant component. We deduce that a.s. $|R_\tau|/n$ converges weakly to $(1-\rho)\delta_{1-\rho} + \rho\delta_0$. More quantitatively, for any fixed $0 < \varepsilon < \rho$, with high probability, either $|R_\tau| \leq c \log n$ or $|R_\tau| \in ((1-\rho-\varepsilon)n, (1-\rho+\varepsilon)n)$. Thus there exists a sharp threshold at $\lambda p = 1$ on the behavior of the epidemic.

4.7.3 Dynamic on the configuration model

Now let P be a probability distribution on integers with positive finite second moment. We assume instead that $G = G_n$ has distribution $\widehat{\mathcal{G}}(\mathbf{d}_n)$ where $\mathbf{d}_n$ satisfies (H_2). We consider independent Bernoulli random variables (ξ_e) on the edges of the multi-graph, independent of G_n, $\mathbb{P}(\xi_e = 1) = 1 - \mathbb{P}(\xi_e = 0) = p$.

Now, conditioned on the degree sequence $\mathbf{d}_n^p$ of G_n^p, G_n^p has distribution $\widehat{\mathcal{G}}(\mathbf{d}_n^p)$. Note that $\mathbf{d}_n^p$ is a random degree sequence. It is not hard to check that a.s. $\mathbf{d}_n^p$ satisfies (H_2) with limit degree distribution

$$Q(k) = \sum_{\ell=k}^{\infty} P(\ell) \binom{\ell}{k} p^k (1-p)^{\ell-k}.$$

In other words, if M has distribution Q and N has distribution P, then $M = \sum_{i=1}^{N} \xi_i$, where (ξ_i) are independent Bernoulli variables.

Hence, by theorem 3.15, the rooted graph $[H_\tau, 1]$ converges weakly to $\mathrm{GWT}_*(Q)$. Denote by ψ the generating function of Q and φ the generating function of P : we have $\psi(z) = \varphi(pz + (1-p))$. From corollary 4.2, the threshold for non-extinction of a $\mathrm{GWT}_*(Q)$ is $\psi''(1) > \psi'(1)$, it can be rewritten has $p^2\varphi''(1) > p\varphi'(1)$ or

$$\mathbb{E}D\left(D - \frac{p+1}{p}\right) > 0.$$

where D has distribution P (indeed $\varphi'(1) = \mathbb{E}D$ and $\varphi''(1) = \mathbb{E}D(D-1)$). Therefore, if $\mathbb{E}D\left(D - \frac{p+1}{p}\right) < 0$, we deduce that $|R_\tau|$ converges to the size of a $\mathrm{GWT}_*(Q)$ whose tail distribution can be estimated by using theorem 4.10. On the contrary, if $\mathbb{E}D\left(D - \frac{p+1}{p}\right) > 0$, then we can adapt the argument of theorem 4.14(ii), $|R_\tau|/n$ converges a.s. to $(1-\rho)\delta_{1-\rho} + \rho\delta_0$ where ρ is given by (4.4) with ψ replacing φ and $\widehat{\psi} = \psi'(z)/\psi'(1)$ replacing $\widehat{\varphi}$.

Chapter 5

Continuous length combinatorial optimization

`To be continued...`

5.1 Issues of combinatorial optimization

Consider a finite network $G = (V, E, \omega)$ with marks $\omega(v)$, $\omega(e)$ in $\mathbb{R}_+$. We can conveniently think as such marks as lengths, weights, costs or rewards.

A *matching* M of G is a subset of edges $M \subset E$ such that no two edges in M have a common adjacent vertex. (Beware that this definition of a matching differs from the one we have already used in the context of configuration model). We denote by $\mathcal{M}(G)$ the set of matchings of G. The *maximal weight of a matching* of G is

$$\max_{M \in \mathcal{M}(G)} \sum_{e \in M} \omega(e). \tag{5.1}$$

A matching reaching the above maximum is called a *maximal matching*. For $\omega \equiv 1$, the above is called the *matching number* of G, it is simply the cardinal of a largest matching of G.

