On independent sets in random graphs

Amin Coja-Oghlan and Charilaos Efthymiou
University of Warwick, Mathematics and Computer Science, Coventry CV4 7AL, UK
{a.coja-oghlan,c.efthymiou}@warwick.ac.uk

November 10, 2010

Abstract

The independence number of a sparse random graph $G(n, m)$ of average degree $d = 2m/n$ is well-known to be $\alpha(G(n, m)) \sim 2n \ln(d)/d$ with high probability. Moreover, a trivial greedy algorithm w.h.p. finds an independent set of size $(1 + o(1)) \cdot n \ln(d)/d$, i.e., half the maximum size. Yet in spite of 30 years of extensive research no efficient algorithm has emerged to produce an independent set with $(1 + \varepsilon)n \ln(d)/d$, for any fixed $\varepsilon > 0$. In this paper we prove that the combinatorial structure of the independent set problem in random graphs undergoes a phase transition as the size $k$ of the independent sets passes the point $k \sim n \ln(d)/d$. Roughly speaking, we prove that independent sets of size $k > (1 + \varepsilon)n \ln(d)/d$ form an intricately ragged landscape, in which local search algorithms are bound to get stuck. We illustrate this phenomenon by providing an exponential lower bound for the Metropolis process, a Markov chain for sampling independents sets.

Key words: random graphs, independent set problem, Metropolis process, phase transitions.
1 Introduction and Results

1.1 Probabilistic analysis and the independent set problem

In the early papers on the subject, the motivation for the probabilistic analysis of algorithms was to alleviate the glum of worst-case analyses by a brighter ‘average-case’ scenario. This optimism was stirred by early analyses of simple, greedy-type algorithms, showing that these perform rather well on randomly generated input instances, at least for certain ranges of the parameters. Examples of such analyses include Grimmett and McDiarmid [13] (independent set problem), Achlioptas and Molloy [2] (graph coloring), and Frieze and Suen [10] (k-SAT). Yet, remarkably, in spite of 30 years of research, for many problems no efficient algorithms, howsoever sophisticated, have been found to outperform those simple greedy algorithms.

The independent set problem in random graphs $G(n, m)$ is a case in point. Recall that $G(n, m)$ is a graph on $n$ vertices obtained by choosing $m$ edges uniformly at random (without replacement). We say that $G(n, m)$ has a property with high probability if the probability that the property holds tends to 1 as $n \to \infty$. One of the earliest results in the theory of random graphs is a non-constructive argument showing that for $m = \frac{1}{2} \binom{n}{2}$ the independence number of $G(n, m)$ is $\alpha(G(n, m)) \sim 2 \log_2(n)$ w.h.p. [7, 5, 20]. Grimmett and McDiarmid [13] analyzed a simple algorithm that just constructs an inclusion-maximal independent set greedily on $G(n, m)$: it yields an independent set of size $(1 + o(1)) \log_2 n$ w.h.p., about half the maximum size. But no algorithm is known to produce an independent set of size $(1 + \varepsilon) \log_2 n$ for any fixed $\varepsilon > 0$ in polynomial time with a non-vanishing probability, neither on the basis of a rigorous analysis, nor on the basis of experiments or other evidence. In fact, devising such an algorithm is probably the most prominent open problem in the algorithmic theory of random graphs [9, 16]. (However, note that one can find a maximum independent set w.h.p. by trying all $n^{O(lnn)}$ possible sets of size $2 \log_2 n$.)

Matters are no better for sparse random graphs. If we let $d = 2m/n$ denote the average degree, then non-constructive arguments yield $\alpha(G(n, m)) \sim \frac{2 \ln(d)}{d} \cdot n$ for the regime $1 \ll d = o(n)$. In the case $d \gg \sqrt{n}$, the proof of this fact is via a simple second moment argument [5, 20]. By contrast, for $1 < d \ll \sqrt{n}$, the second moment argument breaks down and additional methods such as large deviations inequalities are needed [8]. Yet in either case, no algorithm is known to find an independent set of size $(1 + \varepsilon) \frac{\ln(d)}{d} \cdot n$ in polynomial time with a non-vanishing probability, while ‘greedy’ yields an independent set of size $(1 + o(1)) \frac{\ln(d)}{d} \cdot n$ w.h.p. In the sparse case, the time needed for exhaustive search scales as $\exp(\frac{1}{2} \ln(d) \ln(n/d))$, i.e., the complexity grows as $d$ decreases.

In the present paper we explore the reason for the apparent ‘hardness’ of finding large independent sets in random graphs. The focus is on sparse random graphs, both conceptually and computationally the most difficult case. We exhibit a phase transition in the structure of the problem that occurs as the size of the independent sets passes the point $\frac{\ln(d)}{d} \cdot n$ up to which efficient algorithms are known to succeed. Roughly speaking, we show that independent sets of of sizes bigger than $(1 + \varepsilon) \frac{\ln(d)}{d} \cdot n$ form an intricately ragged landscape, in which local-search algorithms are likely to get stuck. Thus, ironically, instead of showing that the ‘average case’ scenario is brighter, we end up suggesting that random graphs provide a rich source of worst-case examples.

Taking into account the (substantially) different nature of the independent set problem, our work complements the results obtained in [11] for random constraint satisfaction problem such as k-SAT or graph coloring and in [21] for and a wide class of symmetric CSPs. Furthermore, very recently B. Rossman [23] obtained a monotone circuit lower bound for the clique problem on random graphs that is exponential in the size of the clique. The setup of [23] is somewhat orthogonal to our contribution, as we are concerned with the case that the size of the desired object (i.e., the independent set) is linear in the number of vertices, while [23] deals with the case that the size of the clique is $O(1)$ in terms of the order of the graph. Nevertheless, the punchline of viewing random graphs as a source of computational
hardness is similar.

1.2 Results

Throughout the paper we will be dealing with sparse random graphs where the average degree \( d = 2m/n \) is ‘large’ remains bounded as \( n \to \infty \). To express this, we will refer to functions \( \varepsilon_d \) that tends to zero as \( d \) gets large. In this case \( \alpha(G(n, m)) = (2 - \varepsilon_d) \ln d / d \cdot n \) and the greedy algorithm finds independent sets of size \((1 + \varepsilon_d) \ln d / d \cdot n \) w.h.p., where \( \varepsilon_d, \varepsilon_d' \to 0 \) for large \( d \). However, no efficient algorithm is known to find independent sets of size \((1 + \varepsilon') \ln d / d \cdot n \) for any fixed \( \varepsilon' > 0 \).

For a graph \( G \) and an integer \( k \) we let \( S_k(G) \) denote the set of all independent sets of size exactly \( k \). What we will show is that in \( G(n, m) \) the set \( S_k(G(n, m)) \) undergoes a phase transition as \( k \sim \frac{\ln d}{d} \cdot n \).

