Phase transitions for the cavity approach to the clique problem on random graphs *

Alexandre Gaudillière[1] Benedetto Scoppola[2] Elisabetta Scoppola[3]
Massimiliano Viale[4]

[1] LATP, Université de Provence, CNRS
39 rue F. Joliot-Curie - 13013 Marseille, France
gaudilli@cmi.univ-mrs.fr

[2] Dipartimento di Matematica, University of Rome "Tor Vergata"
Via della Ricerca Scientifica - 00133 Rome, Italy
scoppola@mat.uniroma2.it

[3] Dipartimento di Matematica, University of Rome "Roma Tre"
Largo San Murialdo, 1 - 00146 Rome, Italy
scoppola@mat.uniroma3.it

[4] Dipartimento di Fisica, University of Rome la Sapienza"
P.le Aldo Moro, 2 00185 Rome, Italy
massimiliano.viale@gmail.com

Abstract

We give a rigorous proof of two phase transitions for a disordered system designed to find large cliques inside Erdös random graphs. Such a system is associated with a conservative probabilistic cellular automaton inspired by the cavity method originally introduced in spin glass theory.

AMS 2010 subject classification: 60C05, 82B26, 82B44.

Key words: Phase transitions, disordered systems, random graphs, cliques, probabilistic cellular automaton.

*Supported by GDRE 224 GREFI-MEFI and the European Research Council through the "Advanced Grant" PTRELSS 228032

Contents

1 Introduction

The largest clique problem (LCP), is the problem to find the largest complete subgraph of a given graph G. Let $G = (V, E)$ be a graph. A graph g is a *subgraph of* G, $g \subset G$, if its vertex set $V(g) \subset V$ and its edges $E(g) \subset E$. A subgraph $g = (V(g), E(g))$ is *complete* if for any $i, j \in V(g)$ then $(i, j) \in E(g)$. We will denote by $\mathcal{K}(G)$ the set of *complete subgraphs* or *cliques* of G and by $\mathcal{M}axCl(G)$ the set of the largest cliques in G:

$$\mathcal{M}axCl(G) := \{g \in \mathcal{K}(G) : |V(g)| = \max_{g' \in \mathcal{K}(G)} |V(g')| \tag{1}$$

where $|B|$ denotes the cardinality of the set B. We call *clique number* of the graph G, $\omega(G)$, the cardinality of the vertex set of any largest clique in G, i.e., $\omega(G) = |V(g)|$ with $g \in \mathcal{M}axCl(G)$. Solving the LCP for a given graph $G(V, E)$ implies finding $\omega(G)$, and both problems are in fact in the same complexity class. Note that we are not strictly following the definition in [1] since we are using the term clique also for a non maximal complete subgraph of G. The LCP is one of the main example of NP-hard problem. It has been proven (see e.g. [GJ] and references therein) to be polinomially equivalent to the k-satisfiability problem and it is equivalent to many other well known difficult problems in combinatorial optimization.

It is well known that the LCP remains difficult also when restricted to typical instances of Erdös random graphs with finite fixed density p, i.e. of graphs with n vertices, $|V| = n$, in which each pair $(i,j) \in V \times V$ belongs to the edges set E with independent probability p. In particular it is well known that in such a random graph G it is very easy to find complete subgraphs $g \in \mathcal{K}(G)$ of size $|g| = -\frac{\log n}{\log p}$ but is difficult to find cliques that exceed this size, see below. The clique number $\omega(G)$ is almost deterministic, in a sense that will be stated more precisely below, and it is roughly speaking twice the size of the cliques that are easy to find. Recently some progress has been made in [8] to understand the intricated landscape of the LCP for Erdös random graph, and therefore to show why at the moment there are no algorithms able to find cliques of a size that exceeds the easy one.

In a previous paper [6] , in collaboration with Antonio Iovanella, two of us introduced an algorithm to find cliques inspired by the cavity method developed in the study of spin glasses. This Markov Chain Monte Carlo exhibits very good numerical performances, in the sense that, although asymptotically it is not able to find cliques larger than the easy ones, for finite size effects it find cliques very near to the largest also for quite large graphs. The idea of the algorithm is the following: starting from a non feasible (i.e. non clique) configuration σ of k vertices of V, the algorithm chooses the next configuration assigning to each new set σ' of k vertices of V a probability proportional to $e^{-\beta[H_0(\sigma,\sigma')+h(k-q(\sigma,\sigma'))]}$, where β is a parameter called *inverse temperature*. The function H_0 is a non negative quantity defined by the number of missing edges between the two configurations, i.e. the number of pairs (i,j) with $i \in \sigma$, $j \in \sigma'$ and $i \neq j$ such that $(i,j) \notin E(G)$. The quantity $q(\sigma, \sigma')$ represents the overlap between σ and σ' and then $k - q(\sigma, \sigma')$ is the number of vertices in σ' that are not in σ. The transition probabilities depend therefore also on the positive parameter h.

The presence of βH_0 in the transition probabilities, when β is large, makes very low the probability to reach new configurations σ' that are badly connected with σ, while a large h depresses the configurations σ' with many different vertices with respect to σ.

From a statistical mechanics point of view the dynamics above has various interesting features. First, the dynamics is conservative, since it is defined on the space of configurations with k vertices, moreover since the whole configuration can be renewed in a single step the resulting MCMC can be considered a *canonical* (or *conservative*) probabilistic cellular automaton (PCA). Rigorous results on canonical PCA are quite rare in the literature. Second, $H(\sigma, \sigma') = H_0 + h(k-q)$ is in some sense the Hamiltonian of a disordered system of pair of configurations, and the combined action of H_0 and $h(k-q)$ makes the energy landscape quite complicate. Third, good numerical performances stimulate a deeper understanding of the dynamics.

For this reasons we decided to study in more detail the statistical mechanical system described by the chain in the case of random graphs. We prove several results. First of all it can be proved rigorously that for suitable values of k, including the interesting case $k = \omega(G)$, the annealed analysis corresponds to the quenched one.

Then it can be proved the existence in the plane (β, h) of a nontrivial phase diagram. More precisely, the system exhibits a first order phase transition while the pair β, h crosses a line $h_c(\beta)$. At $h > h_c(\beta)$ the phase is characterized by pairs of configurations σ, σ' with $\sigma = \sigma'$ and with a given density of missing links in σ, depending on β. At $h < h_c(\beta)$ the phase is characterized by pairs of disjoint configurations σ, σ' with again a particular value for the density of missing links between σ and σ' depending on β.

Moreover, in the region below the critical line $h_c(\beta)$, a second phase transition is present, and again it has a transparent "physical" interpretation: for temperature $T = \frac{1}{\beta}$ below a critical value T_c the system tend to oscillate indefinitely between two fixed configurations σ and σ', while above T the new configuration at each step is typically different from the configurations previously visited by the system.

This detailed control on the features of the system is achieved by a careful evaluation of the thermodynamics. In particular the proof of the existence of the phase transitions can be performed in a relatively easy way, computing the annealed partition function of the system. The self averaging of the system, i.e., the equivalence between quenched and annealed, is more complicate to prove rigorously, involving the computation of the second moment of the partition function, and it is more a brute-force computation. We will present it in some detail in the paper, in an almost pedagogical way, because, as far as we know, there are few cases in the literature where a phase transition for a disordered system can be controlled rigorously. Moreover the way we achieve this result, although based on classical argument like the saddle point method, has some technical details that are quite interesting and may be useful also in different contexts.

Of course this analysis gives important information on the choice of the parameter used in simulation, and in a following paper we will discuss its application to the study of the convergence to equilibrium of the dynamics.

To be more precise we need now some definition.

1.1 Random graphs and the clique number

In this section we fix definitions and notations on random graphs and we recall well known results on the clique number.

For all $p \in [0,1]$ consider the *probability space* given by an infinite sequence of independent Bernoulli variables of parameter p, i.e., $\omega \in \Omega := \{0,1\}^{\mathbb{N}}$, $\omega = (a_1, a_2, ..., a_l, ...)$ with $a_l \in \{0,1\}$, with σ-algebra generated by $A_l^j := \{\omega : \ a_l = j\}$, $j = 0,1$ and with probability measure

$$\mathbb{P}(\omega : \ a_{i_1} = j_1, ...a_{i_k} = j_k) = p_{j_1}...p_{j_k} \quad \text{with } p_1 = p, \ p_0 = 1-p.$$

Given a set of vertices $V = \{1, ..., n\}$ we associate to it the probability space Ω_n given by the first $\binom{n}{2}$ Bernoulli variables in Ω describing the edges between vertices in V, with the obviuos ordering $(1,2), (1,3), (2,3), ..., (1,n), (2,n), ..., (n-1,n)$. In this way we represent with Ω the probability space usually denoted by $\mathcal{G}(\mathbb{N}, p)$, i.e., the infinite random graph.

For any $G \in \mathcal{G}(\mathbb{N}, p)$ and $n \in \mathbb{N}$ we denote by G_n the subgraph of G *spanned* by the set $V_n := \{1, 2, ..., n\}$, i.e., the subgraph of G containing all the edges of G that join two vertices in V_n. By definition G_n is Ω_n measurable. We will denote by $\mathbb{P}$ and $\mathbb{E}$ the probability and the mean value respectively, on this probability space.

The following well known result on the clique number can be found in [1](Corollary 11.2, pg 286):

Proposition 1.1 *For a.e. $G \in \mathcal{G}(\mathbb{N},p)$ there is a constant $m_0(G)$ such that if $n \geq m_0(G)$ then*

$$\left|\omega(G_n) - 2\log_b n + 2\log_b \log_b n - 2\log_b(\frac{e}{2}) - 1\right| < \frac{3}{2}$$

with $b := \frac{1}{p}$.

The main tool in the proof of this Proposition is the study of the random variable $Y_r(n)$ defined as the number of complete subgraphs of G_n with r vertices, i.e., the number of r-cliques in G_n with the second moment method . Indeed its mean value is given by:

$$\mathbb{E}Y_r(n) = \binom{n}{r} p^{\binom{r}{2}} =: f(r,n) \simeq b^{r\log_b n - \frac{r^2}{2}}$$

with $b := \frac{1}{p}$. The function $f(r,n)$, as a function of r, has its maximum in $r \simeq \log_b n$ and drops rather suddenly below 1, by increasing r, say at $r \sim 2\log_b n$. Moreover, again by an explicit calculation, $Y_r(n)$ satisfies the following inequality

$$\frac{varY_r(n)}{(\mathbb{E}Y_r(n))^2} \leq br^4 n^{-2} + 2(\mathbb{E}Y_r)^{-1},$$

when $(1+\eta)\log_b n < r < 3\log_b n$, for $\eta \in (0,1)$. With the Borel-Cantelli lemma, it is easy to show that, given $\varepsilon \in (0, \frac{1}{2})$, for almost every graph $G \in \mathcal{G}(\mathbb{N},p)$ there is a constant $m_0(G)$ such that if $n \geq m_0(G)$ and $n'_r \leq n \leq n_{r+1}$ then $\omega(G_n) = r$, with

$$n_r := \max\{n \in \mathbb{N} :\ f(r,n) \leq r^{-(1+\varepsilon)}\} \quad n'_r := \min\{n \in \mathbb{N} :\ f(r,n) \geq r^{1+\varepsilon}\}. \qquad (2)$$

Indeed the size k of the interesting cliques can be parametrized by a real $c \in (1,2]$ since the relation between n and the size k of the cliques that we want to study is given by

$$\ln n = k\frac{\ln 1/p}{c}, \quad \text{with } c \in (1,2].$$

As emerges in (2) it's more efficient to use k as parameter, instead of n, to study the asymptotic behavior for large graphs and so for any $\bar{c} > 1$ we define

$$\mathcal{S}_{\bar{c}} := \{(n_k)_{k>0} :\ \lim_{k\to\infty} \frac{\ln n_k}{k} = \frac{\ln 1/p}{\bar{c}}\} \qquad (3)$$

This means that if we define $c_k = k\frac{\ln 1/p}{\ln n_k}$ we consider sequences n_k such that $\lim_{k\to\infty} c_k = \bar{c}$. This is actually a particular asymptotic regime that could be generalized.

Let Y be a *random function* on the probability space Ω associating to a pair (n,k) a random variable $Y(n,k)$ on Ω_n depending on k, for instance the number of k cliques in G_n, considered before.