Define similarly, an *independent set* S of G is a subset of vertices $S \subset V$ such that no two vertices in S have a common adjacent edge. We denote by $\mathcal{I}(G)$ the set of matchings of G. The *maximal weight of an independent set* of G is

$$\max_{I \in \mathcal{I}(G)} \sum_{v \in I} \omega(v). \tag{5.2}$$

An independent set reaching the above maximum is called a *maximal independent set*. For $\omega \equiv 1$, the above is called the *independent set number* of G, it is the cardinal of largest independent set in G.

Assume that G is connected. A *spanning tree* T of G is a subtree of G with vertex set V. If $\mathcal{T}(G)$ is the set of spanning trees of G, the *minimal length of a spanning tree* of G is

$$\max_{T\in\mathcal{T}(G)} \sum_{e\in E} \omega(e)\mathbf{1}(e \in T). \tag{5.3}$$

A spanning tree reaching the above minimum is called a *minimal spanning tree (MST)*. If all weights are distinct, the MST is unique. We shall denote by $\mathrm{MST}(G)$ the minimal spanning tree of G.

From an algorithmic point of view, the three above network functionals are quite different. Finding a maximal weight independent set is an NP-hard problem, finding a maximal matching has complexity which is polynomial in the size of the network. Finally, they are greedy algorithms which find the minimal spanning tree of a network.

In this chapter, we will try to understand the links between local weak convergence and these network functionals. We should consider a sequence of finite networks having a local weak limit. Our main goal will be to compute the asymptotic value of these functions as the size of the networks grows large.

Note first that these functions are obviously invariant under network isomorphisms. Also, taking for example the MST, if $\rho = U(G)$ and

$$L(G) = \sum_{e\in E} \omega(e)\mathbf{1}(e \in \mathrm{MST}(G)).$$

is the total length of the MST, we find

$$\frac{L(G)}{|V|} = \frac{1}{2|V|}\sum_{v\in V}\sum_{e\in E: v\in e} \omega(e)\mathbf{1}(e\in\mathrm{MST}(G)) = \frac{1}{2}\mathbb{E}_\rho \sum_{e\in E:\varnothing\in e} \omega(e)\mathbf{1}(e\in\mathrm{MST}(G)),$$

where under ρ, $\varnothing$ is uniformly distributed on V.

This remark invites us to study the function on rooted networks

$$(G,\varnothing) \mapsto \sum_{e\in E:\varnothing\in e} \omega(e)\mathbf{1}(e\in\mathrm{MST}(G)).$$

We are however immediately confronted to the problem that it is not a priori obvious to define $\mathrm{MST}(G)$ on an arbitrary infinite network. We shall see that in some cases, it is possible to define in a natural way the combinatorial structures : maximal independent set, maximal matching and minimal spanning tree on infinite networks. There will be two strategies :

(i) give an explicit construction ;

(ii) give an iterative construction which is shown to converge for some networks.

5.2 Limit of random networks

In the context of our random graphs, there is a natural limit unimodular network, the *Galton-Watson network with degree distribution P and weights distribution Q*. Precisely, let $Q \in \mathcal{P}(\mathbb{R}_+)$ and $P \in \mathcal{P}(\mathbb{Z}_+)$ with finite positive first moment. Consider a Galton-Watson tree with degree distribution P. Put independently marks on edges and vertices which i.i.d. variables with law Q. We obtain this way a random rooted network. We shall denote by $\mathrm{GWN}_*(P,Q)$ the law on $\mathbb{G}_*(\mathbb{R}_+)$ of the equivalence class of this random rooted network.

Note that in our context, we will only care either about the weights on vertices (independent set) or on the edges (matchings, spanning trees).