To state the result for \( k \) smaller than \( \frac{\ln d}{d} \cdot n \), we need the following concept. Let \( S \) be a set of subsets of \( V \), and let \( \gamma > 0 \) be an integer. We say that \( S \) is \( \gamma \)-connected if for any two sets \( \sigma, \tau \in S \) there exist \( \sigma_1, \ldots, \sigma_N \in S \) such that \( \sigma_1 = \sigma, \sigma_N = \tau \), and \( \text{dist}(\sigma_t, \sigma_{t+1}) \leq \gamma \) for all \( 1 \leq t < N \). If \( S_k(G(n, m)) \) is \( \gamma \)-connected for some \( \gamma = O(1) \), one can easily define various simple Markov chains on \( S_k(G) \) that are ergodic.

**Theorem 1** There is \( \varepsilon_d \to 0 \) such that \( S_k(G(n, m)) \) is \( O(1) \)-connected w.h.p. for any

\[
k \leq (1 - \varepsilon_d) \frac{\ln d}{d} \cdot n.
\]

The proof actually yields an efficient algorithm for finding a path of length \( O(n) \) between two given independent sets of size \( k \) w.h.p.

By contrast, our next result shows that for \( k > \frac{\ln d}{d} \cdot n \) the set \( S_k(G(n, m)) \) is not just disconnected w.h.p., but that it shatters into exponentially many, exponentially tiny pieces.

**Definition 1** We say that \( S_k(G(n, m)) \) **shatters** if there exist constants \( \gamma, \zeta > 0 \) such that w.h.p. \( S_k(G(n, m)) \) admits a partition into classes so that:

1. Each class contains at most \( \exp(-\gamma n) |S_k(G(n, m))| \) independent sets.

2. If two independent sets belong to different classes, their distance is at least \( \zeta n \).

**Theorem 2** There is \( \varepsilon_d \to 0 \) so that \( S_k(G(n, m)) \) shatters for all \( k \) with

\[
(1 + \varepsilon_d) \frac{\ln d}{d} \cdot n \leq k \leq (2 - \varepsilon_d) \frac{\ln d}{d} \cdot n.
\]

Theorems [1] and [2] deal with the geometry of a single ‘layer’ \( S_k(G(n, m)) \) of independent sets of a specific size. The following two results explore if/how a ‘typical’ independent set in \( S_k(G(n, m)) \) can be extended to a larger one. To formalize the notion of ‘typical’, we let \( \Lambda_k(n, m) \) signify the set of all pairs \((G, \sigma)\), where \( G \) is a graph on \( n \) vertices \( V = \{1, \ldots, n\} \) and \( \sigma \) of \( G \) is an independent set of size \( k \). Let \( \mathcal{U}_k(n, m) \) be the probability distribution induced on \( \Lambda_k(n, m) \) by the following experiment.

Choose a graph \( G = G(n, m) \).

If \( \alpha(G) \geq k \), choose an independent set \( \sigma \in S_k(G) \) uniformly at random.

\footnote{The reason why we need to speak about \( d \) ‘large’ is that the sparse random graph \( G(n, m) \) is not connected. This implies, for instance, that algorithms can find independent sets of size \((1 + \varepsilon_d)n \ln(d)/d\) for some \( \varepsilon_d \to 0 \) by optimizing carefully over the small tree components of \( G(n, m) \). Our results/proofs actually carry over to the case that \( d = d(n) \) tends to infinity as \( n \) grows, but to keep matters as simple as possible, we will focus on fixed \( d \) in this extended abstract.}
We say a pair $(G, \sigma)$ chosen from the distribution $U_k(n, m)$ has a property $\mathcal{P}$ with high probability if the probability of $(G, \sigma) \in \mathcal{P}$ tends to one as $n \to \infty$.

**Definition 2** Let $\gamma, \delta \geq 0$, let $G$ be a graph, and let $\sigma$ be an independent set of $G$. We say that $(G, \sigma)$ is $(\gamma, \delta)$-expandable if $G$ has an independent set $\tau$ such that $|\tau| \geq (1 + \gamma)|\sigma|$ and $|\tau \cap \sigma| \geq (1 - \delta)|\sigma|$.

**Theorem 3** There is $\varepsilon_d \to 0$ such that for any $\varepsilon_d \leq \varepsilon \leq 1 - \varepsilon_d$ the following is true. For $k = (1 - \varepsilon)\ln d \cdot n$ a pair $(G, \sigma)$ chosen from the distribution $U_k(n, m)$ is $(2\varepsilon/(1 - \varepsilon), 0)$-expandable w.h.p.

Theorem 3 shows that w.h.p. in a random graph $G(n, m)$ almost all independent sets of size $k = (1 - \varepsilon)\ln d \cdot n$ are contained in some bigger independent set of size $(1 + \varepsilon)\ln d \cdot n$. That is, they can be expanded beyond the critical size $\ln d \cdot n$ where shattering occurs. However, as $k$ approaches the critical size $\ln d \cdot n$, i.e., as $\varepsilon \to 0$, the typical potential for expansion shrinks.

**Theorem 4** There is $\varepsilon_d \to 0$ such that for $\varepsilon_d \leq \varepsilon \leq 1 - \varepsilon_d$ and $k = (1 + \varepsilon)\ln d \cdot n$ w.h.p. a pair $(G, \sigma)$ chosen from the distribution $U_k(n, m)$ is not $(\gamma, \delta)$-expandable for any $\gamma > \varepsilon_d$ and

$$\delta < \gamma + \frac{2(\varepsilon - \varepsilon_d)}{1 + \varepsilon}.$$

In other words, Theorem 3 shows that for $k = (1 + \varepsilon)\ln d \cdot n$, a typical $\sigma \in S_k(G(n, m))$ cannot be expanded to an independent set of size $(1 + \gamma)k$, $\gamma > \varepsilon_d$ without first reducing its size below

$$(1 - \delta)k = (1 - \varepsilon - \gamma(1 + \varepsilon) + 2\varepsilon_d)\ln d \cdot n < \ln d \cdot n.$$  

(However, a random independent set of size $k \leq (2 - \varepsilon_d)\ln d n/d$ is typically not inclusion-maximal because, for instance, it is unlikely to contain all isolated vertices of the random graph $G(n, m)$.)

Metaphorically, the above results show that finding a maximum independent set in $G(n, m)$ is like climbing a ragged mountain range in dense fog, in quest of the highest summit. Beyond the ‘plateau level’ $k \sim \ln d \cdot n$ there is an abundance of smaller ‘peaks’, i.e., independent sets of sizes $(1 + \varepsilon)k$, almost all of which are not expandable. Thus, ascending blindly, we are bound to get stuck way below the overall summit of height $2\ln d \cdot n$. Furthermore, once we find ourselves on one of the lesser peaks of height $(1 + \varepsilon)k$, we will have to return all the way to plateau level to try anew.

The algorithmic equivalent of such a mountaineer is a Markov chain called the *metropolis process*. For a given graph $G$ its state space is the set of all independent sets of $G$. Let $I_t$ be the state at time $t$. In step $t + 1$, the process chooses a vertex $v$ of $G$ uniformly at random. If $v \in I_t$, then with probability $1/\lambda$ the next state is $I_{t+1} = I_t \setminus \{v\}$, and with probability $1 - 1/\lambda$ we let $I_{t+1} = I_t$, where $\lambda \geq 1$ is a ‘temperature’ parameter. If $v \notin I_t \cup N(I_t)$, then $I_{t+1} = I_t \cup \{v\}$. Finally, if $v \in N(I_t)$, then $I_{t+1} = I_t$. The probability of an independent set $S$ of $G$ in the stationary distribution equals $\lambda^{|S|}/Z(G, \lambda)$, where

$$Z(G, \lambda) = \sum_{k=0}^{n} \lambda^k \cdot |S_k(G)|$$

is the partition function. Hence, the larger $\lambda$, the higher the mass of large independent sets. Let

$$\mu(G, \lambda) = \sum_{k=0}^{n} k\lambda^k \cdot |S_k(G)| / Z(G, \lambda)$$

denote the average size of an independent set of $G$ under the stationary distribution.