Definition 1.2 *A random function Y on the probability space Ω is called $\bar{c}$-asymptotically self averaging, if the random variables $\frac{Y(n_k,k)}{\mathbb{E}Y(n_k,k)}$ converge almost surely to 1 for any $(n_k)_{k>0}$ in $\mathcal{S}_{\bar{c}}$ as $k\to\infty$, uniformly in $\mathcal{S}_{\bar{c}}$.*

This means that there exists $\tilde{\Omega}\subset\Omega$ with $\mathbb{P}(\tilde{\Omega})=1$ such that for every $\omega\in\tilde{\Omega}$ and any $(n_k)_{k>0}\in\mathcal{S}_{\bar{c}}$ the random variable $\frac{Y(n_k,k)(\omega)}{\mathbb{E}Y(n_k,k)}$ converges to 1.

Note that, by the Borel-Cantelli lemma a sufficient condition for self-averaging is the following:

$$\frac{varY(n_k,k)}{(\mathbb{E}Y(n_k,k))^2}<n_k^{-\alpha}e^{o(k)} \tag{4}$$

for some $\alpha>1$ with $o(k)$ uniform in $(n_k)\in\mathcal{S}_{\bar{c}}$. Indeed for any $\varepsilon>0$ we have that

$$P\Big(|\frac{Y(n,k)}{\mathbb{E}Y(n,k)}-1|>\varepsilon\ \text{ for some } n=n_k,\ (n_k)_{k>0}\in\mathcal{S}_{\bar{c}}\Big)\le$$

$$\le e^{\frac{k\ln 1/p}{c}+o(k)}\frac{1}{\varepsilon^2}var(\frac{Y(n_k,k)}{\mathbb{E}Y(n_k,k)})\le\frac{1}{\varepsilon^2}e^{k\frac{\ln 1/p}{c}(\alpha-1)+o(k)}$$

is summable on k, since $(n_k)_{k>0}\in\mathcal{S}_{\bar{c}}$ implies that $c_k:=k\frac{\ln 1/p}{\ln n_k}$ converges to $\bar{c}$, so that with at most finitely many exception on k, we have that $|\frac{Y(n_k,k)}{\mathbb{E}Y(n_k,k)}-1|<\varepsilon$ for any $(n_k)_{k>0}\in\mathcal{S}_{\bar{c}}$.

We also note that if Z is $\bar{c}$-asymptotically self averaging we have that $\ln Z(n_k,k)-\ln\mathbb{E}Z(n_k,k)$ converges almost surely to 0.

1.2 The cavity algorithm

Let $V=\{1,...,n\}$ and define for each unordered pair in $V\times V$

$$J_{ij}=\begin{cases}0 & \text{if } (i,j)\in E\\ 1 & \text{if } (i,j)\notin E\end{cases} \tag{5}$$

We consider the space $\mathcal{X}^{(n)}:=\{0,1\}^{\{1,\dots,n\}}$ of lattice gas configurations on V and we will denote by the same letter a configuration $\sigma\in\mathcal{X}^{(n)}$ and its support $\sigma\subseteq V$. On this configuration space $\mathcal{X}$ we can consider an Ising Hamiltonian with an antiferromagnetic interaction between non-neighbor sites:

$$H(\sigma) := \sum_{i,j\in V,\, i\neq j} J_{ij}\sigma_i\sigma_j - h\sum_{i\in V}\sigma_i \tag{6}$$

where $h > 0$. It is immediate to prove that when $h < 2$ the minimal value of $H(\sigma)$ is obtained on configurations with support on the vertices of a maximum clique. In the case of a random graph G, i.e., when the interaction variables J_{ij} are i.i.d.r.v., the Hamiltonian (6) is similar to the Hamiltonian of the Sherrington-Kirkpatrick(SK) model. The main differences are that we use lattice gas variables instead of spin variables and, more important, the interaction is given by Bernoulli variables.

For each $\sigma \in \mathcal{X}^{(n)}$ we define its *cavity field* (or *molecular field*) as the field created in each site i by all the sites in the configuration σ:

$$h_i(\sigma) = \sum_{j\neq i} J_{ij}\sigma_j + h(1-\sigma_i) \qquad \forall i \in V. \tag{7}$$

We consider the canonical case, i.e., for any integer $k < n$ we define the *canonical configuration space*

$$\mathcal{X}_k^{(n)} := \{\sigma \in \mathcal{X}^{(n)} : \sum_{i\in V}\sigma_i = k\} \tag{8}$$

For each pairs of configurations $\sigma, \sigma' \in \mathcal{X}_k^{(n)}$ we can define the *pair hamiltonian*:

$$H(\sigma,\sigma') = \sum_{i,j\in V, i\neq j} J_{ij}\sigma_i\sigma'_j + h\sum_i (1-\sigma_i)\sigma'_i = \sum_i h_i(\sigma)\sigma'_i \tag{9}$$

This hamiltonian is non-negative and vanishes when $\sigma = \sigma'$ and its support is a k-clique.

For every $\sigma, \sigma' \in \mathcal{X}_k^{(n)}$ the transition probabilities of the *cavity algorithm* are given by:

$$P(\sigma,\sigma') = \frac{e^{-\beta H(\sigma,\sigma')}}{\sum_{\tau\in\mathcal{X}_k} e^{-\beta H(\sigma,\tau)}} = \frac{e^{-\beta H(\sigma,\sigma')}}{Z_\sigma}, \tag{10}$$

with

$$Z_\sigma = \sum_{\tau\in\mathcal{X}_k} e^{-\beta H(\sigma,\tau)}. \tag{11}$$

By an immediate computation we can check that the detailed balance condition w.r.t. the *invariant measure* on $\mathcal{X}_k^{(n)}$

$$\mu(\sigma) = \frac{\sum_{\tau\in\mathcal{X}_k^{(n)}} e^{-\beta H(\sigma,\tau)}}{\sum_{\tau,\sigma\in\mathcal{X}_k^{(n)}} e^{-\beta H(\sigma,\tau)}} = \frac{Z_\sigma}{Z} \tag{12}$$

is verified with *partition function* Z:

$$Z(n,k) = Z(\mathcal{X}_k^{(n)}) = \sum_{\sigma\in\mathcal{X}_k^{(n)}} Z_\sigma = \sum_{\sigma,\tau\in\mathcal{X}_k^{(n)}} e^{-\beta H(\sigma,\tau)}. \tag{13}$$

We will denote by $\mu(.)$ the mean w.r.t. this stationary measure. For large β, this stationary measure is exponentially concentrated on cliques.

Note that at each step all the sites are updated; this dynamics could be considered a canonical version of probabilistic cellular automata (PCA). Given a fixed configuration σ the probability measure on $\mathcal{X}_k^{(n)}$ given by $\pi_\sigma(.) := P(\sigma,.)$ can be considered in the frame of the *Fermi statistics*. Indeed the cavity fields $h_i(\sigma)$ have values $e_{l,r} = l + rh$ with $l \in \{0,1,...,k\}$ and $r \in \{0,1\}$. We define

$$\mathcal{I}_{l,1} := \{i \in V : h_i(\sigma) = l + h\} \qquad l = 0,...,k. \tag{14}$$

By equation (9) we have

$$H(\sigma,\tau) = \sum_i h_i(\sigma)\tau_i = \sum_{l,r} e_{l,r} \sum_{i\in\mathcal{I}_{l,r}} \tau_i =: \sum_{l,r} e_{l,r} n_{l,r} \tag{15}$$

where $n_{l,r}$ denotes the occupation of the level (or cell) (l,r). On the other hand each level consists of $g_{l,r} := |\mathcal{I}_{l,r}|$ subcells (or sublevels) containing at most one particle since for every $i \in \mathcal{I}_{l,r}$ we have $\tau_i \in \{0,1\}$. This means that instead of configurations in $\mathcal{X}_k^{(n)}$ we can consider the occupation numbers of the levels $\{n_{l,r}\}_{l=0,...,k,\, r=0,1}$. This statistical system is called a *Fermi gas*, see [4] for more detail on sampling for the Fermi statistics, and thus on the realization of this single step of the dynamics.

As far as the energy levels corresponding to sites not in σ, i.e., with $r = 1$, are concerned, we have that their number of sublevels, $g_{l,1} = |\{i \in V : \; h_i(\sigma) = l + h\}|$, is "almost deterministic", as discussed in [6]. Indeed they follow a binomial law, and precise results can be found in Lemma 5.4 in Section 5.

A final remark on probability measures can be useful. The invariant measure $\mu(\sigma)$ is not a Gibbs measure, as usual with PCA, but we can define a Gibbs measure on pairs of configurations, i.e., on $\mathcal{X}_k^{(n)} \times \mathcal{X}_k^{(n)}$ as $\mu_2(\sigma,\sigma') = \frac{1}{Z} e^{-\beta H(\sigma,\sigma')}$ with the same partition function $Z(n,k)$ given in (13). Actually the invariant measure μ can be considered the marginal of μ_2. The probability measure $\pi_\sigma(.) = \frac{e^{-\beta H(\sigma,.)}}{Z_\sigma}$ introduced above in the discussion on the Fermi statistics can be considered as the *conditioned measure* on $\mathcal{X}_k^{(n)}$, since we have the relation:

$$\mu_2(\sigma,\tau) = \mu(\sigma)\pi_\sigma(\tau). \tag{16}$$

1.3 Results

The main results presented in this paper are summarized by the following:

Theorem 1.3 *For each $p \in (0,1)$ and $\beta \in (0,\infty]$*

i) let n and k be integers such that $c := k\frac{\ln 1/p}{\ln n} > 1$, defining $\tilde{h} := \frac{h}{k}$, there is a critical value of $\tilde{h}$ defined by

$$\tilde{h}_c = \frac{1}{\beta}\Big(\frac{f(2\beta)}{2} - f(\beta) + \frac{\ln(1/p)}{c}\Big) \tag{17}$$

with $f(\beta) := -\ln\left[p + (1-p)e^{-\beta}\right]$ for which

$$\ln(\mathbb{E}Z(\mathcal{X}_k^{(n)})) = \begin{cases} k^2\frac{\ln(1/p)}{c} + k - f(2\beta)\frac{k(k-1)}{2} - k\ln k + o(k) & \text{if } \tilde{h} > \tilde{h}_c \\ 2k^2\frac{\ln(1/p)}{c} + 2k - \beta\tilde{h}k^2 - f(\beta)k^2 - 2k\ln k + o(k) & \text{if } \tilde{h} < \tilde{h}_c \end{cases} \tag{18}$$

ii) if $c \in (1,2]$

$$\frac{var Z(\mathcal{X}_k^{(n)})}{(\mathbb{E}Z(\mathcal{X}_k^{(n)}))^2} \le n^{-2}e^{o(k)} \tag{19}$$

with $o(k)$ independent of n and c, so that the partition function Z is $\bar{c}$-asymptotically self averaging for $\bar{c} \in (1,2]$;

iii) the line $\tilde{h}_c$ corresponds to a first order phase transition, in particular the phase with $\tilde{h} > \tilde{h}_c$ is characterized by configurations σ with $H(\sigma,\sigma) \sim k^2 f'(2\beta)$ and the phase with $\tilde{h} < \tilde{h}_c$ is characterized by pairs of disjoint configurations (σ,σ') with $H(\sigma,\sigma') \sim k^2 f'(\beta) + k^2\tilde{h}$;

iv) if $\bar{c} \in (1,2]$ and for any $(n_k)_{k>0} \in \mathcal{S}_{\bar{c}}$ and $\tilde{h} \neq \tilde{h}_{\bar{c}}$ define the entropy $S(n_k,k) := -\sum_{\sigma\in\mathcal{X}_k^{(n_k)}} \mu(\sigma)\ln(\mu(\sigma))$, then $\frac{1}{k^2}S(n_k,k)$ converges almost surely to the following non random function:

$$\bar{s} = \begin{cases} \frac{\ln 1/p}{\bar{c}} - \frac{f(2\beta)}{2} + \beta f'(2\beta) & \text{in the parameter region } (A):\ \tilde{h} > \tilde{h}_{\bar{c}} \\ 2\frac{\ln 1/p}{\bar{c}} - f(\beta) + \beta f'(\beta) & \text{in the region } (B):\ \tilde{h} < \tilde{h}_{\bar{c}} \text{ and } \beta > \beta_{\bar{c}} \\ \frac{\ln 1/p}{\bar{c}} & \text{in the region } (C):\ \tilde{h} < \tilde{h}_{\bar{c}} \text{ and } \beta < \beta_{\bar{c}} \end{cases} \tag{20}$$

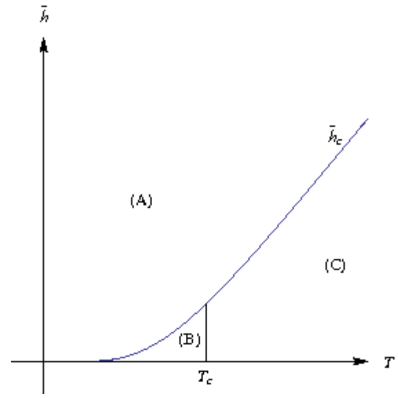

Figure 1: Phase dygram for the cavity algorithm

where $\beta_{\bar{c}}$ is a zero of the function

$$C(\beta) = \frac{\ln 1/p}{\bar{c}} - f(\beta) + \beta f'(\beta) \tag{21}$$

so that also $S(n_k, k)$ is $\bar{c}$-asymptotically self-averaging. The function $\bar{s}$ is discontinuous along the line $\tilde{h} = \tilde{h}_{\bar{c}}$ and has a discontinuity in its first derivative at $T_{\bar{c}} = \frac{1}{\beta_{\bar{c}}}$ corresponding to a "low temperature phase transition". The asymptotic value of the entropy in the phase $\tilde{h} < \tilde{h}_{\bar{c}}$ and $\beta < \beta_{\bar{c}}$ is maximal since $|\mathcal{X}_k^{(n_k)}| \asymp e^{k^2 \frac{\ln 1/p}{c_k}}$.