Consider a sequence of finite networks $G_n = (V_n, E_n, \omega_n)$. The empirical distribution of the vertex and edge weights Q_n^v and Q_n^e are respectively

$$Q_n^v = \frac{1}{|V_n|} \sum_{v \in V_n} \delta_{\omega_n(v)} \quad \text{and} \quad Q_n^v = \frac{1}{|E_n|} \sum_{e \in E_n} \delta_{\omega_n(e)}$$

We shall say that the vertex or edge weights of G_n are *uniformly integrable* if Q_n^v or Q_n^e is uniformly integrable, i.e. if

$$\lim_{t\to\infty} \sup_{n\geq 1} \int |x| \mathbf{1}_{|x|\geq t} \, dQ_n^{v/e}(x) = 0.$$

For example, consider a random multi-graph $G_n \overset{d}{\sim} \widehat{\mathcal{G}}(\mathbf{d}_n)$ where $(\mathbf{d}_n)$ satisfies (H_p), for some $p > 2$. We could turn G_n into a network by adding independently i.i.d. weights on vertices and edges with common law Q. Then, by theorem 3.28, it is not hard to check that a.s. $U(G_n)$ converges weakly to $\mathrm{GWN}_*(P,Q)$.

5.3 The minimal spanning tree

The minimal spanning tree is an example of a problem of combinatorial optimization where it is possible to define explicitly the limit random structure. `To be continued...`

5.4 Maximal weight independent set

`To be continued...`

We now give a example of a combinatorial optimization which can be solved thanks to a fixed point analysis. As in (5.3), for a finite network G, we set

$$I(G) = \max_{S \in \mathcal{I}(G)} H(S),$$

where

$$H(S) = \sum_{v \in S} \omega(v)$$

We define the $\mathcal{P}(\mathbb{R}_+)$ to $\mathcal{P}(\mathbb{R}_+)$ mapping :

$$A : F \to \mathcal{L}(Y),$$

where

$$Y = \left(\omega - \sum_{i=1}^{\widehat{N}} X_i \right)^+ ,$$

and $(X_i)_{i \geq 1}$ iid with law F, independent of $(\omega, \widehat{N})$ with law $Q \otimes \widehat{P}$. The next result is a slight generalization of Gamarnik et al. (2006).

Theorem 5.1 (Maximal weight independent set - unique fixed point) *Let $G_n = (V_n, E_n, \omega_n)$ be a sequence of finite networks with vertex set $|V_n| = n$. Assume that $U(G_n)$ converges to $\mathrm{GWN}_*(P, Q)$ with $0 < \int x dP < \infty$ and Q has a density with respect to Lebesgue measure. Assume further that the vertex weights of G_n are uniformly integrable. If $L \in \mathcal{P}(\mathbb{R}_+)$ is the unique fixed point of A^2, then*

$$\lim_{n \to \infty} \frac{I(G_n)}{n} = \mathbb{E} \omega \mathbf{1}_{\omega > \sum_{i=1}^{N} X_i},$$

with $(X_i)_{i \geq 1}$ iid with law L, independent of (ω, N) with law $Q \otimes P$.

The important and very restrictive assumption is that A^2 has a unique fixed point.

5.4.1 Proof of theorem 5.1

Step one : Iterated map analysis. In this paragraph, we prove that for any initial measure $F \in \mathcal{P}(\mathbb{R}_+)$, $A^t(F)$ converges as integer t goes to infinity. As for more usual iterated maps $f^t(x)$ with f from $[0, 1]$ to $[0, 1]$, the use of monotony will play a crucial role.