Our above results on the structure of the sets $S_k(G(n, m))$ imply that w.h.p. the mixing time of the metropolis process is exponential if the parameter $\lambda$ is tuned so that it tries to ascend to independent sets bigger than $\ln d \cdot n$.
Corollary 1 There is $\varepsilon_d \to 0$ such that for $\lambda > 1$ with

$$(1 + \varepsilon_d) \ln \frac{d}{d} \cdot n \leq \mathbb{E}[\mu(G(n, m), \lambda)] \leq (2 - \varepsilon_d) \ln \frac{d}{d} \cdot n.$$ \hfill (2)

the mixing time of the metropolis process on $G(n, m)$ is $\exp(\Omega(n))$ w.h.p.

1.3 Related work

To our knowledge, the connection between transitions in the geometry of the ‘solution space’ (in our case, the set of all independent sets of a given size) and the apparent failure of local algorithms in finding a solution has been pointed first out in the statistical mechanics literature \cite{15, 18, 11}. In that work, which mostly deals with CSPs such as $k$-SAT, the shattering phenomenon goes by the name of ‘dynamic replica symmetry breaking.’ Our present work is clearly inspired by the statistical mechanics ideas, although we are unaware of explicit contributions from that line of work addressing the independent set problem in the case of random graphs with average degree $d \gg 1$. Generally, the statistical mechanics work is based on deep, insightful, but, alas, mathematically non-rigorous techniques.

The recent work in \cite{3}, on reconstruction thresholds on weighted independent sets on trees, predicted some of the results of this work on the basis of the heuristic connection between reconstruction on trees and clustering of solution space of r-CSPs. Our result here come as a verification of all of these predictions.

In the case that the average degree $d$ satisfies $d \gg \sqrt{n}$, the independent set problem in random graphs is conceptually somewhat simpler than in the case of $d = o(\sqrt{n})$. The reason for this is that for $d \gg \sqrt{n}$ the second moment method can be used to show that the number of independent sets is concentrated about its mean. As we will see in Corollary \cite{2} below, this is actually untrue for sparse random graphs.

The work that is most closely related to ours is a remarkable paper of Jerrum \cite{14}, who studied the metropolis process on random graphs $G(n, m)$ with average degree $d = 2m/n > n^{2/3}$. The main result is that w.h.p. there exists an initial state from which the expected time for the metropolis process to find an independent set of size $(1 + \varepsilon) \ln d \cdot n$ is superpolynomial. This is quite a non-trivial achievement, as it is a result about the initial steps of the process where the states might potentially follow a very different distribution than the stationary distribution. The proof of this fact is via a concept called ‘gateways’, which is somewhat reminiscent of the expandability property in the present work. However, Jerrum’s proof hinges upon the fact that the number of independent sets of size $k \sim (1 + \varepsilon) \ln d \cdot n$ is concentrated about its mean. The arguments from the present work can be used to extend Jerrum’s result to the sparse case, where the expected time until a large independent set is found is fully exponential in $n$ w.h.p., via Theorem \cite{5} below. (Instead of showing this, we chose to present Corollary \cite{1} which just yields a bound on the mixing time of the Metropolis process.)

For extremely sparse random graphs, namely $d < e \approx 2.718$, finding a maximum independent set in $G(n, m)$ is easy. More specifically, the greedy matching algorithm of Karp and Sipser \cite{17} can easily be adapted so that it yields a maximum independent set w.h.p. But this approach does not generalize to average degrees $d > e$ (see, however, \cite{12} for a particular type of weighted independent sets).

The results of the present paper extend the main results from Achlioptas and Coja-Oghlan \cite{1}, which dealt with constraint satisfaction problems such as $k$-SAT or graph coloring, to the independent set problem. This requires new ideas, because the natural questions are somewhat different (for instance, the concept of ‘expandability’ has no counterpart in CSPs). Furthermore, in \cite{1} we conjectured but did not manage to prove the counterpart of Theorem \cite{1} on the connectivity of $S_k(G(n, m))$.

On a technical level, we owe to \cite{1} the idea of analyzing the distribution $\mathcal{U}_k(n, m)$ via a different distribution $P_k(n, m)$, the so-called ‘planted model’ (see Section \cite{2} for details). However, the proof that this approximation is indeed valid (Theorem \cite{5} below) requires a rather different approach. In \cite{1} we derived the corresponding result from the second moment method in combination with sharp threshold
results. By contrast, here we use an indirect approach that reduces the problem of estimating the number $|S_k(G(n, m))|$ of independent sets of a given size to the problem of (very accurately) estimating the independence number $\alpha(G(n, m))$. Indeed, the argument used here carries over to other problems, particularly random $k$-SAT, for which it yields a conceptually simpler proof than given in [1].

2 Approaching the distribution $U_k(n, m)$

2.1 The planted model

The main results of this paper deal with properties of ‘typical’ independent sets of a given size in a random graph, i.e., the probability distribution $U_k(n, m)$. In the theory of random discrete structures often the conceptual difficulty of analyzing a probability distribution is closely linked to the computational difficulty of sampling from that distribution (e.g., [22, Chapter 9]). This could suggest that analyzing $U_k(n, m)$ is a formidable task, because for $k > (1 + \varepsilon)n \ln(d)/d$ there is no efficient procedure known for finding an independent set of size $k$ in a random graph $G(n, m)$, let alone for sampling one at random. In effect, we do not know of an efficient method for sampling from $U_k(n, m)$.

To get around this problem, we are going to ‘approximate’ the distribution $U_k(n, m)$ by another distribution $P_k(n, m)$ on $\Lambda_k(n, m)$, the so-called planted model, that is easy to sample from. This distribution is induced by the following experiment:

Choose a subset $\sigma \subset [n]$ of size $k$ uniformly at random.
Choose a graph $G$ with $m$ edges in which $\sigma$ is an independent set uniformly at random.
Output the pair $(G, \sigma)$.