The phase diagram is summarized in Figure 1.

Remark 1.4 *We can write*

$$Z = \sum_E e^{-\beta E} N_E \tag{22}$$

with E running on all the possible values of the energy $H(\sigma, \tau)$ and N_E being the number of pairs of configurations σ, τ with $H(\sigma, \tau) = E$. For $\beta = \infty$ this implies that $Z = N_0$. Hence the self-averaging of Z implies the self-averaging of the number of cliques of any size k corresponding to $c \in (1, 2]$. This generalizes the Bollobas result quoted above.

Remark 1.5 *Even though the relevant case for the clique problem is $c \in (1, 2]$, for $c > 2$ we can prove (see Appendix C) that, with $\bar{\beta}_c < \infty$ the unique solution of*

$$f(2\bar{\beta}_c) - \frac{1}{2} f(4\bar{\beta}_c) = \frac{1}{c} \ln 1/p,$$

for all $\beta < \bar{\beta}_c$ we still have the estimate (19). Therefore quenched quantities behave like annealed ones for $\beta < \bar{\beta}_c$. In addition we can prove the existence of a second value for the *inverse temperature, say $\hat{\beta}_c > \bar{\beta}_c$ such that for $\beta > \hat{\beta}_c$ and $\beta > 2\hat{\beta}_c$, respectively for $\tilde{h} > \tilde{h}_c$ and $\tilde{h} < \tilde{h}_c$, quenched quantities certainly differ from annealed ones. Indeed if $\hat{\beta}_c > \bar{\beta}_c$ is such that*

$$f(2\hat{\beta}_c) - 2\hat{\beta}_c f'(4\hat{\beta}_c) = \frac{1}{c}\ln 1/p,$$

the estimated entropy for μ_2 obtained from the annealed quantities turns out to be asymptotically negative, i.e., for $(n_k)_{k>0}$ in $\mathcal{S}_{\bar{c}}$,

$$\lim_{k\to+\infty}\frac{1}{k^2}\left(\ln\mathbb{E}[Z(n_k,k)] - \beta\frac{\partial\ln\mathbb{E}[Z(n_k,k)]}{\partial\beta}\right) < 0$$

for $\beta > \hat{\beta}_{\bar{c}}$ in the case $\tilde{h} > \tilde{h}_{\bar{c}}$ and $\beta > 2\hat{\beta}_{\bar{c}}$ in the case $\tilde{h} < \tilde{h}_{\bar{c}}$. Then, for $\beta > \hat{\beta}_{\bar{c}}$ and $\beta > 2\hat{\beta}_{\bar{c}}$ respectively, quenched quantities certainly differ from annealed ones. We actually expect a third phase transition at these temperatures: conversely, for $\beta < \hat{\beta}_{\bar{c}}$ and $\beta < 2\hat{\beta}_{\bar{c}}$ respectively, quenched quantities should behave like quenched ones.

Notation:

For notation convenience in what follows we adopt the following simplification: given $\bar{c} > 1$ and $(n_k)_{k>0} \in \mathcal{S}_{\bar{c}}$ we write $n = n_k$ and $c = c_k$

2 The annealed partition function $\bar{Z}$

In the case of random graphs is not difficult to compute the annealed partition function:

$$\bar{Z} := \mathbb{E}Z = \mathbb{E}\Big[\sum_{\sigma,\tau\in\mathcal{X}_k} e^{-\beta H(\sigma,\tau)}\Big]. \tag{23}$$

Let $I := \sigma\cap\tau$, and q be the overlap $q := |I|$. We denote by $H_0(\sigma,\tau)$ the first part of the pair hamiltonian, i.e., the pair hamiltonian evaluated for $h = 0$:

$$H_0(\sigma,\tau) = \sum_{i,j\in V,\, i\neq j} J_{ij}\sigma_i\sigma'_j = \sum_{\{i,j\}} J_{ij}(\sigma_i\tau_j + \sigma_j\tau_i) \tag{24}$$

The quantity $\sigma_i\tau_j + \sigma_j\tau_i$ takes values $0, 1, 2$ as given in Table 1, where we denote by $S := \sigma\backslash I$, $T := \tau\backslash I$ and $C := (\sigma\cup\tau)^c$. So $\sigma_i\tau_j + \sigma_j\tau_i = 1$ for $\{i,j\}$ in the set of unordered pairs $\mathcal{E}_1 = (S\times I)\cup(T\times I)\cup(S\times T)$ and $\sigma_i\tau_j + \sigma_j\tau_i = 2$ for $\{i,j\} \in \mathcal{E}_2 = I\times I$. By using

	S	I	T	C
S	0	1	1	0
I	1	2	1	0
T	1	1	0	0
C	0	0	0	0

Table 1: Values of $\sigma_i\tau_j + \sigma_j\tau_i$

the independence of the random variables J_{ij} we can conclude:

$$\bar{Z} = \sum_{q=0}^{k} \sum_{\sigma,\tau\in\mathcal{X}_k:\, |I|=q} e^{-\beta h(k-q)} \prod_{\{i,j\}\in\mathcal{E}_1} \mathbb{E}e^{-\beta J_{ij}} \prod_{\{i,j\}\in\mathcal{E}_2} \mathbb{E}e^{-2\beta J_{ij}}.$$

Since $\mathbb{E}e^{-\beta J_{ij}} = e^{-f(\beta)}$ and $\mathbb{E}e^{-2\beta J_{ij}} = e^{-f(2\beta)}$ with $f(\beta) = -\ln[p + (1-p)e^{-\beta})]$ and defining

$$\Theta(q) := \ln\Big[\binom{n}{2k-q}\binom{2k-q}{q}\binom{2(k-q)}{k-q}\Big] \tag{25}$$

we have

$$\bar{Z} = \sum_{q=0}^{k} e^{\Theta(q)} e^{-\beta h(k-q) - f(\beta)(2q(k-q)+(k-q)^2) - f(2\beta)\frac{q(q-1)}{2}} = \sum_{q=0}^{k} e^{\Theta(q)+\Phi(q)} \tag{26}$$

with

$$\Phi(q) := -\beta h(k-q) - f(\beta)(k^2-q^2) - f(2\beta)\frac{q(q-1)}{2} \tag{27}$$

We collect in Appendix A the main properties of the function $f(\beta)$.

As far as the entropic term Θ is concerned we can use the Stirling formula $n! = (\frac{n}{e})^n n^{\frac{1}{2}}\sqrt{2\pi}e^{\mathcal{O}(\frac{1}{12n})}$ to obtain the following asymptotic behaviors for m and k large with $m = o(n)$ and $g < k$:

$$\binom{n}{m} = \exp\{m\ln n - m\ln m + m + o(m)\}, \quad \ln(n-m) = \ln n - \frac{m}{n} + \mathcal{O}(\frac{m^2}{n^2}) \tag{28}$$

$$\binom{k}{g} = \exp\{k\ln k - (k-g)\ln(k-g) - g\ln g + o(k)\}. \tag{29}$$

With the definition $\ln n = k\frac{\ln(1/p)}{c}$ we can write:

$$\Theta(q) = (2k-q)k\frac{\ln(1/p)}{c} + 2k - q - q\ln q - 2(k-q)\ln(k-q) + o(k)$$

We can estimate $\bar{Z}$ for large k with the saddle point method looking for the maximum of the function $\Theta(q) + \Phi(q)$. Indeed $\ln \bar{Z} = \max_{q\in[0,k]}(\Theta(q) + \Phi(q)) + \mathcal{O}(\ln k)$ We have $\Theta(q)+\Phi(q) = a(q)-b(q)$, where $a(q)$ is a polynomial with degree less or equal 2, i.e., with

$$a(q) = (2k-q)k\frac{\ln(1/p)}{c} + 2k - q - \beta h(k-q) - f(\beta)(k^2-q^2) - f(2\beta)\frac{q(q-1)}{2},$$

and $b(q)$ the remaining part:

$$b(q) = q\ln q + 2(k-q)\ln(k-q) + o(k)$$

By noting that $f(\beta)$ is a concave function so that $f(\beta) - \frac{f(2\beta)}{2} > 0$ we have that $a(q)$ is a convex parabola, and so with maximum in 0 or k, while $b(q)$ is non negative and $|b(q)| \leq 4k\ln k$. We have that, for sufficiently large k, the maximum of $a(q)$ is obtained in $q_{max} = k$ if $\tilde{h} > \tilde{h}_c$, (see equation (17) for the definition of $\tilde{h}_c$) and in $q_{max} = 0$ if $\tilde{h} < \tilde{h}_c$. By simple calculations we have that these points k and 0 correspond to the maximum also for the function $\Theta(q)+\Phi(q) = a(q)-b(q)$ when $\tilde{h} > \tilde{h}_c$ and $\tilde{h} < \tilde{h}_c$, respectively. Indeed in the two different cases it is immediate to verify that in a neighborhood of q_{max}, i.e., in the intervals $[k-k^\alpha, k]$ and $[0, k^\alpha]$ with $\alpha \in (0,1)$, respectively, the following estimates hold for the variations of the functions a and b: for sufficiently large k, there exists a positive ε such that

$$|\Delta a(q)| := |a(q+1) - a(q)| > \varepsilon k, \quad |\Delta b(q)| = o(k)$$

and for q outside these intervals $a(q) < a(q_{max}) - \varepsilon k^{1+\alpha}$, so that $a(q) - b(q) \leq a(q) < a(q_{max}) - \varepsilon k^{1+\alpha} < a(q_{max}) - b(q_{max})$.

Summarizing we have, for $\tilde{h} > \tilde{h}_c$:

$$\ln \bar{Z} = k^2\frac{\ln(1/p)}{c} + k - f(2\beta)\frac{k(k-1)}{2} - k\ln k + o(k) \tag{30}$$

and for $\tilde{h} < \tilde{h}_c$

$$\ln \bar{Z} = 2k^2\frac{\ln(1/p)}{c} + 2k - \beta\tilde{h}k^2 - f(\beta)k^2 - 2k\ln k + o(k) \tag{31}$$

For large k we obtain that $\frac{1}{k^2}\ln \bar{Z}$ is a continuous function with a discontinuous derivative in β when $\tilde{h} = \tilde{h}_c$, corresponding to a line of a first order phase transition as discussed in Section 4.