Lemma 5.2 *The mapping A is continuous (for the topology of weak convergence).*

Proof. The $\mathcal{P}(\mathbb{R}_+)^2$ to $\mathcal{P}(\mathbb{R}_+)$ functions which maps (F, G) to the law of $\max(X, Y)$ and $X + Y$ where (X, Y) has distribution $F \otimes G$ are continuous functions. It follows that for every integer $n \geq 0$, the $\mathcal{P}(\mathbb{R}_+)$ to $\mathcal{P}(\mathbb{R}_+)$ function which maps F to the law of $\sum_{i=1}^{n} X_i$ where (X_i) iid with distribution F, is a continuous function. We then write, for any $m \geq 1$ and any bounded continuous function f,

$$\left| \mathbb{E} f \left(\sum_{i=1}^{\widehat{N}} X_i \right) - \sum_{n=0}^{m} \widehat{P}(n) \mathbb{E} f \left(\sum_{i=1}^{n} X_i \right) \right| \leq \|f\|_\infty \widehat{P}((m, \infty)).$$

For m large enough, the right hand side is arbitrarily small. By composition, it then becomes clear that the mapping A is continuous. $\square$

We define the following partial order relation on $\mathcal{P}(\mathbb{R}_+)$, we write

$$F \leq_{st} G$$

if for all $t \in \mathbb{R}_+$, $F(t,\infty) \leq G(t,\infty)$. This is called stochastic domination. Note that if there exists a coupling (X,Y) of (F,G) such that a.s. $X \leq Y$ then $F \leq_{st} G$. The converse is also true.

Theorem 5.3 (Strassen) *If $F \leq_{st} G$ then there exists a coupling (X,Y) of (F,G) such that $X \leq Y$.*

Proof. Define the pseudo-inverse of F and G as, for $x \in [0,1]$,

$$F^{\leftarrow}(x) = \inf\{t \geq 0 : F(t,\infty) \leq x\} \quad \text{and} \quad G^{\leftarrow}(x) = \inf\{t \geq 0 : G(t,\infty) \leq x\}.$$

If t is a continuity point of the non-increasing function $x \mapsto F(x,\infty)$ and U is uniform on $[0,1]$ then

$$\mathbb{P}(F^{\leftarrow}(U) > t) = \mathbb{P}(U < F(t,\infty)) = F(t,\infty).$$

Since there is at most a countable set of discontinuity points of $x \mapsto F(x,\infty)$, we deduce that $X = F^{\leftarrow}(U)$ and $Y = G^{\leftarrow}(U)$ have distributions F and G respectively. Also by assumption, $F^{\leftarrow}(x) \leq G^{\leftarrow}(x)$, in particular, $X \leq Y$. $\square$

Lemma 5.4 *The map A is non-increasing : if $F \leq_{st} G$ then $A(F) \geq_{st} A(G)$.*

Proof. From Strassen theorem, there exists a coupling (X,Y) of F and G such that $X \leq Y$. Consider an iid sequence $(X_i,Y_i)_{i\geq 1}$ of such couplings so that for all integer i, $X_i \leq Y_i$. Let (ω,N) be independent of $(X_i,Y_i)_i$ with law $Q \otimes P$, then

$$\left(\omega - \sum_{i=1}^N X_i\right)^+ \geq \left(\omega - \sum_{i=1}^N Y_i\right)^+.$$

The left hand side has distribution $A(F)$ while the right hand side has distribution $A(G)$. We have thus found a coupling of $A(F)$ and $A(G)$ that fulfills the conditions of the remark before Strassen Theorem. $\square$

Lemma 5.5 *As integer t goes to infinity, $A^{2t}(\delta_0)$ and $A^{2t}(Q)$ converge.*

Proof. Since $\delta_0 \leq_{st} A(F) \leq_{st} Q$, $\delta_0 \leq_{st} A^2(\delta_0)$. By Lemma 5.4, A^2 is non-decreasing and we get

$$\delta_0 \leq_{st} A^2(\delta_0) \leq_{st} A^4(\delta_0) \leq_{st} \cdots$$

In particular for any $s \geq 0$, $A^{2t}(s,\infty)$ is non-decreasing and converges to say $g_0(s)$. For fixed t, $s \mapsto A^{2t}(s,\infty)$ is non-increasing in s, hence g_0 is also non-increasing. Also from $A(F) \leq_{st} Q$, we deduce that $g_0(s) \leq Q(s,\infty)$ and $\lim_{s\to\infty} g_0(s) = 0$. It follows that for all continuity points s of g_0, $1 - g_0(s)$ is the partition function of some probability measure L_0. From Portemanteau theorem 3.2(v), we deduce that A^{2t} converges weakly to L_0.