In other words, the probability assigned to a given pair $(G_0, \sigma_0) \in \Lambda_k(n, m)$ is

$$P_{P_k(n, m)}[(G_0, \sigma_0)] = \left[\binom{n}{k} \cdot \left(\frac{\binom{n}{k} - \binom{k}{m}}{m}\right)\right]^{-1},$$

i.e., $P_k(n, m)$ is nothing but the uniform distribution on $\Lambda_k(n, m)$. The key result that allows us to study the distribution $U_k(n, m)$ is the following

**Theorem 5** There is $\varepsilon_d \to 0$ such that for $k < (2 - \varepsilon_d)n \ln(d)/d$ we have

$$P_{U_k(n, m)}[E] \leq P_{P_k(n, m)}[E] \cdot \exp\left(14n\sqrt{\ln^5 d/d}\right) \quad \text{for any event } E \subset \Lambda_k(n, m).$$

To establish Theorem 5 we need to find a way to compare $P_k(n, m)$ and $U_k(n, m)$. Suppose that $k < (2 - \varepsilon_d)n \ln(d)/d$ is such that $\alpha(G(n, m)) \geq k$ w.h.p. Then the probability of a pair $(G_0, \sigma_0) \in \Lambda_k(n, m)$ under the distribution $U_k(n, m)$ is

$$P_{U_k(n, m)}[(G_0, \sigma_0)] \sim \left[\binom{n}{k} \cdot |S_k(G_0)|\right]^{-1}$$

(4) (because we first choose a graph uniformly, and then an independent set of that graph). Hence, the probabilities assigned to $(G_0, \sigma_0)$ under (3) and (4) coincide (asymptotically) iff

$$|S_k(G_0)| \sim \binom{n}{k} \left(\frac{\binom{n}{k} - \binom{k}{m}}{m}\right) / \binom{n}{m}.$$

(5) A moment’s reflection shows that the expression on the r.h.s. of (5) is precisely the expected number $E|S_k(G(n, m))|$ of independent sets of size $k$. Thus, $P_k(n, m)$ and $U_k(n, m)$ coincide asymptotically iff the number $|S_k(G(n, m))|$ of independent sets of size $k$ is concentrated about its expectation.
This is indeed the case in ‘dense’ random graphs with $m \gg n^{3/2}$. For this regime one can perform a ‘second moment’ computation to show that $|S_k(G(n, m))| \sim E|S_k(G(n, m))|$ w.h.p., whence the measures $P_k(n, m)$ and $U_k(n, m)$ are interchangeable. This fact forms (somewhat implicitly) the foundation of the proofs in [14].

By contrast, in the sparse case $m \ll n^{3/2}$ a straight second moment argument fails utterly. As it turns out, this is because the number $|S_k(G(n, m))|$ simply does not concentrate anymore. In fact, maybe somewhat surprisingly Theorem 5 can be used to infer the following corollary, which shows that in sparse random graphs the expectation $E|S_k(G(n, m))|$ ‘overestimates’ the typical number of independent sets by an exponential factor w.h.p.

**Corollary 2** There exist functions $\varepsilon_d \to 0$ and $g(d) > 0$ such that for $(1 + \varepsilon_d)n \ln(d)/dn \leq k < (2 - \varepsilon_d)n \ln(d)/d$ we have

$$|S_k(G(n, m))| \leq E|S_k(G(n, m))| \cdot \exp(-g(d)n) \quad \text{w.h.p.}$$

Conversely, in order to prove Theorem 5 we need to bound the ‘gap’ between the typical value of $|S_k(G(n, m))|$ and its expectation from above. This estimate can be summarized as follows.

**Proposition 1** There is $\varepsilon_d \to 0$ such that for $k < (2 - \varepsilon_d)n \ln(d)/d$ we have

$$|S_k(G(n, m))| \geq E|S_k(G(n, m))| \cdot \exp\left(14n \sqrt{\ln 5 \frac{d}{d^3}}\right) \quad \text{w.h.p.}$$

Using Proposition 1, equations (3) and (4), and a double counting, one completes the proof of Theorem 5 fairly easily.

## 3 Preliminaries

In this section we present some preliminary, technical, results which are going to be very useful in rest of this work. The first theorem we present is known as the Chernoff bounds.

**Theorem 6** Let $I_1, I_2, \ldots, I_n$ be independent Bernoulli trials such that $Pr[I_i = 1] = p_i$, for $1 \leq i \leq n$ and $0 \leq p_i \leq 1$. Then, for $X = \sum_{i=1}^{n} I_i$, $\mu = E[X]$ and $0 < \delta \leq 1$

$$Pr[X < (1 - \delta)\mu] \leq \exp \left(-\mu \delta^2 / 2\right) \quad (6)$$

and for $0 < \delta < 2e - 1$

$$Pr[X > (1 + \delta)\mu] \leq \exp \left(-\mu \delta^2 / 4\right). \quad (7)$$

**Theorem 7** Let $I_1, I_2, \ldots, I_n$ be Bernoulli trials such that for $1 \leq i \leq n$ $Pr[I_i = 1] = p_i$, where $0 < p_i < 1$. Furthermore, assume that the variables are negatively dependent, i.e.

$$Pr[I_i = 1|I_j = 1] \leq Pr[I_i = 1] \quad \text{for any } i, j.$$

For the variable $Y = \sum_{i=1}^{n} I_i$, the concentration bounds of Theorem 6 still apply.

In many cases we can avoid a large deal of complex derivations if instead of $G(n, m)$ we consider the following, closely related model of random graphs. Let $G^*(n, m)$ be random graph model defined by inserting into a set of $n$ vertices $m$ edges as follows: Each of the $m$ edges chooses its incident vertices u.a.r. and with replacement. After inserting all the $m$ edges, delete all the self loops and replace the multi-edges with a single edge.

The model $G^*(n, m)$ is related to the model $G(n, m)$ the following lemmas describe.
Lemma 1 Let $\mathcal{A}$ be a decreasing graph property. For $m = cn$, where $c$ is a positive fixed constant, and for $m' = m(1 - \delta)$ it holds that

$$\Pr[G(n, m') \in A] \geq \Pr[G^*(n, m) \in A] - \exp\left(-m\delta^2/17\right)$$

for any $0 < \delta < 1$.

**Proof:** Let $X$ be the number of edges in $G^*(n, m)$ after removing self loops and multi-edges. For any decreasing graph property $\mathcal{A}$, it holds that $\Pr[G^*(n, m) \in A \mid X > q] \leq \Pr[G(n, q) \in A]$. Applying the law of total probability and the previous inequality we get that for any integer $m' > 0$ it holds

$$\Pr[G^*(n, m) \in A] \leq \Pr[G(n, m') \in A] + \Pr[X < m'].$$ -

The lemma follows by deriving appropriate bounds for $\Pr[X < m']$ when $m' = (1 - \delta)m$.

Consider the process that creates $G^*(n, m)$ and let $Y$ be the number of edges that are not a self-loop or form multi-edges with some others (before the deletions). It always holds, $Y \leq X$, i.e. $Y$ is a lower bound on the number of edges in $G^*_{n, m}$ after the deletion of multiple edges and the self loops. The lemma follows by deriving an appropriate concentration result for $Y$.

Let $I_j$ be the indicator random variable such that $I_j = 1$ if the edge $j$ is placed such that it is not a self-loop or does not form multi-edge with some others, while $I_j = 0$ otherwise. It is direct to get that $E[I_j] \geq 1 - \frac{m}{n^2}$. Also, it is direct that $Y = \sum_j I_j$ and $E[Y] \geq (1 - \frac{m}{n^2})m$.