$j\backslash i$	SS’	SI’	ST’	SC’	IS’	II’	IT’	IC’	TS’	TI’	TT’	TC’	CS’	CI’	CT’	CC’
SS’	0	1	1	0	1	2	2	1	1	2	2	1	0	1	1	0
SI’		2	1	0	2	3	2	1	2	3	2	1	1	2	1	0
ST’			0	0	2	2	1	1	2	2	1	1	1	1	0	0
SC’				0	1	1	1	1	1	1	1	1	0	0	0	0
IS’					2	3	3	2	1	2	2	1	0	1	1	0
II’						4	3	2	2	3	2	1	1	2	1	0
IT’							2	2	2	2	1	1	1	1	0	0
IC’								2	1	1	1	1	0	0	0	0
TS’									0	1	1	0	0	1	1	0
TI’										2	1	0	1	2	1	0
TT’											0	0	1	1	0	0
TC’												0	0	0	0	0
CS’													0	1	1	0
CI’														2	1	0
CT’															0	0
CC’																0

Table 2: The value of $\sigma_i\tau_j + \sigma_j\tau_i + \sigma'_i\tau'_j + \sigma'_j\tau'_i$ for different i, j

3 The asymptotic self-averaging of Z

The proof of self averaging of Z is a crude calculation based on elementary arguments. We first evaluate the second moment of Z proving that asymptotically it behaves like $\bar{Z}^2$. An upper bound for $\frac{varZ}{\bar{Z}^2}$ is obtained with a more detailed computation based on the same tools.

3.1 The second moment of Z

We evaluate the second moment of Z:

$$\mathbb{E}(Z^2) = \mathbb{E}\Big(\sum_{\sigma,\tau,\sigma',\tau' \in \mathcal{X}_k^{(n)}} e^{-\beta[H(\sigma,\tau)+H(\sigma',\tau')]}\Big). \tag{32}$$

By defining, as before, $q = |\sigma \cap \tau|$ (and similarly q') we have

$$H(\sigma,\tau) + H(\sigma',\tau') = \sum_{\{i,j\}} J_{ij}(\sigma_i\tau_j + \sigma_j\tau_i + \sigma'_i\tau'_j + \sigma'_j\tau'_i) - h(2k - q - q') \tag{33}$$

The quantity $\sigma_i\tau_j + \sigma_j\tau_i + \sigma'_i\tau'_j + \sigma'_j\tau'_i$ takes values $0, 1, 2, 3, 4$ as in the Table 2, where we use the previous notation, i.e., $I := \sigma \cap \tau$, $S := \sigma \backslash I$, $T := \tau \backslash I$, $C := (\sigma \cup \tau)^c$ and similarly for the sets I', S', T' and C'. We also use the notation SS' for the set $S \cap S'$ and so on. The table is symmetric due to the symmetry in the exchange $i \leftrightarrow j$ so we write only the upper triangle. For every $l \in \{1, 2, 3, 4\}$ again we denote by $\mathcal{E}_l$ the set of unordered pairs $\{i, j\}$ where $\sigma_i\tau_j + \sigma_j\tau_i + \sigma'_i\tau'_j + \sigma'_j\tau'_i = l$. By the table we have: $\mathcal{E}_4 = II' \times II'$,

	S'	I'	T'
S	1	2	3
I	4	5	6
T	7	8	9

Table 3: Index of the intersections

$\mathcal{E}_3 = (SI' \times \tau I') \cup (IS' \times I\tau') \cup (II' \times IT') \cup (II' \times TI')$, and so on. With these notations we can write for the second moment of Z:

$$\mathbb{E}(Z^2) = \sum_{q=0}^{k} \sum_{q'=0}^{k} e^{-\beta h(2k-q-q')} \sum_{\substack{\sigma,\tau:\, |\sigma\cap\tau|=q, \\ \sigma',\tau':\, |\sigma'\cap\tau'|=q'}} \prod_{\{i,j\}\in\mathcal{E}_1} \mathbb{E}e^{-\beta J_{ij}} \times$$

$$\times \prod_{\{i,j\}\in\mathcal{E}_2} \mathbb{E}e^{-2\beta J_{ij}} \prod_{\{i,j\}\in\mathcal{E}_3} \mathbb{E}e^{-3\beta J_{ij}} \prod_{\{i,j\}\in\mathcal{E}_4} \mathbb{E}e^{-4\beta J_{ij}} \tag{34}$$

For shortness we will denote by g_r the cardinality of the intersection of the different subsets involved in this table, where the index $r \in \{1, 2, ..., 9\}$ is fixed in Table 3, e.g $g_1 := |SS'|$. The cardinalities g_r have the following constraints:

$$g_1 + g_2 + g_3 \le k - q, \quad g_4 + g_5 + g_6 \le q, \quad g_7 + g_8 + g_9 \le k - q, \tag{35}$$

$$g_1 + g_4 + g_7 \le k - q', \quad g_2 + g_5 + g_8 \le q', \quad g_3 + g_6 + g_9 \le k - q'. \tag{36}$$

The cardinality of the sets given by intersections with C or C' is obtained by difference:

$$g_S := |SC'| = k - q - (g_1 + g_2 + g_3), \quad g_{S'} := |S'C| = k - q' - (g_1 + g_4 + g_7), \tag{37}$$

$$g_I := |IC'| = q - (g_4 + g_5 + g_6), \quad g_{I'} := |I'C| = q' - (g_2 + g_5 + g_8), \tag{38}$$

$$g_T := |TC'| = k - q - (g_7 + g_8 + g_9), \quad g_{T'} := |T'C| = k - q' - (g_3 + g_6 + g_9), \tag{39}$$

By defining $g = g_1 + g_2 + ... + g_9$ and $M(g_1, ..., g_9, n, q, q')$ the multinomial coefficient

$$M(g_1, ..., g_9, n, q, q') = \frac{n!}{g_1!...g_9!g_S!g_I!g_T!g_{S'}!g_{I'}!g_{T'}!(n - (4k - q - q' - g))!}$$

we can write

$$\mathbb{E}(Z^2) = \sum_{q=0}^{k} \sum_{q'=0}^{k} \sum_{g=0}^{(2k-q)\wedge(2k-q')} \sum_{g_1,...,g_9} M(g_1, ..., g_9, n, q, q') e^{\Phi(q)+\Phi(q')+\Psi(q,q',g,g_1,...,g_9)} \quad (40)$$

where the sum over $g_1, ..., g_9$ satisfies the constraints (35), (36) and $g_1 + g_2 + ... + g_9 = g$ and with Φ defined in (27) and Ψ given by:

$$\Psi = \frac{g_5(g_5-1)}{2}\Big(2f(2\beta) - f(4\beta)\Big) + \frac{1}{2}\sum_{r=1}^{9} g_r C_r$$

where the coefficients C_r are defined as follows:

$$C_1 = \Big(2f(\beta) - f(2\beta)\Big)(g_5 + g_6 + g_8 + g_9) \quad (41)$$

$$C_3 = \Big(2f(\beta) - f(2\beta)\Big)(g_4 + g_5 + g_7 + g_8) \quad (42)$$

$$C_7 = \Big(2f(\beta) - f(2\beta)\Big)(g_2 + g_3 + g_5 + g_6) \quad (43)$$

$$C_9 = \Big(2f(\beta) - f(2\beta)\Big)(g_1 + g_2 + g_4 + g_5) \quad (44)$$

$$C_2 = \Big(2f(\beta) - f(2\beta)\Big)(g_4 + g_6 + g_7 + g_9) + \Big(f(\beta) + f(2\beta) - f(3\beta)\Big)(g_5 + g_8) \quad (45)$$

$$C_4 = \Big(2f(\beta) - f(2\beta)\Big)(g_2 + g_3 + g_8 + g_9) + \Big(f(\beta) + f(2\beta) - f(3\beta)\Big)(g_5 + g_6) \quad (46)$$

$$C_6 = \Big(2f(\beta) - f(2\beta)\Big)(g_1 + g_2 + g_7 + g_8) + \Big(f(\beta) + f(2\beta) - f(3\beta)\Big)(g_4 + g_5) \quad (47)$$

$$C_8 = \Big(2f(\beta) - f(2\beta)\Big)(g_1 + g_3 + g_4 + g_6) + \Big(f(\beta) + f(2\beta) - f(3\beta)\Big)(g_2 + g_5) \quad (48)$$

$$C_5 = \Big(2f(\beta) - f(2\beta)\Big)(g_1 + g_3 + g_7 + g_9) + \Big(f(\beta) + f(2\beta) - f(3\beta)\Big)(g_2 + g_4 + g_6 + g_8) \quad (49)$$

We denote by $\mathcal{P}$ the region of parameters $q, q', g, g_1, ..., g_9$ defined by the constraints $0 \le q \le k$, $0 \le q' \le k$, $g = g_1 + ... + g_9$ and (35) and (36).

Lemma 3.1 *For any $\beta \in (0, \infty)$ and for any $q, q', g_1, ..., g_9, g \in \mathcal{P}$ and for $c \le 2$ we have*

$$\Psi(q, q', g_1, ..., g_9, g) \le \bar{\Psi}(q, q', g, g_5) \quad (50)$$

where

$$\bar{\Psi} = \frac{g_5(g_5-1)}{2}\Big(2f(2\beta) - f(4\beta)\Big) + \frac{1}{2}\Big(f(\beta) + f(2\beta) - f(3\beta)\Big)((k \wedge g) + g_5)(g - g_5) \quad (51)$$

The proof of Lemma 3.1 is given in Appendix C.

As far as the entropic term is concerned we can write

$$\sum_{g_1,g_2,g_3,g_4,g_6,g_7,g_8,g_9} \frac{n!}{g_1!...g_9!g_S!g_I!g_T!g_{S'}!g_{I'}!g_{T'}!(n-(4k-q-q'-g))!} \leq$$

$$\leq \exp\{(4k-q-q'-g)\ln n - \ln\Big((q-g)!(q'-g)!\Big) - 2\ln\Big((k-q-g)!(k-q'-g)!\Big) + 8\} =: e^{\bar{\Theta}_2} \tag{52}$$

where the sum is under the condition $g_1 + g_2 + g_3 + g_4 + g_6 + g_7 + g_8 + g_9 = g - g_5$ and with the notation $m! = 1$ if $m \leq 1$ and where, for the sum of g_r with $r \neq 5$, we used the estimate

$$\sum_{g_1,g_2,g_3,g_4,g_6,g_7,g_8,g_9} \frac{1}{g_1!g_2!g_3!g_4!g_6!g_7!g_8!g_9!} = \frac{8^{g-g_5}}{(g-g_5)!} \leq e^8.$$

With these estimates we can write

$$\mathbb{E}(Z^2) \leq \sum_{q=0}^{k}\sum_{q'=0}^{k}\sum_{g=0}^{(2k-q)\wedge(2k-q')}\sum_{g_5=0}^{g\wedge q\wedge q'} e^{\bar{\Theta}_2(q,q',g,g_5)+\Phi(q)+\Phi(q')+\bar{\Psi}(q,q',g,g_5)} \tag{53}$$

To evaluate this sums again we look for the maximum of the exponent. Define for notation convenience $\Phi_2(q,q') = \Phi(q) + \Phi(q')$.

Lemma 3.2 *The maximum of the function $\bar{\Theta}_2 + \Phi_2 + \bar{\Psi}$ on the parameter region defined by the constraints is obtained for $q = q'$ i.e., for $q = q', g, g_5$ in the three dimensional polyhedron $\bar{\mathcal{P}}$ defined by*

$$0 \leq q \leq k, \quad g_5 \leq g \leq 2k - q, \quad 0 \leq g_5 \leq q \tag{54}$$

and represented in Figure 2. Moreover $(\bar{\Theta}_2 + \Phi_2 + \bar{\Psi})(q,q,g,g_5)$ reaches its maximum on $\bar{\mathcal{P}}$ in $(k,k,0,0)$ if $\tilde{h} \geq \tilde{h}_c$ and in $(0,0,0,0)$ if $\tilde{h} < \tilde{h}_c$.

The proof of Lemma 3.2 is given again in Appendix C. Note that when $g = 0$ and $q = q'$ we have the expected relations $\Theta_2 = 2\Theta + \mathcal{O}(\frac{k}{n})$, and $\Psi = 0$.