The same argument carries over with Q since we have $A^2(Q) \leq_{st} Q$. □

Proposition 5.6 *If $L \in \mathcal{P}(\mathbb{R}_+)$ is the unique fixed point of A^2, then for any $F \in \mathcal{P}(\mathbb{R}_+)$, as integer t goes to infinity, $A^t(F) \Rightarrow L$ and $A(L) = L$.*

Proof. By Lemma 5.5, $A^{2t}(Q)$ and $A^{2t}(\delta_0)$ converge to L_Q and L_0 respectively. By lemma 5.2, $A^2(A^{2t}(Q)) = A^{2t+2}(Q)$ and $A^2(A^{2t}(\delta_0)) = A^{2t+2}(\delta_0)$ converge to $A^2(L_Q) = L_Q$ and $A^2(L_0) = L_0$. We deduce that $L = L_0 = L_Q$. Now for any $F \in \mathcal{P}(\mathbb{R}_+)$, $\delta_0 \leq_{st} A(F) \leq_{st} Q$ and composing by A^{2t} we deduce that $A^{2t+1}(F)$ converges to L. Applying the same argument to $G = A(F)$ we deduce the statements. □

Step two : Independent set on finite trees. Let $G = (V, E, \omega)$ be a finite rooted graph network, with root denoted by $\o$. We define the rooted payoff as

$$X(G) = \max_{S \in \mathcal{I}(G)} H(S) - \max_{S \in \mathcal{I}_*(G)} H(S),$$

where $\mathcal{I}_*(G)$ is the set of independent sets S in $\mathcal{I}(G)$ which do not contain the root. From the definition of $X(G)$, if S^* is a maximal weight independent set in $\mathcal{I}(G)$ (i.e. $H(S_*) = I(G)$) then $X(G) > 0$ implies $\o \in S^*$, while $X(G) = 0$ implies that there exists a maximal weight independent set S^* such that $\o \notin S^*$.

Now, with $\mathbb{N}^f = \cup_{k\geq 0}\mathbb{N}^k$, let $(N_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^f}$ be a collection of integers. We build a forest on $\mathbb{N}^f$ by connecting each vertex $\mathbf{i}$ to its offsprings $(\mathbf{i}, 1), \cdots, (\mathbf{i}, N_{\mathbf{i}})$. We define T the rooted tree on $V \subset \mathbb{N}^f$ with root $\o$ as the connected component of $\o$. The weight on vertex $\mathbf{i}$, $\omega(\mathbf{i})$, is simply denoted by $\omega_{\mathbf{i}}$.

Proposition 5.7 *If T is finite, then*

$$X(T) = \left(\omega_{\o} - \sum_{i=1}^{N_{\o}} X(T_i)\right)^+,$$

where $T_1, \cdots, T_{N_{\o}}$ are the rooted subtrees rooted at $1, \cdots, N_{\o}$.

Proof. Let S^* be such that $H(S^*) = \max_{S\in\mathcal{I}_*(T)} H(S)$. Then $S^* \cap T_i$ is a maximal independent set for T_i : $H(S^* \cap T_i) = I(T_i)$. It follows

$$\max_{S\in\mathcal{I}_*(T)} H(S) = \sum_{i=1}^{N_{\o}} I(T_i) = \sum_{i=1}^{N_{\o}} \max_{S\in\mathcal{I}(T_i)} H(S).$$