Also, we should remark that the random variables $I_j$ are negatively dependent with each other. Thus, by Theorem 7 we have that $\Pr[Y < (1 - \delta)m] \leq \exp\left(-m\delta^2/17\right)$. The lemma follows. $\diamond$

Lemma 2 Let the variable $X_k(G)$ denote the number of independent sets of the graph $G$. For $m = \Theta(n)$, it holds that

$$E[X_k(G^*_{n,m})] \sim E[X_k(G_{n,m})].$$

Finally we present a lemma that it will be very useful in the course of this paper.

**Lemma 3 (Expectation.)** Let $G^*_{n,m}$ be of expected degree $d = 2m/n$ and for real $\epsilon$ let

$$k = \frac{2n}{d} \left( \ln d - \ln \ln d + 1 - \ln 2 - \epsilon \right)$$

If $X_{n,m}(k)$ is the number of independent sets of size $k$ in $G^*_{n,m}$, then it holds that

$$E[X_{n,m}(k)] = \exp\left\{ k \left( \epsilon - \ln \left( 1 - \frac{\ln d - 1 + \ln 2 + \epsilon}{\ln d} \right) - \frac{1}{2}(k/n) - O(k/n^2 \ln d) \right) \right\}.$$ -

**Proof:** Clearly it holds

$$E[X_{n,m}(k)] = \binom{n}{k} \left( 1 - (k/n)^2 \right)^m.$$ -

Let $s = \frac{k}{n}$. We have that

$$\ln \binom{n}{k} = -n(s \ln s + (1 - s) \ln(1 - s)) + o(n)$$

$$= ns(- \ln s + 1 - s/2 - O(s^2)) + o(n)$$

$$= k \left( \ln d - \ln \ln d - \ln 2 + 1 - \ln(1 - q_d) - k_\epsilon/2n \right) + O((k/n)^2)$$

where $q_d = \frac{\ln \ln d - 1 + \ln 2 + \epsilon}{\ln d}$. Using the fact that $m = \frac{d}{4}n$ we get that

$$\ln(1 - s^2)^m = -ns(ds/2 + ds^3/4 + O(ds^5))$$

$$= -k \left( \ln d - \ln \ln d - \ln 2 + 1 - \epsilon + O(\ln d/k(n)^2) \right).$$

The lemma follows after straightforward calculations. $\diamond$

---

A graph property $\mathcal{A}$ is increasing iff given that $\mathcal{A}$ holds for a graph $G(V, E)$, then $\mathcal{A}$ holds for any $G(V, E')$: $E' \supseteq E$. 7
4 Proof of Theorem 5

Lemma 4 Suppose that there is a number \( q \in [0,1] \) such that \( \mathbb{P}_{\Lambda_k(n,m)}[E] < q \). Then
\[
|\{(G, \sigma) \in \Lambda_k(n,m) \cap E\}| < q \cdot |\Lambda_k(n,m)|.
\]

Proof: It is direct that for every \( \sigma \subset [n] \) of size \( k \), the number of graphs that have independent set \( \sigma \) is exactly \( \binom{n}{k}-\binom{n}{\lceil k/2 \rceil} \). This implies that \( \mathbb{P}_{\Lambda_k(n,m)} \) is the uniform distribution over all the members of \( \Lambda_k(n,m) \). The lemma follows. \( \diamond \)

Proof of Theorem 5: Assume that a \( \mathbb{P}_{\Lambda_k(n,m)}[E] = \zeta \), where \( E \in \Lambda_k(n,m) \). Assume further that \( \mathbb{P}_{\Lambda_k(n,m)}[E] < q \) for some \( q \in [0,1] \). Proposition 3 implies that with high probability it holds \( |S_k(G_{n,m})| \geq \frac{1}{2} \exp \left( -14n \sqrt{\ln^5 \frac{d}{d^3}} \right) E[X_{n,m}(k)] \). Thus, the assumption on \( \mathbb{P}_{\Lambda_k(n,m)}[E] \) implies that
\[
|\{(G, \sigma) \in \Lambda_k(n,m) \cap E\}| \geq \frac{\zeta}{2} \exp \left( -14n \sqrt{\ln^5 \frac{d}{d^3}} \right) |\Lambda_k(n,m)|.
\]
Combining the above inequality with Lemma 4 we get that
\[
\frac{\zeta}{2} \exp \left( -14n \sqrt{\ln^5 \frac{d}{d^3}} \right) |\Lambda_k(n,m)| \leq |\{(G, \sigma) \in \Lambda_k(n,m) \cap E\}| \leq q |\Lambda_k(n,m)|.
\]
It is direct that it should hold \( \zeta \leq 2q \exp \left( 14n \sqrt{\ln^5 \frac{d}{d^3}} \right) \). The theorem follows. \( \diamond \)

4.1 Proof of Proposition 1

Since the second moment method fails to yield a lower bound on the typical number \( |S_k(G(n,m))| \) of independent sets, we need to invent a less direct approach to prove Proposition 1. Of course, the demise of the second moment argument also presented an obstacle to Frieze in his proof that
\[
\alpha(G(n,m)) \geq (2 - \varepsilon_d)n \ln(d)/d \quad \text{w.h.p.} \tag{8}
\]
However, unlike the number \( |S_k(G(n,m))| \) of independent sets, the size \( \alpha(G(n,m)) \) of the largest one actually is concentrated about its expectation. In fact, an arsenal of large deviations inequalities applies (e.g., Azuma’s and Talagrand’s inequality), and [8] uses these to bridge the gap left by the second moment argument. Unfortunately, these large deviations inequalities draw a blank on \( |S_k(G(n,m))| \). Therefore, we are going to derive the desired lower bound on \( |S_k(G(n,m))| \) directly from (8).

To simplify our derivations we consider the model of random graphs \( G^*(n,m) \) and we show the following proposition.

Proposition 2 There is \( \varepsilon_d \to 0 \) such that for \( k < (2 - \varepsilon_d)n \ln(d)/d \) we have
\[
|S_k(G^*(n,m))| \geq E|S_k(G^*(n,m))| \cdot \exp \left( -14n \sqrt{\ln^5 \frac{d}{d^3}} \right) \quad \text{w.h.p.}
\]
Then, Proposition 1 follows by Lemma 4 for \( \delta = n^{-1/5} \), and Lemma 2.

Given some integer \( k > 0 \) and \( q \in [0,1] \), let \( X_{n,m}(k) = |S_k(G^*(n,m))| \) and let
\[
M_{k,q} = \max\{m \in \mathbb{N}|Pr[X_{n,m}(k) > 0] \geq 1 - q\}.
\]
**Lemma 5** Consider some integer $k > 0$, $q \in [0, 1]$ and some $m < M_{k,q}$. It holds that

$$\Pr \left[ X_{n,m}(k) < h^{-1} \cdot E[X_{n,m}(k)] \right] \leq 2q$$

where $h(n) > 2E[X_{n,M_{k,q}}(k)]$.

**Proof:** Consider the process where we have $n$ vertices and at each step we add an edge by selecting u.a.r., with replacement, its endpoints. It is clear that after having inserted $m$ edges and having deleted the multiple edges and selfloops, the resulting graph would be $G_{n,m}^*$.