With these lemmas we immediately obtain

$$\ln \mathbb{E}(Z^2) = \max_{q,g,g_5\in\bar{\mathcal{P}}} (\bar{\Theta}_2 + \Phi_2 + \bar{\Psi})(q,q,g,g_5) + \mathcal{O}(\ln k) = 2\ln \mathbb{E}Z + \mathcal{O}(\ln k) \tag{55}$$

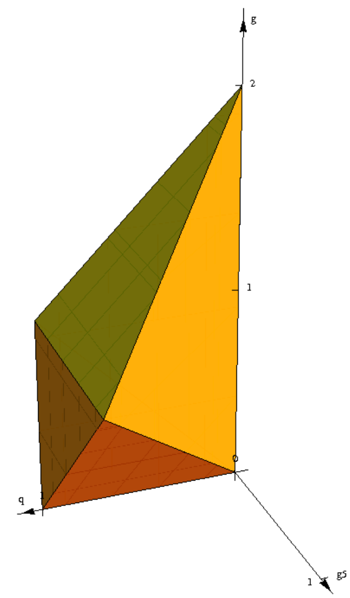

Figure 2: The polyhedron $\bar{\mathcal{P}}$

3.2 Self averaging

To evaluate the quantity $\frac{varZ}{(\mathbb{E}Z)^2}$ we can write

$$(\mathbb{E}Z)^2 = \sum_{q=0}^{k}\sum_{q'=0}^{k}\sum_{g=0}^{(2k-q)\wedge(2k-q')}\sum_{g_1,\dots,g_9}\frac{n!\; e^{\Phi(q)+\Phi(q')}}{g_1!\dots g_9! g_S! g_I! g_T! g_{S'}! g_{I'}! g_{T'}!(n-(4k-q-q'-g))!}$$

and note that, as in the case of the clique number, the terms corresponding to $g=0$ and $g=1$ are identical in $\mathbb{E}(Z^2)$ and in $(\mathbb{E}Z)^2$, indeed $\Psi = 0$ not only for $g=0$ but also in the case $g=1$. Therefore

$$\frac{varZ}{(\mathbb{E}Z)^2} \leq \frac{1}{\bar{Z}^2}\sum_{q=0}^{k}\sum_{q'=0}^{k}\sum_{g=2}^{(2k-q)\wedge(2k-q')}\sum_{g_5=0}^{g\wedge q\wedge q'} e^{\bar{\Theta}_2+\Phi_2+\bar{\Psi}} \tag{56}$$

Lemma 3.3 *The maximum of the function* $(\bar{\Theta}_2+\Phi_2+\bar{\Psi})(q,q',g,g_5)$ *on the parameter region* $\bar{\mathcal{P}}$ *with the additional constraint* $g \geq 2$ *is equal to*

$$\begin{cases} -f(2\beta)k(k-1)+(2k-2)\ln n - 2\ln(k-2)! + o(k) & \textit{if } \tilde{h} > \tilde{h}_c \\ -2\beta hk - 2f(\beta)k^2 + (4k-2)\ln n - 4\ln(k-2)! + o(k) & \textit{if } \tilde{h} < \tilde{h}_c \end{cases}$$

The proof of this Lemma is analogous to that of Lemma 3.2 given in the Appendix C.

With this Lemma we conclude the self averaging result. Consider first the case $\tilde{h} \geq \tilde{h}_c$:

$$\frac{varZ}{(\mathbb{E}Z)^2} \leq \exp\Big\{-2\Big[k\ln n + k - f(2\beta)\frac{k(k-1)}{2} - k\ln k + o(k)\Big] +$$

$$-f(2\beta)k(k-1) + (2k-2)\ln n - 2\ln(k-2)! + o(k)\Big\} \leq$$

$$\leq \exp\Big\{-2k + 2k\ln k - 2\ln n - 2(k-2)\ln(\frac{k-2}{e}) + o(k)\Big\} =$$

$$= \exp\Big\{-2\ln n + o(k)\Big\} = e^{-2\ln n + o(k)}$$

and using the asymptotic $\ln\frac{k}{k-2} \sim \frac{2}{k}$ we obtain the self averaging in this case. In the case $\tilde{h} < \tilde{h}_c$ the calculation is similar:

$$\frac{varZ}{(\mathbb{E}Z)^2} \leq \exp\Big\{-2\Big[2k\ln n + 2k - \beta\tilde{h}k^2 - f(\beta)k^2 - 2k\ln k + o(k)\Big]$$

$$-2\beta hk - 2f(\beta)k^2 + (4k-2)\ln n - 4\ln(k-2)! + o(k)\Big\}$$

$$\leq \exp\Big\{-4k + 4k\ln k - 2\ln n - 4(k-2)\ln(\frac{k-2}{e}) + o(k)\Big\} \leq e^{-2\ln n + o(k)}$$

4 Phase transition across $\tilde{h}_c(T)$

By the previous results on the self averaging of Z with

$$\bar{Z} = \begin{cases} \exp\left\{k^2\left[\frac{\ln 1/p}{c} - \frac{f(2\beta)}{2}\right] + o(k^2)\right\} & \text{if } \tilde{h} > \tilde{h}_c \\ \exp\left\{k^2\left[2\frac{\ln 1/p}{c} - f(\beta) - \beta\tilde{h}\right] + o(k^2)\right\} & \text{if } \tilde{h} < \tilde{h}_c \end{cases} \tag{57}$$

we can conclude that the line $\tilde{h}_c(T)$ represented in Figure 1 corresponds to a line of a first order phase transition. Indeed the function $\ln\bar{Z}$ turns out to be continuous with a discontinuous derivative in β when $\tilde{h} = \tilde{h}_c$ and we have $-\frac{\partial}{\partial\beta}\ln Z = \mu_2(H(\sigma,\tau))$ converges almost surely to

$$-\frac{\partial}{\partial\beta}\ln\bar{Z}\begin{cases} k^2 f'(2\beta) + o(k^2) & \text{if } \tilde{h} > \tilde{h}_c \\ k^2(f'(\beta) + \tilde{h}) + o(k^2) & \text{if } \tilde{h} < \tilde{h}_c \end{cases} \tag{58}$$

By the convexity property of the function $\ln Z$ we can conclude with standard arguments that $\frac{\partial}{\partial\beta}\lim_{k\to\infty}\ln Z = \lim_{k\to\infty}\frac{\partial}{\partial\beta}\ln Z$ and so the same result can be obtained by evaluating directly the mean $\mathbb{E}\Big[\mu_2(H(\sigma,\tau))\Big]$

If we look at the model on the state space of couple of configurations, with Gibbs measure $\mu_2(\sigma,\tau) = \frac{1}{Z} e^{-\beta H(\sigma,\tau)}$, the two phases correspond to two different mean energies.

As far as the second derivative is concerned we have

$$\frac{\partial^2}{\partial\beta^2} \ln Z = var_{\mu_2}(H) = \begin{cases} -k^2 f"(2\beta) + o(k^2) & \text{if } \tilde{h} > \tilde{h}_c \\ -k^2 f"(\beta) + o(k^2) & \text{if } \tilde{h} < \tilde{h}_c \end{cases} \tag{59}$$

and again the same result can be obtained by evaluating directly the mean on the J_{ij} of the variance w.r.t. the pair measure μ_2.

In a similar way we can study the overlap $q(\sigma,\tau)$ by computing $\frac{\partial}{\partial\tilde{h}} \ln Z$. Indeed $\mu_2(q(\sigma,\tau)) = k + \frac{1}{\beta k}\frac{\partial}{\partial\tilde{h}} \ln Z$; we obtain

$$-\frac{1}{\beta k}\frac{\partial}{\partial\tilde{h}} \ln \bar{Z} \begin{cases} 0 & \text{if } \tilde{h} > \tilde{h}_c \\ k & \text{if } \tilde{h} < \tilde{h}_c \end{cases} \tag{60}$$

so that the two phases have not only different mean energies but also different mean overlap.

5 A low temperature phase transition

We prove in this section the last claim of our main theorem. The proof is divided in three steps. First, we made a few prelimiray remarks on the computation of the annealed partition function $\bar{Z}$ and we deduce an almost sure concentration property of the Gibbs measure μ_2. Second, we translate this concentration property in a concentration property of the marginal law μ. Last, we evaluate the free entropy of the measure π_σ for the typical configurations σ by proving a last large deviation estimate.

5.1 Concentration of the Gibbs measure μ_2

An alternative way to compute the annealed partition function consists in counting the mean number $\mathcal{N}(q,l_1,l_2)$ of pairs of configurations (σ,τ) with a given overlap $q = |I| := |\sigma\cap\tau|$, a given number l_1 of missing links inside I, and a given number l_2 of missing links between I and $T := \tau\setminus\sigma$, between $S := \sigma\setminus\tau$ and T, as well as S and I. We get

$$\bar{Z} = \sum_{q=0}^{k} \sum_{l_1=0}^{\frac{q(q-1)}{2}} \sum_{l_2=0}^{k^2-q^2} e^{-\beta[2l_1+l_2+h(k-q)]}\mathcal{N}(q,l_1,l_2). \tag{61}$$

To evaluate $\mathcal{N}(q, l_1, l_2)$ we use the following argument. Consider the obvious extension of the definition (24) of $H_0(\sigma, \tau)$ to a generic pair A, B of subsets of V:

$$H_0(A, B) = \sum_{i,j \in V,\, i \neq j} J_{i,j} \mathbf{1}_A(i) \mathbf{1}_B(j).$$

For $a \in \{0, ..., k\}$ and $l \in \{0, ..., ka\}$ let $\mathcal{A} := \{A \subset \sigma^c : \ |A| = a, \text{ and } H_0(A, \sigma) = l\}$, then

$$\mathbb{E}|\mathcal{A}| = \binom{n-k}{a} \binom{ak}{l} (1-p)^l p^{ak-l} = e^{ak[\frac{\ln 1/p}{c} - I_p(\frac{l}{ak})] + o(k^2)} \tag{62}$$

where we denote by I_p the large deviation functional

$$I_p : x \in [0, 1] \mapsto x \ln \frac{x}{1-p} + (1-x) \ln \frac{1-x}{p}. \tag{63}$$

In Appendix A the main properties of this function are recalled; we just mention here that the function $I_p(x)$ is related to the function $f(\beta)$ used in section 2 by a Legendre transform, indeed we are doing the same computation of $\bar{Z}$ by using different variables. A similar computation can be given for subsets of σ and so we can easily conclude that

$$\mathcal{N}(q, l_1, l_2) = n_I(q, l_1) n_{S,T}(q, l_2)$$

with

$$n_I(q, l_1) = \binom{n}{q} \exp\{-\frac{q(q-1)}{2} I_p(\frac{2l_1}{q(q-1)})\} \tag{64}$$

$$n_{S,T}(q, l_2) = \binom{n-q}{2(k-q)} \binom{2(k-q)}{k-q} \exp\{-(k^2 - q^2) I_p(\frac{l_2}{k^2 - q^2})\} \tag{65}$$

and so

$$\mathcal{N}(q, l_1, l_2) = e^{\Theta(q) - \frac{q(q-1)}{2} I_p(x_1) - (k^2 - q^2) I_p(x_2) + o(k)}. \tag{66}$$

with $\Theta(q)$ defined in (25), $x_1 = \frac{2l_1}{q(q-1)}$ and $x_2 = \frac{l_2}{k^2 - q^2}$, The sums over l_1 and l_2 can be written as sums over x_1 and x_2 and can be estimated with the saddle point method. We then obtain for $\bar{Z}$ the expression given in (30) and (31) by using the fact that f is the Legendre transform of $-I$. In the previous sections we used a different approach to estimate Z and $\bar{Z}$ because this decomposition becomes not easily dealt with as soon as second moment estimates are involved. However, the decomposition proposed here is useful to prove a concentration property for the Gibbs measure μ_2.

Three simple remarks will be used in what follows.