Similarly, if S^* is now such that $H(S^*) = \max_{S\in\mathcal{I}(T):\o\in S} H(S)$, then $H(S^*\cap T_i) = \max_{S\in\mathcal{I}_*(T_i)} H(S)$. We get

$$\max_{S\in\mathcal{I}(T):\o\in S} H(S) = \omega_\o + \sum_{i=1}^{N_\o} \max_{S\in\mathcal{I}_*(T_i)} H(S) = \omega_\o - \sum_{i=1}^{N_\o} X(T_i) + \sum_{i=1}^{N_\o} \max_{S\in\mathcal{I}(T_i)} H(S).$$

Finally, we subtract our two last expressions,

$$\max_{S\in\mathcal{I}(T):\o\in S} H(S) - \max_{S\in\mathcal{I}_*(T)} H(S) = \omega_\o - \sum_{i=1}^{N_\o} X(T_i).$$

□

Corollary 5.8 *Assume that T has distribution $\mathrm{GWT}(\widehat{P})$ and that $(\omega_{\mathbf{i}})_{\mathbf{i}\in\mathbb{N}^f}$ are iid with law Q. Let $t \geq 1$ be an integer,*

$$X((T)_t) \stackrel{d}{=} A^t(Q).$$

Proof. The subtrees of the offsprings of the root $\o$, $(T_i)_{i\geq 1}$ are iid $\mathrm{GWT}(\widehat{P})$. Thus we have $X((T)_t) = \left(\omega - \sum_{i=1}^{N_\o} X((T_i)_{t-1})\right)^+$. Now by construction, $X((T)_0) \stackrel{d}{=} Q$. By recursion, we deduce that $X(T_t) \stackrel{d}{=} A^t(Q)$. □

Step three: Independent set with boundary conditions. In order to deal with maximal independent sets of graphs that are not necessarily trees but "locally tree-like", we shall generalize the above argument to trees with arbitrary "boundary conditions". More precisely, for a rooted graph G and $t \geq 1$ integer, we define $\partial(G)_t = (G)_t \backslash (G)_{t-1}$ as the set of vertices at distance exactly t from the root. If $B \in \mathcal{I}(G) \cap \partial(G)_t$ we define

$$X_t(G,B) = \max_{S\in\mathcal{I}(G):S\cap\partial(G)_t = B} H(S) - \max_{S\in\mathcal{I}_*(G):S\cap\partial(G)_t = B} H(S).$$

If $t = 0$, then B is either the root or the empty set, and we set $X_0(G,B) = H(B)$. As in the Step II, we consider a rooted tree T on $V \subset \mathbb{N}^f$ with root $\o$ as the connected component of $\o$. The analog of proposition 5.7 to boundary conditions is the following :

Proposition 5.9 *Let $t \geq 1$ be an integer, T be as above and $B \in \mathcal{I}(T) \cap \partial(T)_t$, then*

$$X_t(T,B) = \left(\omega_\o - \sum_{i=1}^{N_\o} X_{t-1}(T_i, B_i)\right)^+,$$

where $T_1, \cdots, T_{N_\o}$ are the rooted subtrees rooted at $1, \cdots, N_\o$ and $B_i = B \cap T_i \in \mathcal{I}(T_i) \cap \partial(T_i)_{t-1}$.

Proof. The proof of proposition 5.7 obviously applies here also. □

Corollary 5.10 *Let $B \in \mathcal{I}(T) \cap \partial(T)_t$. If t is even then $X_t(T,\emptyset) \leq X_t(T,B) \leq X((T)_t)$. If t is odd then $X_t(T,\emptyset) \geq X_t(T,B) \geq X((T)_t)$. In particular, for any $t \geq 1$ and $B \in \mathcal{I}(T) \cap \partial(T)_{2t}$,*

$$X((T)_{2t-1}) \leq X_{2t}(T,B) \leq X((T)_{2t}).$$

Proof. We note that $X_0(T,\emptyset) = 0 \leq X_0(T,B) \leq X((T)_0) = \omega_\varnothing$. For general integer t, we write $X_t(T,B) = \left(\omega_\varnothing - \sum_{i=1}^{N_\varnothing} X_{t-1}(T_i,B_i)\right)^+$, and the first two statements follow by recursion on t. For the last statement, we notice that $X_t(T,\emptyset) = X((T)_{t-1})$. □