W.l.o.g. assume that $m$, $k$, $q$ are such that $m \leq M_{k,q}$. We look the above process at two specific moments, the first is exactly after having inserted the $m$-th edge and the second is exactly after having inserted the $M_{k,q}$-th edge. Let $G_{n,m}^*$ and $G_{n,M_{k,q}}^*$ be the two graphs we see at these specific moments, after having deleted the multiple edges and selfloops, and let $X_{n,m}(k)$ and $X_{n,M_{k,q}}(k)$ be the number of independent sets of size $k$ in each graph, correspondingly. It holds that

$$E[X_{n,M_{k,q}}(k)|X_{n,m}(k)] = X_{n,m}(k)(1 - (k/n)^2)^{M_{k,q} - m}.$$ 

We let the event $E_1 = "X_{n,m}(k) < E[X_{n,m}(k)]/h(n)"$ and assume that $Pr[E_1] = \zeta$ for some $\zeta \in [0, 1]$.

By the Markov inequality we get that

$$Pr \left[ X_{n,M_{k,q}}(k) < 2E[X_{n,m}(k)]h(n)(1 - (k/n)^2)^{M_{k,q} - m} | E_1 \right] > 1/2.$$ 

The above inequality and the assumption that $Pr[E_1] = \zeta$ implies that

$$Pr \left[ X_{n,M_{k,q}}(k) < 2E[X_{n,m}(k)]h(n)(1 - (k/n)^2)^{M_{k,q} - m} \right] > \zeta/2.$$ 

Since $h(n) > 2E[X_{n,M_{k,q}}(k)]$, we get that $Pr[X_{n,M_{k,q}}(k) < 1] > \zeta/2$. It is direct that it holds $\zeta \leq 2q$. Thus, $Pr[E_1] \leq 2q$. \hfill \Box

**Definition 3** For the integer $k > 0$ let

$$\delta_k = 2(n/k) \ln(n/k) + 2(n/k) - O(\ln(n/k)\sqrt{n/k})$$

Improving on the main Theorem in \[8\] we get the following theorem.

**Theorem 8** For the integer $k$ and $G^*(n, m)$ of expected degree $d \leq \delta_k$ it holds that

$$Pr[\alpha(G^*(n, m)) < k] < \exp \left( -n \ln^2(d/d^2) \right).$$

Also, for $d = \delta_k$ it holds that $E[|S_k(G^*_{n,m})|] \leq \exp \left( 14n \ln^5(d/d^3) \right)$.

The proof of Theorem \[8\] appears in Section \[4.2\]. Proposition \[2\] follows by using Lemma \[5\] and taking $M_{k,q} = \frac{\delta_k n}{2}$, the values of $q$ and $h(n)$ follow from Theorem \[8\].

Recently, V. Dani and C. Moore \[6\] provided an improvement on the result of Theorem \[8\]. In particular, they showed that $G^*_{n,m}$ of expected degree $d \leq 2(n/k) \ln(n/k) + 2(n/k) - O(\sqrt{n/k})$ has an independent set of size $k$ w.h.p. This new result can be used to improve slightly Theorem \[5\].
4.2 Proof of Theorem

Definition 4 For $G^*(n, m)$ of average degree $d$ and for $\epsilon \in \mathbb{R}$ let $k_\epsilon$ denote the size of independent sets such that

$$\lim_{n \to \infty} \ln E[X_{n,m}(k_\epsilon)]/k_\epsilon = \epsilon$$

We should remark that setting $f_d(s) = \lim_{n \to \infty} \ln E[X_{n,m}(sn)]/(sn)$, where $s \in (0, 1)$ it is direct to show that for a given $d$ the function $f_d(s)$ is decreasing. In particular it holds that

$$\frac{d}{ds} f(s) \leq -\left(\frac{d}{2(1-s^2)} + 1/s\right).$$

Thus for $\epsilon < \epsilon'$ we get that $k_\epsilon > k_{\epsilon'}$.

Lemma 6 Consider the graph $G^*_{n,m}$ of expected degree $d = 2m/n$. For $\epsilon \geq 7\sqrt{\frac{\ln^3 d}{d}}$ it holds that

$$P[\alpha(G^*_{n,m}) \geq k_\epsilon] \geq 1 - \exp\left(-\epsilon^2 \frac{n}{49d \ln d}\right).$$

The proof of Lemma 4 appears in Section 4.3. All the above imply that for $G^*_{n,m}$ of expected degree $d$ the size of the maximum independent set is w.h.p. at least $k_{cr} = k(7\sqrt{\ln^3 d/d})$ and

$$E[X_{n,m}(k_{cr})] \leq \exp\left(14n\sqrt{\ln^5 d/d^5}\right).$$

Theorem 8 follows from Lemma 7.

Lemma 7 The graph $G^*_{n,m}$ of expected degree $d$ with probability at least $1 - \exp\left(-n\ln^2 d/d^2\right)$ has an independent set of size $k$ if

$$d \leq 2(n/k) \ln(n/k) + 2(n/k) - O(\ln(n/k)\sqrt{\ln(n/k)})$$

Proof: For given $k$, let $s = k/n$. In the derivations below assume that $s$ is sufficiently small constant. It is direct to show that

$$\frac{1}{n} \ln E[X_{n,m}(k)] = -h(s) + d \ln(1 - s^2)/2 - (1 - o(1)) \ln n/2n,$$

where $h(x) = x \ln x + (1 - x) \ln(1 - x)$, for $x \in [0, 1]$. Let $d^+ = d^+(k)$ be such that

$$-h(s) + d^+ \ln(1 - s^2)/2 = 0.$$

It is direct to get that

$$d^+ = 2\frac{h(s)}{\ln(1 - s^2)} = 2s^{-1}\ln s^{-1} + 2s^{-1} - 1 - O(s \ln s^{-1}).$$

Assume that we have a graph $G^*(n, m')$ of expected degree $d_k = 2s^{-1}\ln s^{-1} + 2(s^{-1} - C \ln s^{-1}/\sqrt{s}$, where $C$ is a sufficiently large constant. It is easy to check that it also holds that $k = 2n(1 - o_{d_k}(1)) \ln d_k/d_k$, where $o_{d_k}(1)$ is a quantity that tends to zero as $d_k$ increases.

It holds that $d^+ > d_k$. Add to $G^*(n, m')$ edges at random so as to increase the expected degree to $d^+$, i.e. we insert $(d^+ - d_k)n/2$ edges at random. Each independent set in $G^*(n, m')$ is also an independent set of $G^*(n, m)$ with probability $(1 - s^2)^{(d^+ - d_k)n/2}$. Then it is easy to get that

$$E[X_{n,m}(k)] = (1 - s^2)^{(d^+ - d_k)n/2} E[X_{n,m'}(k)].$$
The proof the claim appears after the proof of this lemma. Furthermore, we show that

Let \( b_i \) decrease with \( i \) because of the factor \((i+1)\) in the dominator. For intermediate \( i \) it grows due to the factor \((1 + q_i (2i + 1)/n^2)) \). Finally, when the difference \( k - i \) becomes small, \( b_i \) declines. Thus, \( a_i \) achieve its largest value either at \( i_1 = \min\{i \geq 1| b_i < 1\} \), or at \( i_2 = \max\{i < k| b_i > 1\} + 1 \).

Claim 1 Let \( s = k/n \), then \( i_1 \in [sk, sk(1 + s \ln^{3/2} d)] \).