Remark 5.1 *As far as equation (62) is concerned, we note that the set of values of density of missing links having positive entropy is given by*

$$X_c := \{x \in [0,1] : \frac{\ln 1/p}{c} > I_p(x)\}. \tag{67}$$

Remark 5.2 *Note that by the definition of cavity field (7) we have*

$$\sum_{i \in T} h_i(S) = \sum_{i \in S} h_i(T) = H_0(S,T)$$

so that $H_0(S,T) = \frac{\sum_{i\in T} h_i(S) + \sum_{i \in S} h_i(T)}{2}$

Remark 5.3 *The the density of missing links is typically constant in subsets of a given set. More precisely let* A *and* B *be a pair of subsets of* V *with* $|A| = ak$, $|B| = bk$ *with* $a, b > 0$ *and let* $\rho \in (0,1)$. *Then for every* $A_1 \subset A$ *with* $|A_1| = a_1 k$, $a_1 \in (0,a)$ *and for every* $\rho_1 \neq \rho$ *there exists* $\delta(\rho_1) > 0$ *such that*

$$\mathbb{P}(H_0(A_1,B) = \rho_1 a_1 bk^2 | H_0(A,B) = \rho abk^2) < e^{-\delta(\rho_1)abk^2}$$

The proof is an immediate consequence of the convexity of I_p *if we note that*

$$\mathbb{P}(H_0(A_1,B) = \rho_1 a_1 bk^2 | H_0(A,B) = \rho abk^2) =$$

$$= \mathbb{P}(H_0(A_1,B) = \rho_1 a_1 bk^2,\ H_0(A_2,B) = \rho_2(a-a_1)bk^2) e^{abk^2 I_p(\rho) + o(k^2)} \leq$$

$$\leq e^{-a_1 bk^2 I_p(\rho_1) - (a-a_1)bk^2 I_p(\rho_2) + abk^2 I_p(\rho) + o(k^2)}$$

with $A_2 := A \backslash A_1$ *and* $\rho_2 = \rho \frac{a}{a-a_1} - \rho_1 \frac{a_1}{a-a_1}$.

For (σ, τ) in $\mathcal{X}_k^{(n)} \times \mathcal{X}_k^{(n)}$, let us denote again the overlap by $q(\sigma,\tau)$ and by $H_0(\sigma,\tau)$ the number of missing links between σ and τ. We define also

$$\bar{Q}(\sigma,\tau) := \{i \in T :\ h_i(S) \notin X_c\} \cup \{i \in S :\ h_i(T) \notin X_c\} \tag{68}$$

that is the set of points in $S \cup T$ with non-typical number of missing links to the other set, and let $\bar{q}(\sigma,\tau) := |\bar{Q}(\sigma,\tau)|$. Even if the energy $H(\sigma,\tau)$ does not depend on this parameter

$\bar{q}(\sigma,\tau)$, as we will show in Section 5.3, the value of $\bar{q}(\sigma,\tau)$ is crucial to perform entropy estimates at low temperature when $\tilde{h} < \tilde{h}_c$. We then set, for any $\delta > 0$,

$$\mathcal{Q}_\delta := \begin{cases} [1-\delta, 1] & \text{if } \tilde{h} > \tilde{h}_{\bar{c}}, \\ [0,\delta] & \text{if } \tilde{h} < \tilde{h}_{\bar{c}}. \end{cases} \tag{69}$$

$$\mathcal{H}_\delta := \begin{cases} [f'(2\beta)-\delta, f'(2\beta)+\delta] & \text{if } \tilde{h} > \tilde{h}_{\bar{c}}, \\ [f'(\beta)-\delta, f'(\beta)+\delta] & \text{if } \tilde{h} < \tilde{h}_{\bar{c}}. \end{cases} \tag{70}$$

and

$$\bar{\mathcal{Q}}_\delta := \begin{cases} [0,1] & \text{if} \quad \tilde{h} > \tilde{h}_{\bar{c}} \quad \text{or} \quad \tilde{h} < \tilde{h}_{\bar{c}} \text{ and } T > T_c, \\ [1-2\delta, 1] & \text{if} \quad \tilde{h} < \tilde{h}_{\bar{c}} \text{ and } T < T_c. \end{cases} \tag{71}$$

With these intervals of parameters we can define a set of typical pairs of configuration:

$$\Sigma_{2,\delta} := \left\{ (\sigma,\tau) \in \mathcal{X}_k^{(n)} \times \mathcal{X}_k^{(n)} : \ \frac{q(\sigma,\tau)}{k} \in \mathcal{Q}_\delta, \quad \frac{\bar{q}(\sigma,\tau)}{2(k-q(\sigma,\tau))} \in \bar{\mathcal{Q}}_\delta, \quad \frac{1}{k^2} H_0(\sigma,\tau) \in \mathcal{H}_\delta \right\} \tag{72}$$

where we define $\frac{\bar{q}(\sigma,\tau)}{2(k-q(\sigma,\tau))} = 0$ when $q(\sigma,\tau) = k$. By the self-averaging property of Z we have

$$\mathbb{E}\left[\mu_2(\Sigma_{2,\delta}^c)\right] = \mathbb{E}\left[\frac{1}{Z} \sum_{(\sigma,\tau)\in\Sigma_{2,\delta}^c} e^{-\beta H(\sigma,\tau)}\right] = \frac{1}{\bar{Z} e^{o(k)}} \mathbb{E}\left[\sum_{(\sigma,\tau)\in\Sigma_{2,\delta}^c} e^{-\beta H(\sigma,\tau)}\right]. \tag{73}$$

Now, using the previous decomposition (61) and the fact that I_p is strictly convex and more precisely that $I_p'' \geq 2$, we get, with the saddle point method, and using the remarks 5.1, 5.2, 5.3 that

$$\mathbb{E}\left[\mu_2(\Sigma_{2,\delta}^c)\right] \leq e^{-C\delta^2 k^2} \tag{74}$$

for k large enough and a suitable constant C. Indeed

$$\mathbb{E}\left[\mu_2(\Sigma_{2,\delta}^c)\right] \leq \frac{1}{\bar{Z} e^{o(k)}} \sum_{q=0,\ldots,k}{}' \sum_{x_1\in[0,1]} \sum_{x_2\in[0,1]} e^{-\beta[2l_1+l_2+h(k-q)]} \mathcal{N}(q,l_1,l_2) +$$

$$+ \frac{1}{\bar{Z} e^{o(k)}} \sum_{q=0,\ldots,k} \sum_{x_1\in[0,1]}{}' \sum_{x_2\in[0,1]}{}' e^{-\beta[2l_1+l_2+h(k-q)]} \mathcal{N}(q,l_1,l_2) + e^{-2C\delta^2 k^2} \tag{75}$$

where the sums $\sum'$ are with the restriction given by $\sigma \in (\Sigma_{2,\delta})^c$, i.e., $\sum_{q=0,\ldots,k}{}'$ is with the condition $q \notin \mathcal{Q}_\delta$ and $\sum_{x_i}'$ are with the condition x_1 is such that $I_p(x_1) + 2\beta x_1 > f(2\beta) + C(\delta)$ in the case $\tilde{h} > \tilde{h}_c$ and in the case $\tilde{h} < \tilde{h}_c$ with the condition x_2 is such that $I_p(x_2) + \beta x_2 > f(\beta) + C(\delta)$, with $C(\delta) \geq \frac{\delta^2}{2} \min_{x \in [0,1]} I"_p(x) \geq \frac{\delta^2}{2} 2$. Moreover the last term $e^{-2C\delta^2 k^2}$ estimates the mean of the measure of the pairs (σ, τ) such that $\frac{q(\sigma,\tau)}{k} \in \mathcal{Q}_\delta$, $\frac{1}{k^2} H_0(\sigma, \tau) \in \mathcal{H}_\delta$ but $\frac{\bar{q}(\sigma,\tau)}{2(k - q(\sigma,\tau))} \notin \bar{\mathcal{Q}}_\delta$. This can be obtained only in the case $\tilde{h} < \tilde{h}_{\bar{c}}$ and $T < T_c$ and in this regime we have that $f'(\beta) \notin X_c$ and the main contribution to Z is given by pairs of disjoint sets S and T with $|H_0(S,T) - f'(\beta)| < B\delta$ for a suitable constant B. By remarks 5.1 and 5.2 we can conclude that S (and T) can be decomposed into two disjoint parts $S = S_1 \cup S_2$, with $|S_i| \geq \delta k$ with different density of missing link to T. The estimate then follows by remark 5.3.

We conclude, with Markov inequality and Borel-Cantelli lemma, that, almost surely,

$$\mu_2(\Sigma^c_{2,\delta}) \leq e^{-C\delta^2 k^2/2} \tag{76}$$

for k large enough.

5.2 Concentration of the marginal measure μ

Starting from $\Sigma_{2,\delta}$ we want to define a set Σ_δ of *typical configurations* in $\mathcal{X}_k^{(n)}$ with the property that $\sigma \in \Sigma_\delta$ implies that π_σ is concentrated on the configurations τ such that $(\sigma, \tau) \in \Sigma_{2,\delta}$.

To give a precise definition of this set Σ_δ we can proceed as follows. For all σ in $\mathcal{X}_k^{(n)}$ and α , $\bar{\alpha}$ and ρ in $[0,1]$, define

$$s_\sigma(\alpha, \bar{\alpha}, \rho) := \frac{1}{k^2} \ln \left| \left\{ \tau \in \mathcal{X}_k^{(n)} : \frac{q(\sigma,\tau)}{k} = \alpha, \frac{\bar{q}(\sigma,\tau)}{2(k-\alpha)} = \bar{\alpha}, \frac{H_0(\sigma,\tau)}{k^2} = \rho \right\} \right| \tag{77}$$

$$\phi_\sigma(\alpha, \bar{\alpha}, \rho) := s_\sigma(\alpha, \bar{\alpha}, \rho) - \beta(\rho + \tilde{h}(1-\alpha)) \tag{78}$$

$$\phi^*_\sigma := \max\{\phi_\sigma(\alpha, \bar{\alpha}, \rho) : \alpha, \bar{\alpha}, \rho \in [0,1]\} \tag{79}$$

Since the number of possible values of $\alpha, \bar{\alpha}$ and ρ for which $s_\sigma(\alpha, \bar{\alpha}, \rho)$ can be non-negative and finite is only polynomial in k, we note that, for all positive δ and large enough k, the quantity $\ln Z_\sigma$ is such that

$$k^2 \phi^*_\sigma \leq \ln Z_\sigma \leq k^2 \phi^*_\sigma + \delta k^2. \tag{80}$$

We then set, for any $\delta > 0$,

$$\Sigma_\delta := \Big\{\sigma \in \mathcal{X}_k^{(n)} : \text{ there exist } \alpha \in \mathcal{Q}_\delta,\ \bar{\alpha} \in \bar{\mathcal{Q}}_\delta,\ \rho \in \mathcal{H}_\delta \text{ such that } \phi_\sigma(\alpha, \bar{\alpha}, \rho) \geq \phi_\sigma^* - \delta\Big\} \tag{81}$$

Note that for a given positive δ, for k large enough and for all σ in Σ_δ^c,

$$\pi_\sigma\left(\{\tau \in \mathcal{X}_k^{(n)} : \frac{q(\sigma,\tau)}{k} \in \mathcal{Q}_\delta, \frac{\bar{q}(\sigma,\tau)}{2(k - q(\sigma,\tau))} \in \bar{\mathcal{Q}}_\delta, \frac{H_0(\sigma,\tau)}{k^2} \in \mathcal{H}_\delta\}\right) \leq e^{-\delta k^2/2} \tag{82}$$

this means that we have a concentration property of π_σ implying that for configurations $\sigma \notin \Sigma_\delta$ the measure π_σ is concentrated on values of $(\alpha, \bar{\alpha}, \rho)$ not in $\mathcal{Q}_\delta \times \bar{\mathcal{Q}}_\delta \times \mathcal{H}_\delta$. Now, due to (16) we have, for a given $\delta > 0$ and k large enough,

$$\mu_2(\Sigma_{2,\delta}^c) = \sum_{\sigma \in \Sigma_\delta} \mu(\sigma) \sum_{\tau:(\sigma,\tau)\in\Sigma_{2,\delta}^c} \pi_\sigma(\tau) + \sum_{\sigma \in \Sigma_\delta^c} \mu(\sigma) \sum_{\tau:(\sigma,\tau)\in\Sigma_{2,\delta}^c} \pi_\sigma(\tau) \geq \mu(\Sigma_\delta^c)(1 - e^{-\delta k^2/2}). \tag{83}$$

We conclude, using the concentration property of μ, that, almost surely,

$$\mu(\Sigma_\delta^c) \leq e^{-C\delta^2 k^2/3} \tag{84}$$

for k large enough.