Step four : End of proof of theorem 5.1. Let $(X_i)_{i\geq 1}$ be iid with law L, independent of (ω, N) with law $Q \otimes P$, and

$$\gamma = \mathbb{E}\omega \mathbf{1}_{\omega > \sum_{i=1}^N X_i}.$$

We may define S_n^* as the uniformly sampled maximal weight independent set of G_n, i.e. S_n^* is uniformly sampled on the set of independent sets $S \in \mathcal{I}(G_n)$ such that $H(S) = I(G_n)$. If $\varnothing$ denotes a uniformly chosen root on $[n]$, we have

$$\mathbb{E}I(G_n) = n\mathbb{E}\omega_\varnothing \mathbf{1}_{\omega_\varnothing \in S_n^*}.$$

Fix $\varepsilon > 0$, by proposition 5.6, there exists an integer t such that for all integers $s \geq 2t-1$

$$|\mathbb{E}\omega \mathbf{1}_{\omega > \sum_{i=1}^N X_i} - \gamma| < \varepsilon, \tag{5.4}$$

where $(X_i)_{i\geq 1}$ iid with law $A^s(Q)$, independent of (ω, N_*) with law $Q \otimes P_*$, (uniform integrability in s comes from $\omega \mathbf{1}_{\omega > \sum_{i=1}^N X_i} \leq \omega$).

Since $(G_n, \varnothing)$ converges to $\mathrm{GWN}_*(P,Q)$,

$$\lim_n \mathbb{P}((G_n,\varnothing)_{2t-1} \text{ is a tree}) = 1.$$

Thus, writing for ease of notation G_n instead of $(G_n, \varnothing)$, by uniform integrability,

$$\lim_n |\mathbb{E}\omega_\varnothing \mathbf{1}_{\omega_\varnothing \in S_n^*} - \mathbb{E}\omega_\varnothing \mathbf{1}_{\omega_\varnothing \in S_n^*} \mathbf{1}_{(G_n)_{2t+1} \text{ is a tree}}| = 0. \tag{5.5}$$

Now, if the event $\{(G_n)_{2t+1} \text{ is a tree}\}$ holds, we may write

$$\begin{aligned}
\mathbf{1}_{\omega_\varnothing \in S_n^*} &= \sum_{B \in \mathcal{I}(G_n) \cap \partial(G_n)_{2t}} \mathbf{1}_{\omega_\varnothing \in S_n^*} \mathbf{1}_{S_n^* \cap \partial(G_n)_{2t} = B} \\
&= \sum_{B \in \mathcal{I}(G_n) \cap \partial(G_n)_{2t}} \mathbf{1}_{X_{2t}(G_n,B) > 0} \mathbf{1}_{S_n^* \cap \partial(G_n)_{2t} = B} \\
&\in [\mathbf{1}_{X((G_n)_{2t-1}) > 0}, \mathbf{1}_{X((G_n)_{2t}) > 0}],
\end{aligned}$$

where we have applied corollary 5.10. On the event $\{(G_n)_{2t+1} \text{ is a tree}\}$, we denote by $N_\o$ the degree of the root and by

$$((G_{n,1})_{2t}, \cdots, (G_{n,N_\o})_{2t})$$

the rooted subtrees of depth $2t$ rooted at the adjacent vertices of the root, and similarly for depth $2t-1$. From proposition 5.7, on the event $\{(G_n)_{2t+1} \text{ is a tree}\}$,

$$\omega_\o \mathbf{1}_{\omega_\o > \sum_{i=1}^{N_\o} X((G_{n,i})_{2t})} \le \omega_\o \mathbf{1}_{\omega_\o \in S_n^*} \le \omega_\o \mathbf{1}_{\omega_\o > \sum_{i=1}^{N_\o} X((G_{n,i})_{2t-1})}.$$

Now, we use again the assumption that $(G_n, \o)$ converges to $\mathrm{GWN}_*(P,Q)$. It implies that $(\omega_\o, N_\o)$ has limit law $Q \otimes P$ and, conditioned on $N_\o$, the vector $((G_{n,1})_{2t}, \cdots, (G_{n,N_\o})_{2t})$ converges to independent $\mathrm{GWN}(\widehat{P}, Q)$.