The proof the claim appears after the proof of this lemma. Furthermore, we show that \( i_2 \) has a lower bound \( i_2' = \left\lfloor k (1 - \sqrt{k/n}) \right\rfloor \) (obviously \( i_2 \leq k \)). More specifically, it holds that

\[
\frac{1}{n} \ln E[X_{n,m}(k)] = \ln(1 - s^2)(d_k - d^-)/2 - o(1)
\geq -\frac{s^2}{1 - s^2}(d_k - d^-)/2 \geq \frac{2C'}{5} s^{3/2} \ln s^{-1}.
\]

The lemma follows by using the fact that \( s = 2(1 - o_d(1)) \ln d/d \) and by taking sufficiently large \( C \). \( \Diamond \)

4.3 Proof of Lemma 6

Lemma 8 (Second Moment Bound.) For the graph \( G^*_{n,m} \) with sufficiently large expected degree \( d = 2m/n \) and \( \epsilon \geq 3\sqrt{\ln^3 d/d} \) it holds that

\[
Pr[X_{n,m}(k) > 0] \geq \exp \left( -\frac{57 \ln^{11/2} d}{d^3} k_e \right) .
\]

Proof: It holds that

\[
Pr[X_{n,m}(k) > 0] \geq \frac{E^2[X_{n,m}(k_e)]}{E[X_{n,m}(k_e)]^2} .
\]

The lemma will follow by bounding appropriately the r.h.s. of the above inequality. Note first that for such \( \epsilon \) it holds that \( E[X_{n,m}(k_e)] = \exp \left( 6n \sqrt{\ln^5 d/d^2} \right) \). When there is no danger of confusion, we abbreviate \( k_e \) to \( k \). We have that

\[
\frac{E[X_{n,m}^2(k)]}{E^2[X_{n,m}(k)]} = \sum_{i=0}^{k} \binom{k}{i} \binom{n-i}{k-i} (1 - 2(k/n)^2 + (i/n)^2)^m \frac{a_i}{\binom{n}{k} (1 - (k/n)^2)^{2m}} = \sum_{i=0}^{k} a_i
\]

where

\[
a_i = \binom{k}{i} \binom{n-k}{k-i} (1 - 2(k/n)^2 + (i/n)^2)^m \binom{n}{k} (1 - (k/n)^2)^{2m} .
\]

Let

\[
b_i = \frac{a_{i+1}}{a_i} = \left( \frac{k-i}{i+1}(n-2k+i+1) \right) \left( 1 + q_i \frac{2i+1}{n^2} \right)^m
\]

where \( q_i = (1 + \Theta(\ln^2 d/d^2)) \geq 1 \) is a quantity that tends to 1 with \( d \).

It is not hard to see that for small \( i \), the sequence \( b_i \) decreases with \( i \) because of the factor \((i+1)\) in the dominator. For intermediate \( i \) it grows due to the factor \((1 + q_i (2i + 1)/n^2)) \). Finally, when the difference \( k - i \) becomes small, \( b_i \) declines. Thus, \( a_i \) achieve its largest value either at \( i_1 = \min\{i \geq 1| b_i < 1\} \), or at \( i_2 = \max\{i < k| b_i > 1\} + 1 \).

Claim 1 Let \( s = k/n \), then \( i_1 \in [sk, sk(1 + s \ln^{3/2} d)] \).

The proof the claim appears after the proof of this lemma. Furthermore, we show that \( i_2 \) has a lower bound \( i_2' = \left\lfloor k (1 - \sqrt{k/n}) \right\rfloor \) (obviously \( i_2 \leq k \)). More specifically, it holds that

\[
b_{i_2 - 1} \geq \frac{(k/n)^2 \exp \left( m \frac{2i_2 q_i}{n^2} + O(m(2i_2)^2/n^4) \right)}{3 \ln d/d^2 \exp (2 \ln d - 2 \ln \ln d + 2 - 2 \ln 2 - o_d(1))} \geq \frac{3 \ln d/d^2 \exp (2 \ln d - 2 \ln \ln d + 0.6)}{1}.
\]
Thus, \( i_2 \geq i'_2 \). As far as \( a_{i_2} \) is regarded we get the following:

\[
a_{i_2} = \frac{n-k}{(k-n)^2} \left( \frac{1 - (2n)^2 + (i_2/n)^2}{(1 - (k/n)^2)^m} \right) \\
\leq \frac{(k)_{(k-n)^2}}{E[X_{n,m}(k)]} \\
\leq \left( \frac{e^n n^k}{(k-n)^2} \right)^{k-i_2} \exp \left( -6 \ln^2 d \right) \\
\leq \left( \frac{e^n n^k}{(k-n)^2} \exp \left( -6 \ln^2 d \right) \right)^{k-i_2} \\
\leq \left( \frac{e^n n^k}{(k-n)^2} \exp \left( -2.12 \ln d \right) \right)^{(k/n)^3/2}. 
\]

It is direct that for sufficiently large \( d \), it hold that \( a_{i_2} \leq o(n^{-1}) \). Thus, the main contribution to \( \sum i_a \) comes from the terms with indices around \( i_1 \) while the contribution of the terms with indices around \( i_2 \) is negligible. For \( s = k/n \), it is direct that \((i_1/n)^2 \leq s^4(1 + 3s \ln 3/2 d) \). It holds that

\[
a_{i_1} \leq \left( \frac{1-2s^2+s^2(1+3s \ln 3/2 d)}{(1-s^2)^2} \right)^{m} \\
= \left( 1 + \frac{3s^5 \ln 3/2 d}{(1-s^2)^2} \right)^{dn/2} \\
\leq \exp \left( 3.5s^5 d \ln 3/2 \right) \leq \exp \left( 56 \ln 11/2 d \right). 
\]

The lemma follows by noting that

\[
E[X_{n,m}(k)] \\
\leq k \exp \left( 56 \ln 11/2 d \right) \leq \exp \left( 57 \ln 11/2 d \right). 
\]

\[\Box\]

**Proof of Claim** \[\text{[1]}\] For \( x \in (0, 1) \) let \( H(x) = x \ln x + (1-x) \ln(1-x) \). Also, let \( k_{\pm} = sn i = ak \) and \( m = \frac{1}{2n} \), for \( 0 \leq s, a \leq 1 \) fixed real.

\[
f(a) = \frac{1}{n} \ln \left\{ \binom{n}{k} \binom{n-k}{k-i} (1 - (2n)^2 + (i/n)^2)^m \right\} \\
- H(s) - H(a)s - H(s(1-a)/(1-s))(1-s) + \frac{1}{2} \ln (1 - 2(s)^2 + (as)^2) + o(1). 
\]

Taking the derivative of the function \( f \), above, we get

\[
\frac{d}{da} f(a) = s \left( \ln \frac{1-a}{a} + \ln \frac{s(1-a)}{1-s} - \ln \left( 1 - \frac{s(1-a)}{1-s} \right) \right) + \frac{d\ln s}{1 - 2s^2 + (as)^2}. 
\]