5.3 Conclusion

To estimate the entropy, up to $o(k^2)$

$$S := -\sum_{\sigma \in \mathcal{X}_k} \mu(\sigma) \ln(\mu(\sigma)) = \ln Z - \sum_{\sigma \in \mathcal{X}_k} \frac{Z_\sigma}{Z} \ln Z_\sigma = \ln Z + \beta\mu(F), \tag{85}$$

where, for any σ in $\mathcal{X}_k$, $F(\sigma) := -\frac{1}{\beta} \ln Z_\sigma$ is the free energy associated with π_σ, it is enough to estimate $F(\sigma)$ for all σ in Σ_δ. Indeed, Z is self-averaging and we estimated $\ln \bar{Z}$ up to $o(k)$, moreover we have a polynomial uniform upper bound on F (polynomial in k), and an exponential concentration on Σ_δ. This implies that almost surely, for any $\delta > 0$,

$$|\mu(F) - \mu(F|\Sigma_\delta)| \leq e^{-C\delta^2 k^2/4}. \tag{86}$$

We will estimate $F(\sigma)$, i.e., $\ln Z_\sigma$, uniformely on Σ_δ in the following cases:

(A) $\tilde{h} > \tilde{h}_{\bar{c}}$,

(B) $\tilde{h} < \tilde{h}_{\bar{c}}$, and $C(\beta) < 0$, i.e., $T < T_{\bar{c}}$,

(C) $\tilde{h} < \tilde{h}_{\bar{c}}$, and $C(\beta) > 0$, i.e., $T > T_{\bar{c}}$.

For any positive δ, by definition of Σ_δ for k large large enough we have the following estimate for $\ln Z_\sigma$, for all σ in Σ_δ,

$$\max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} \phi_\sigma(\alpha,\bar{\alpha},\rho) \le \phi^*_\sigma \le \frac{1}{k^2}\ln Z_\sigma \le \phi^*_\sigma + \delta \le \max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} \phi_\sigma(\alpha,\bar{\alpha},\rho) + 2\delta. \tag{87}$$

We estimate $\max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} \phi_\sigma(\alpha,\bar{\alpha},\rho)$ in the three different cases.

Case (A): Since $s_\sigma(\alpha,\rho) \ge 0$ and

$$\max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} s_\sigma(\alpha,\bar{\alpha},\rho) \le \delta\frac{\ln(1/p)}{c} \tag{88}$$

we have

$$-\beta(1+\tilde{h})\delta - \beta f'(2\beta) \le \frac{1}{k^2}\ln Z_\sigma \le -\beta f'(2\beta) + (2+\beta)\delta + \delta\frac{\ln(1/p)}{c}. \tag{89}$$

by using that c goes to $\bar{c}$, we conclude that, almost surely,

$$\lim_{k\to+\infty}\frac{S}{k^2} = \frac{\ln(1/p)}{\bar{c}} - \frac{f(2\beta)}{2} + \beta f'(2\beta). \tag{90}$$

In cases (B) and (C) we will need a concentration result on the numbers of sites i outside σ such that $h_i(\sigma) = j + \tilde{h}k$, i.e., $g_{j,1} = |\mathcal{I}_{j,1}|$ (see (14).

Lemma 5.4 *Let*

$$J_c := \{j \in \mathbb{N} : \frac{j}{k} \in X_c\} \tag{91}$$

with X_c defined in (67). With probability 1, for any $\delta > 0$, if k is large enough then, for all σ in $\mathcal{X}_k^{(n)}$, for $j \in J_c$ we have:

$$\exp\left\{k\left(-\delta + \frac{\ln(1/p)}{c} - I_p\left(\frac{j}{k}\right)\right)\right\} \le g_{j,1} \le \exp\left\{k\left(\delta + \frac{\ln(1/p)}{c} - I_p\left(\frac{j}{k}\right)\right)\right\}; \tag{92}$$

for $j \notin J_c$ we have

$$g_{j,1} \le e^{k\delta}. \tag{93}$$

Proof: The random variable $g_{j,1}$ follows a binomial law with parameters $n-k \leq n$ and

$$\binom{k}{j}(1-p)^j p^{k-j} = e^{-kI_p(j/k)+o(k)}, \tag{94}$$

so that the usual large deviation estimates give

$$\begin{aligned} &\mathbb{P}\left(g_{j,1} \geq \exp\left\{k\left(\delta + \left[\frac{\ln(1/p)}{c} - I_p\left(\frac{j}{k}\right)\right]_+\right)\right\}\right) \\ \leq \quad &\exp\left\{-k\left(\delta + \left[I_p\left(\frac{j}{k}\right) - \frac{\ln(1/p)}{c}\right]_+ + o(1)\right) e^{k(\delta+[\ln(1/p)/c-I_p(j/k)]_+)}\right\} \end{aligned} \tag{95}$$

and, if $\ln(1/p)/c \geq I_p(j/k)$, i.e., $j > j_c$

$$\mathbb{P}\left(g_{j,1} \leq \exp\left\{k\left(-\delta + \frac{\ln(1/p)}{c} - I_p\left(\frac{j}{k}\right)\right)\right\}\right) \leq \exp\left\{-e^{k(\ln(1/p)/c-I_p(j/k)+o(1))}\right\}. \tag{96}$$

Since the number of configurations σ is not larger than $e^{k^2\ln(1/p)/c}$, we obtain our result with the Borel-Cantelli lemma. □

By Lemma 5.4 we can obtain the following results:

In case (B) $T < T_{\bar{c}}$ i.e.,$f'(\beta) \in [0,1]\backslash X_{\bar{c}}$, there exists a constant a_2 such that, almost surely, for all k large enough,

$$\max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} s_\sigma(\alpha,\bar{\alpha},\rho) \leq a_2\delta \tag{97}$$

This immediately follows from (93).

In case (C) $T > T_{\bar{c}}$ i.e., $f'(\beta) \in X_{\bar{c}}$, there exists a constant a_3 such that, almost surely, for all k large enough,

$$\max_{\alpha\in\mathcal{Q}_\delta,\bar{\alpha}\in\bar{\mathcal{Q}}_\delta,\rho\in\mathcal{H}_\delta} s_\sigma(\alpha,\bar{\alpha},\rho) < \frac{\ln 1/p}{\bar{c}} - I_p(f'(\beta)) + a_3\delta \tag{98}$$

The proof of this entropy estimates can be found in Appendix D. It is absolutely standard but we give it not only for completeness but also to show that the point of view of the Fermi statistics is a useful tool. The main idea is that in the asymptotics $k\to\infty$, due to the convexity property of I_p, the entropy is essentially due to the sites i with cavity field $h_i(\sigma)$ such that $\frac{h_i(\sigma)}{k} \in (f'(\beta)+\tilde{h}-\delta, f'(\beta)+\tilde{h}+\delta)$ and the number of such sites is estimated by Lemma 5.4.

With the entropy estimates (97) and (98) we can easily complete our proof.

Case (B): By equations (87) and (97) if σ is in Σ_δ, almost surely, for all k large enough,

$$-\beta f'(\beta) - \beta\delta - \beta\tilde{h} \leq \frac{1}{k^2}\ln Z_\sigma \leq -\beta f'(\beta) + \beta\delta - \beta\tilde{h}(1-\delta) + a_2\delta + 2\delta. \qquad (99)$$

We conclude that, almost surely,

$$\lim_{k\to+\infty} \frac{S}{k^2} = 2\frac{\ln(1/p)}{\bar{c}} - f(\beta) + \beta f'(\beta). \qquad (100)$$

Case (C): Again by equations (87) and (98) if σ is in Σ_δ, almost surely, for all k large enough,

$$\frac{1}{k^2}\ln Z_\sigma \leq -\beta f'(\beta) + \beta\delta - \beta\tilde{h}(1-\delta) + \frac{\ln(1/p)}{c} - I_p(f'(\beta)) + a_3\delta + 2\delta. \qquad (101)$$

We conclude that, almost surely,

$$\lim_{k\to+\infty} \frac{S}{k^2} \geq 2\frac{\ln(1/p)}{\bar{c}} - f(\beta) - \frac{\ln(1/p)}{c} + I_p(f'(\beta)) + \beta f'(\beta) = \frac{\ln(1/p)}{\bar{c}}. \qquad (102)$$

where we used that f is the Legendre transform of $-I_p$. The opposite estimate

$$\lim_{k\to+\infty} \frac{S}{k^2} \leq \frac{\ln(1/p)}{\bar{c}}$$

is trivial.

A The functions $f(\beta)$ and $I_p(x)$

We give here some inequalities for the function $f(\beta) := -\ln[p + (1-p)e^{-\beta}]$ defined in the main theorem.

This is a non negative concave function with $f(0) = 0$, $\lim_{\beta\to\infty} f(\beta) = \ln 1/p = I_p(0)$; from its concavity we immediately obtain the following estimates:

$$f(\beta) > \frac{f(l\beta)}{l} \qquad \forall l > 1$$

$$B(\beta) := f(\beta) + f(2\beta) - f(3\beta) > 0$$

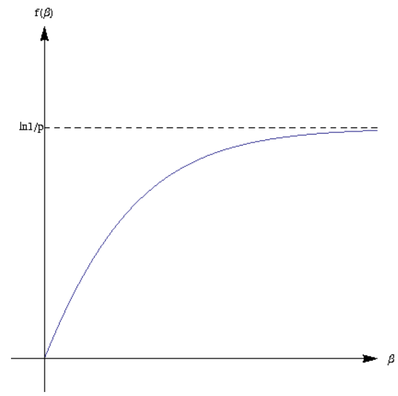

Figure 3: The function $f(\beta)$

since both the functions $F(\beta) := f(\beta) - \frac{f(l\beta)}{l}$ and $B(\beta)$ are strictly increasing function vanishing at zero. Moreover we have:

$$f(2\beta) + f(3\beta) > f(\beta) + f(4\beta)$$

which is an immediate consequence of concavity, and

$$f(\beta) + f(3\beta) - f(2\beta) - \frac{f(4\beta)}{2} > 0$$

since $f(\beta) + f(3\beta) - f(2\beta) - \frac{f(4\beta)}{2} = f(\beta) - \frac{f(2\beta)}{2} + f(3\beta) - \frac{f(2\beta)}{2} - \frac{f(4\beta)}{2}$ again positive by concavity.

For $p \in (0,1)$ and $x \in [0,1]$ consider now the binomial large deviation functional defined in (63), $I_p(x) = x \ln \frac{x}{1-p} + (1-x) \ln \frac{1-x}{p}$. This is a convex non negative function with minimum at $x = 1-p$ and $I_p(0) = \ln 1/p, I_p(1) = \ln 1/(1-p)$. By recalling the asymptotic behavior for the binomial coefficient:

$$\binom{L}{l} \sim (2\pi)^{-1/2} [x^x (1-x)^{1-x}]^{-L} (x(1-x)L)^{-1/2}$$

with $x = \frac{l}{L}$ (see for instance [1] pg.4) we immediately obtain

$$\binom{L}{l} (1-p)^l p^{L-l} = e^{-L I_p(x) + o(L)}.$$

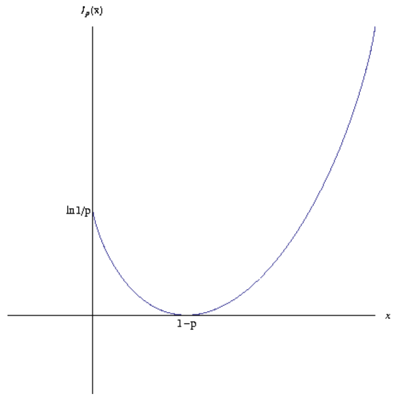

Figure 4: The function $I_p(x)$ for $p = 2/3$

The functions $f(\beta)$ and $I_p(x)$ are related by a Legendre transform. Indeed we have

$$I_p(x) + \beta x \geq f(\beta)$$

where the equality holds only for $x = f'(\beta)$. By evaluating the critical point of the function $I_p(x) + \beta x$ we have

$$I_p'(x) = \ln \frac{xp}{(1-p)(1-x)} = -\beta$$

ans so the critical point is

$$x_0 = \frac{(1-p)e^{-\beta}}{p + (1-p)e^{-\beta}} = f'(\beta)$$

and this is a minimum due to the convexity of $I_p(x)$.

