ICFP M2 - Statistical physics 2 – TD n° 4
## Erdős-Rényi random graphs

Grégory Schehr, Francesco Zamponi

In this TD we shall study the properties of Erdős-Rényi random graphs. They can be defined in two ways (that are equivalent in the thermodynamic limit) :
- each of the $N(N-1)/2$ possible edges is present with probability $p = c/N$.
- M edges are chosen independently, uniformly at random among the possible ones. The thermodynamic limit is taken with $c = 2M/N$.

The plot below represents the average fraction f of vertices in the largest component (i.e. the size of the largest component of the graph divided by N), as a function of c, for different sizes N of the graph, obtained by a simple numerical simulation.
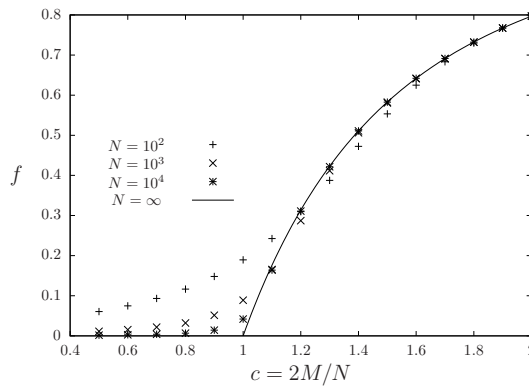
One goal of the TD is to compute the analytic line marked $N = \infty$, that is indeed reached by the numerical results in the thermodynamic limit. It will be easier to use the first definition of the random graph, namely that each edge is present with probability $p = c/N$.

1 Local properties of Erdős-Rényi random graphs

1. Compute the probability q_k that a given vertex i has exactly k neighbours ; give first an exact result at finite N, and simplify it in the thermodynamic limit. What is the average number of neighbours of a given vertex ?

2. Compute the probability that a given vertex i has exactly $k+1$ neighbours, conditionally on the fact that a given edge $\langle i, j \rangle$ adjacent to i is present.

3. Compute the probability that a given vertex has exactly two neighbours, which are also neighbours among themselves. Generalizing this result, argue that the local structure of Erdős-Rényi graphs around a given vertex is tree-like, i.e. with no loops of finite length, with high probability (going to one in the thermodynamic limit).

2 Branching processes

In a Galton-Watson branching process one considers an individual (the ancestor) that has a random number of offsprings (its descendents), who themselves have offsprings, and so on and so forth. One assumes that the numbers of descendents of each individual are i.i.d. random variables, the probability of having k descendents being q_k. As seen in the first part this way of generating a random genealogical tree describes the exploration of the neighborhood of a given vertex in an Erdős-Rényi random graph.

We shall call S the random variable that counts the total number of individuals in a realization of the branching process, and $\mathcal{N}^{(g)}$ the number of individuals in the g-th generation (with $g = 0$ for the ancestor, hence $\mathcal{N}^{(0)} = 1$).

1. Convince yourself that

$$S \overset{\mathrm{d}}{=} 1 + S_1 + \cdots + S_k \ , \qquad \mathcal{N}^{(g+1)} \overset{\mathrm{d}}{=} \mathcal{N}_1^{(g)} + \cdots + \mathcal{N}_k^{(g)} \ , \tag{1}$$

   where in the right hand side k is a random integer drawn with the probability q_k, S_i are independent random variables with the same distribution as S, similarly $\mathcal{N}_i^{(g)}$ are independent copies of $\mathcal{N}^{(g)}$, and we use the convention that an empty sum is equal to zero.

2. Compute the average number of individuals at the g-th generation, $\mathbb{E}[\mathcal{N}^{(g)}]$, and conclude that a phase transition occurs at $c = 1$.

3. We first consider the regime $c < 1$, in which the branching process dies in a finite number of generations with probability one.

   (a) Compute the value of $\mathbb{E}[S]$ as a function of c. Interpret the result in terms of the connected components of an Erdős-Rényi random graph.

   (b) Give the value of the exponent γ_{ER} that describes the divergence of $\mathbb{E}[S]$ when $c \to 1^-$.

4. We consider now the case $c > 1$, and call $f(c)$ the probability that $S = \infty$, i.e. that the branching process persists forever.

   (a) Find a self-consistent equation on $f(c)$.

   (b) Study graphically this equation and draw the shape of $f(c)$.

   (c) Expand the equation when $c \to 1^+$, and determine the exponent β_{ER} that characterizes the behavior of $f(c)$ in this limit.

   (d) Argue that $f(c)$ can be identified with the fraction of vertices in the largest (so-called "giant") component of an Erdős-Rényi random graph.

5. We finally consider the critical case $c = 1$.

   (a) Obtain a self-consistent equation for the generating function $g(x) = \mathbb{E}[x^S]$ for all c, and simplify it in the case $c = 1$.

   (b) Assume that $g(x = 1 - \epsilon) = 1 - a\,\epsilon^\mu + o(\epsilon^\mu)$ when $\epsilon \to 0^+$, with $0 < \mu < 1$. Plug this ansatz in the self-consistent equation and determine the value of a and μ.

   (c) As $\mu < 1$ the generating function is singular around $x = 1$; connect this behavior with the one of $P(s) = \mathbb{P}[S = s]$, that you will assume to be of the form $P(s) \propto 1/s^{\tau-1}$ for $s \to \infty$, and determine the value of the critical exponent $\tau - 1$.

   (d) Let us call $\widehat{P}(s)$ the probability that an uniformly chosen connected component of an Erdős-Rényi random graph contains s vertices. Find a relationship between $P(s)$ and $\widehat{P}(s)$, and conclude that $\widehat{P}(s) \propto 1/s^\tau$ when $s \to \infty$.

   (e) What is the scaling with N of the size of the largest component in a critical Erdős-Rényi random graph?