Note also that, since the law Q of $\omega_\o$ has a density, if Y is independent of $\omega_\o$, then $\mathbb{P}(\omega_\o = Y) = 0$. Hence, from Portemanteau theorem 3.2(v) and corollary 5.8, $\omega_\o \mathbf{1}(\omega_\o > \sum_{i=1}^{N_\o} X((G_{n,i})_{2t}))$ converges weakly to $\omega_\o \mathbf{1}(\omega_\o > \sum_{i=1}^{N_\o} X_i)$ where (X_i) are iid with law $A^{2t}(Q)$, independent of $(\omega_\o, N_\o)$ with law $Q \otimes P$. And similarly, $\omega_\o \mathbf{1}(\omega_\o > \sum_{i=1}^{N_\o} X((G_{n,i})_{2t-1}))$ converges weakly to $\omega_\o \mathbf{1}(\omega_\o > \sum_{i=1}^{N_\o} X_i')$ where (X_i') iid with law $A^{2t-1}(Q)$. Finally, by uniform integrability,

$$\mathbb{E}\omega_\o \mathbf{1}_{\omega_\o > \sum_{i=1}^{N_\o} X_i} \le \liminf_n \mathbb{E}\omega_\o \mathbf{1}_{\omega_\o \in S_n^*} \le \limsup_n \mathbb{E}\omega_\o \mathbf{1}_{\omega_\o \in S_n^*} \le \mathbb{E}\omega_\o \mathbf{1}_{\omega_\o > \sum_{i=1}^{N_\o} X_i'},$$

By (5.4) and (5.5), we get

$$\limsup_n |\gamma - \mathbb{E}\omega_\o \mathbf{1}_{\omega_\o \in S_n^*}| \le \varepsilon.$$

The theorem follows. $\square$

Bibliography

D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454–1508, 2007.

D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439): 509–512, 1999.

A. D. Barbour and L. H. Y. Chen, editors. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. Singapore University Press, Singapore, 2005. Lectures from the Meeting on Stein's Method and Applications: a Program in Honor of Charles Stein held at the National University of Singapore, Singapore, July 28–August 31, 2003.

A. D. Barbour and G. K. Eagleson. Poisson approximation for some statistics based on exchangeable trials. *Adv. in Appl. Probab.*, 15(3):585–600, 1983.

I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001. ISSN 1083-6489.

P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

L. H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975.

F. Chung and L. Lu. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006.

R. Durrett. *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2007.

P. Erdős and T. Gallai. Graphs with prescribed degrees of vertices (hungarian). *Mat. Lapok*, 11:264–274, 1960.

P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

D. Gamarnik, T. Nowicki, and G. Swirszcz. Maximum weight independent sets and matchings in sparse random graphs. Exact results using the local weak convergence method. *Random Structures Algorithms*, 28(1):76–106, 2006.

E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

S. Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18 (1-2):205–225, 2009.

S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. Kronecker graphs: an approach to modeling networks. *Journal of Machine Learning Research*, 2010.

T. Lindvall. *Lectures on the coupling method*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1992. A Wiley-Interscience Publication.

M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pages 161–179, 1995.

C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602, Berkeley, Calif., 1972. Univ. California Press.

R. van der Hofstad. *Random Graphs and Complex Networks.* 2012. available on http://www.win.tue.nl/ rhofstad/.

N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.