In particular we have

$$I_p\text{''}(x) = \frac{1}{x(1-x)} \geq 2$$

B Proof of Lemma 3.1

Indeed to prove (50) we note that, the coefficients C_r in

$$\Psi = \frac{g_5(g_5-1)}{2}\Big(2f(2\beta) - f(4\beta)\Big) + \frac{1}{2}\sum_{r=1}^{9} g_r C_r$$

defined in equations (41), (45), (42), (46), (49), (47), (43), (48), (44), can be estimated by using the concavity of the function $f(\beta)$ so that $0 \leq 2f(\beta) - f(2\beta) \leq f(\beta) + f(2\beta) -$

$f(3\beta) =: B$. Indeed by using the constraints (35) and (36) we can estimate the coefficient:

$$C_r \leq B(k \wedge g) \quad r \neq 5, \qquad C_5 \leq B(g - g_5)$$

so that

$$\Psi \leq g_5(g_5 - 1)\Big(f(2\beta) - \frac{f(4\beta)}{2}\Big) + \frac{1}{2}B((k \wedge g) + g_5)(g - g_5) = \bar{\Psi}$$

C Proofs of Lemmas 3.2 and 3.3

Proof of Lemma 3.2 We look now for the maximum of the function

$$\bar{\Theta}_2 + \Phi_2 + \bar{\Psi} = (4k - q - q' - g)\ln n - \ln\Big((q-g)!(q'-g)!\Big) - 2\ln\Big((k-q-g)!(k-q'-g)!\Big) + C\Big)$$

$$+ \Phi_2 + g_5(g_5 - 1)\Big(f(2\beta) - \frac{f(4\beta)}{2}\Big) + \frac{1}{2}B((k \wedge g) + g_5)(g - g_5) \tag{103}$$

By noting the symmetry of this function in the parameters q and q' and the fact that the constraints are in the form $g \leq (2k - q) \wedge (2k - q')$ and $g_5 \leq q \wedge q' \wedge g$, we immediately can conclude that the maximum is obtained for $q = q'$. So we have only to study the function $a(q, g, g_5) - b(q, g, g_5)$ on the polytope $\bar{\mathcal{P}}$ where

$$a(q, g, g_5) = -2\beta h(k - q) - 2f(\beta)(k^2 - q^2) - f(2\beta)q(q - 1) + g_5(g_5 - 1)\Big(f(2\beta) - \frac{f(4\beta)}{2}\Big) +$$

$$+ \frac{1}{2}B((k \wedge g) + g_5)(g - g_5) + (4k - 2q - g)\ln n \tag{104}$$

$$b(q, g, g_5) = 2\ln\Big((q - g)!\Big) + 4\ln\Big((k - q - g)!\Big)\Big) + C \tag{105}$$

and $\mathcal{P}$ (see Figure 2) is defined by the relations:

$$0 \leq g \leq 2k - q, \quad 0 \leq g_5 \leq g \wedge q, \quad 0 \leq q \leq k \tag{106}$$

We first study the maximum of the function a on $\bar{\mathcal{P}}$. For $g > k$ the hessian of a is given by

$$\begin{pmatrix} [4f(\beta) - 2f(2\beta)] & 0 & 0 \\ 0 & 0 & \frac{1}{2}B \\ 0 & \frac{1}{2}B & [2f(\beta) - f(4\beta) - B] \end{pmatrix} \tag{107}$$

and for $g \le k$ the hessian of a is given by

$$\begin{pmatrix} [4f(\beta) - 2f(2\beta)] & 0 & 0 \\ 0 & B & \frac{1}{2}B \\ 0 & \frac{1}{2}B & [2f(\beta) - f(4\beta) - B] \end{pmatrix} \tag{108}$$

Again by the concavity of the function $f(\beta)$ in both cases we have a positive eigenvalue $\lambda_1 = 4f(\beta) - 2f(2\beta)$ and two real eigenvalues with $\lambda_2 + \lambda_3 > 0$ if $g \le k$ and $\lambda_2\lambda_3 < 0$ if $g > k$ so at least two positive eigenvalues. We can conclude that the maximum of a is obtained on the edges of $\bar{\mathcal{P}}$. By studying the function $a(\mathbf{x})$, with $\mathbf{x} = (q, g, g_5)$, on all the edges we easily check that the maximum actually is obtained on the vertices. To this purpose we used the convexity relations of $f(\beta)$ listed in appendix A. By a direct comparison we obtain that the maximum is obtained in the point $\mathbf{x}_{max} = (k, 0, 0)$ for $\tilde{h} > \tilde{h}_c$ and in $\mathbf{x}_{max} = (0, 0, 0)$ for $\tilde{h} < \tilde{h}_c$ as soon as $f(2\beta) - \frac{1}{2}f(4\beta) < \frac{\ln(1/p)}{c}$. This inequality holds for all β when $c \in (1, 2]$, while in the case $c > 2$ we can simply add the hypothesis $\beta < \bar{\beta}_c$ to conclude.

Fix now $\alpha \in (0, 1)$, in the region $\bar{\mathcal{P}} \cap \{g < k^\alpha\}$ we have that $a(\mathbf{x}) - b(\mathbf{x})$ is a decreasing function of g at q, g_5 fixed and large k, and on the surface $g = g_5$ again is a decreasing function of g for large k. On the other hand we have for $\mathbf{x} \in \bar{\mathcal{P}} \cap \{g > k^\alpha\}$ that $a(\mathbf{x}) < a(\mathbf{x}_{max}) - b(\mathbf{x}_{max})$, so that, as in the discussion of $\bar{Z}$, by noting that the function b is non-negative, we can conclude that the points $\mathbf{x}_{max}$ correspond to maximal values for the function $a(\mathbf{x}) - b(\mathbf{x})$.

D Proof of equation (98)

We have to estimate

$$N(\sigma, \alpha, \rho) := |\{\tau \in \mathcal{X}_k^{(n)} : q(\sigma, \tau) = k\alpha,\, H_0(\sigma, \tau) = k^2\rho\}|, \tag{109}$$

for $\alpha \in [0, \delta]$ and $\rho \in [f'(\beta) - \delta, f'(\beta) + \delta]$ with $f'(\beta) \in X_c$. We have $H_0(\sigma, \tau) = H_0(\sigma, I) + H_0(\sigma, T)$ and so we get $N(\sigma, \alpha, \rho) = \sum_{\rho' \in [\rho - \delta, \rho + 2\delta]} N_1(\sigma, \alpha, \rho') N_2(\sigma, \alpha, \rho')$ with

$$N_1(\sigma, \alpha, \rho') = |\{A \in V \backslash \sigma : \ |A| = (1 - \alpha)k,\, H_0(\sigma, A) = k^2(1 - \alpha)\rho'\}|,$$

$$N_2(\sigma,\alpha,\rho') = |\{A\in\sigma:\ |A|=\alpha k,\ H_0(\sigma,A)=k^2\alpha\frac{\rho-\rho'(1-\alpha)}{\alpha}\}|.$$

The term N_2 is easily estimated from above by $2^k = e^{o(k^2)}$. As far as the term N_1 is concerned we can use the notation of the Fermi statistics and in particular (15), to write

$$N_1(\sigma,\alpha,\rho') = \sum_{\substack{\{n_{j,1}\}_{j=0,1,\dots,k}:\\ \sum_j n_{j,1}=(1-\alpha)k,\\ \sum_j n_{j,1}j=k^2(1-\alpha)\rho'}} \prod_j \binom{g_{j,1}}{n_{j,1}} \tag{110}$$

By using the Stirling formula we can approximate the binomial coefficient as follows:

$$\binom{g}{n} = e^{-g\mathcal{E}(\frac{n}{g})+o(k^2)} \tag{111}$$

with

$$\mathcal{E}(x) := x\ln x + (1-x)\ln(1-x)$$

obtaining:

$$\sum_{\substack{\{n_{j,1}\}_{j=0,1,\dots,k}:\\ \sum_j n_{j,1}=(1-\alpha)k,\\ \sum_j n_{j,1}j=k^2(1-\alpha)\rho'}} \prod_j \binom{g_{j,1}}{n_{j,1}} = \exp\{\max_{\mathbf{x}}\sum_j[-g_{j,1}(\mathcal{E}(x_j)] + o(k^2)\}$$

with $\mathbf{x} = (x_j)_{j\in\{0,1,\dots,k\}}$, where $x_j := \frac{n_{j,1}}{g_{j,1}}$, and the maximum is under the constraints $\sum_j g_{j,1}x_j = (1-\alpha)k$ and $\sum_j g_{j,1}x_j j = k^2(1-\alpha)\rho'$. With the Lagrange multiplier method and standard computation, we can evaluate this maximum by looking at the maximum of the function

$$F(\mathbf{x},\lambda,\mu) = \sum_j g_{j,1}\Big[-\mathcal{E}(x_j) - (\lambda+\mu j)x_j\Big] \tag{112}$$

which is reached in $\bar{\mathbf{x}}$ with $\bar{x}_j = \frac{1}{1+e^{\lambda+\mu j}}$ with λ and μ solution of the equations

$$\sum_j g_{j,1}\bar{x}_j = (1-\alpha)k \qquad \text{and} \qquad \sum_j g_{j,1}\bar{x}_j j = k^2(1-\alpha)\rho'. \tag{113}$$

In this maximum $\bar{\mathbf{x}}$ we have

$$\sum_j[-g_{j,1}(\mathcal{E}(\bar{x}_j)] = \lambda(1-\alpha)k + \mu(1-\alpha)k^2\rho' + o(k^2). \tag{114}$$

By Lemma 5.4 we have that for $j \in J_c$, $\bar{x}_j$ must be exponentially small in k and for any $\delta \geq 0$ we have

$$\sum_{j \in J_c} g_{j,1} x_j = \sum_{j \in J_c} e^{k[\frac{\ln 1/p}{c} - I_p(\frac{j}{k})] - \lambda - k\mu\frac{j}{k} + \mathcal{O}(\delta k)}.$$

Due to the fact that $f'(\beta) \in X_c$, this sum is not exponentially small, i.e.,

$$k[\frac{\ln 1/p}{c} - I_p(\frac{j}{k})] - \lambda - k\mu\frac{j}{k} = \mathcal{O}(\delta k)$$

for some $j \in J_c$, and so we can conclude that

$$\max_{j \in J_c} k[\frac{\ln 1/p}{c} - I_p(\frac{j}{k})] - \lambda - k\mu\frac{j}{k} = k[\frac{\ln 1/p}{c} - f(\mu)] - \lambda = \mathcal{O}(\delta k)$$

that is $\lambda = k[\frac{\ln 1/p}{c} - f(\mu)] + \mathcal{O}(\delta k)$ and so, by (114) that

$$N_1(\sigma, \alpha, \rho') \leq \exp\left\{\{k[\frac{\ln 1/p}{c} - f(\mu)] + \mathcal{O}(\delta k)\}(1-\alpha)k + \mu(1-\alpha)k^2\rho'\right\} =$$

$$= \exp\{k^2(1-\alpha)[\frac{\ln 1/p}{c} - f(\mu) + \mu\rho'] + \mathcal{O}(\delta k^2)\}$$

By recalling that $\rho' \in [\rho - \delta, \rho + 2\delta] = [f'(\beta) - 2\delta, f'(\beta) + 3\delta]$ and the Legendre transformation between f and I_p implying that $\mu f'(\beta) = f(\mu) - I_p(f'(\beta))$ the proof of (97) and (98) follows straightforward.

Acknowledgments: We thank Antonio Iovanella, coauthor of the numerical parts of this project; the first numerical results he found have been the starting point of our investigations. A special thank is due to Fabio Martinelli for discussions and for the nice and stimulating environment he created in Math Department of Università di Roma Tre. Thanks to Prasad Tetali for discussions, bibliographic suggestions and encouragement.

References

[1] B. Bollobas *Random graph*, 2nd ed.,Cambridge University Press, 2001.

[2] F. den Hollander, *Large Deviations*, AMS - Fields Institute Monographs , Providence - Rhode Island, 2000.

[3] M. R. Garey, D.S. Johnson, *Computer and Intractability: A guide to the theory of NP-completeness*, Freeman, New York, 1976.

[4] A.Gaudilliere, J.Reygner *Sampling the Fermi statistics and other conditional product measures (2009)*, to appeaer in Annals of the Institute H. Poincaré, arXiv:0911.4565v1.

[5] K. Huang, *Statistical Mechanics*, Wiley, New York, 1987.

[6] A. Iovanella, B. Scoppola, E. Scoppola, *Some Spin Glass Ideas Applies to the Clique Problem*, JSP 126, 4/5 (2007) 895-915.

[7] M. Jerrum, *Large cliques elude the metropolis process*, Random Structures and Algorithms, 3, 4 (1992), 347–359.

[8] A.Coja-Oghlan, C.Efthymiou, *On independent sets in random graphs*, arXiv:1007.1378v1.