Note that \( f'(s) = \frac{rs^3}{1 - 2s^2 + s^2} > 0 \). For \( a_0 = s + t \) where \( t = s^2 \ln 3/2 d \), it holds that \( \ln(a_0) \geq \ln s + t/(2s) \), \( \ln(1 - a_0) \leq 0 \), \( \ln \frac{1-a_0}{1-s} \leq \ln s \) and \( \ln (1 - \frac{s(1-a)}{1-s}) \geq 2s \). Also, it is direct that \( \frac{d\ln s}{1 - 2s^2 + (as)^2} \leq \frac{3}{2} ds^3 \). Thus, for sufficiently large \( d \), we get that

\[
f'(a_0) \leq -t/2 + 2s^2 + 3r s^3 \leq -2 \frac{\ln 7/2 d}{d^2} + \ln 3d \frac{d^3}{d^2} < 0. 
\]
By the fact that $f'$ is continuous in the interval $(0,1)$ and $f(s)f(a_0) < 0$, there exists at least one $a' \in (s, s + \ln^{3/2} ds^2)$ such that $f(a') = 0$. Also, it is direct that $f(a) > 0$, for $a < a'$, and $f(a) < 0$, for $a > a'$. This implies that there is exactly one point in this interval where $f(a) = 0$. ◊

**Proof of Lemma 6** We work in the same manner as in the proof of Theorem 7.4 in [22], i.e. we apply the Talagrand’s large deviation inequality. Let $\epsilon_2 = \frac{3\ln^{3/2} d}{d^{1/2}}$ and $\epsilon_1 \geq \frac{7\ln^{3/2} d}{d^{1/2}}$. Using Lemma 5 it is direct to show that $k_{r_2} = 2n(1 - a_d(1)) \ln d/d$. It holds that

$$Pr(\alpha(G^*_{n,m}) < k_{r_1} - 1)Pr(\alpha(G^*_{n,m}) \geq k_{r_2}) \leq \exp (- (k_{r_2} - k_{r_1} + 1)^2 / 4k_{r_2})$$

(9)

It is easy to show that $k_{r_2} - k_{r_1}$ is not too small. In particular, setting $r_1 = 7 \sqrt{\ln^3 d/d}$ and $r_2 = \epsilon_2$, it is easy to see that $k_{r_2} - k_{r_1} > k_{r_2} - k_{r_1}$.

**Claim 2** Let $r_1 = \frac{7\ln^{3/2} d}{d^{1/2}}$ and $r_2 = \frac{3\ln^{3/2} d}{d^{1/2}}$. It holds that

$$k_{r_2} - k_{r_1} \geq \frac{4r_1 - 2r_2}{d}n.$$

The proof of the claim appears after this proof. Using the claim above and (9) we get

$$Pr(\alpha(G^*_{n,m}) < k_{r_1} - 1)Pr(\alpha(G^*_{n,m}) \geq k_{r_2}) \leq \exp \left( - \frac{(4(\epsilon_2 - 2\epsilon_1) n/d)^2}{8n \ln d/d} \right) = \exp \left( - \frac{2(\epsilon_2 - 2\epsilon_1)^2 n}{d \ln d} \right).$$

The theorem follows by combining Lemma 8 with the above inequality to get that

$$Pr(\alpha(G^*_{n,m}) < k_{r_1} - 1) \leq \exp \left( - \frac{2\epsilon_1^2}{49d \ln d} n + 160 \ln^{6.5} d / d^4 n \right) \leq \exp \left( - \frac{\epsilon_1^2}{49d \ln d} n \right).

◊

**Proof of Claim 2** Lemma 5 suggest that $k_{r_1}, k_{r_2} > 1.8 \ln d / d$. Let $f(s) = \lim_{n \to \infty} \frac{1}{n} \ln E[X_{n,m}(ns)]$ for $s \in (0,1)$. $f(s)$ is continuous in $(0,1)$. For convenience we let $s_1 = k_{r_1} / n$ and $s_2 = k_{r_2} / n$. By the Mean Value Theorem, there exists $\bar{s} \in [s_1, s_2]$ such that

$$f(s_2) - f(s_1) = f'(\bar{s}).$$

It holds that

$$\frac{d}{ds} f(s) = \frac{d}{ds} \left( - s \ln s - (1 - s) \ln(1 - s) + \frac{d}{2} \ln(1 - s^2) \right) = \frac{d}{\ln s} + \ln(1 - s) - d \frac{s}{1 - s}.$$

For sufficiently large $d$, there exists $x \in [1.8, 2]$ such that $\bar{s} = x \ln d / d$. This implies that $-2 \ln d \leq f(\bar{s}) \leq - \frac{0.6}{2} \ln d$. Using the initial equality, we get that

$$s_2 - s_1 = \frac{f(s_2) - f(s_1)}{f'(\bar{s})} \geq \frac{2r_1 s_1 - r_2 s_2}{\ln d} \geq 2s_1 \frac{r_1 - 2r_2}{0.9 \ln d} \geq 4 \frac{(r_1 - 2r_2)}{d}.$$

The penultimate inequality follows from that fact that $k_{r_1}, k_{r_2}$ are very close to each other, which implies $k_{r_2} < 2k_{r_1}$. For the last inequality we the fact that $s_1 > 1.8 \ln d / d$. The claim follows. ◊

13
5 Proof of Theorem 1

Instead of the random graph model $G(n, m)$ we consider the model $G(n, p)$, where $p = d/n$ for fixed real $d$ and we prove the following theorem.

**Theorem 9** There is $\varepsilon_d \to 0$ such that $S_k(G(n, m))$ is $O(1)$-connected w.h.p. for any

$$ k \leq (1 - \varepsilon_d) \frac{\ln d}{d} \cdot n. $$

Theorem 1 follows by using standard arguments (see Proposition 1.13 in [22]).

**Remark.** For simplicity, we assume that for two adjacent independent sets their Hamming distance is less than $20d$.

**Definition 5** For every vertex $u$ in $G_{np}$ we let $N_u$ denote the set vertices which are adjacent to $u$.

A sufficient condition for establishing the connectivity of $S_k(G_{n,d/n})$ is requiring this space to have the so called Property $\Gamma$:

**Property $\Gamma$.** For any two (arbitrary) $\sigma, \tau \in S_k(G_{n,d/n})$ we can find chains $\sigma, \sigma', \sigma''$ and $\tau, \tau', \tau''$ of independent sets in $S_k(G_{n,d/n})$ and $S_{k+1}(G_{n,d/n})$ connected as in Figure 1. Furthermore, we have that $\sigma'', \tau'' \in S_k(G_{n,d/n})$ and $\operatorname{dist}(\sigma'', \tau'') < \operatorname{dist}(\sigma, \tau)$. In particular it holds that $|\sigma'' \cap \tau''| = |\sigma \cap \tau| + 1$.

**Theorem 2** will follow by showing that with probability $1 - o(1)$ $S_k(G_{n,d/n})$ has Property $\Gamma$. For two independent sets $\sigma, \tau \in S_k(G_{n,d/n})$, we introduce the notion of “augmenting vertex”, which is very closely related to Property $\Gamma$.

**Augmenting vertex.** For the pair $\sigma, \tau \in S_k(G_{n,d/n})$ the vertex $v \in V \setminus (\sigma \cup \tau)$ is augmenting if one of the the following $A, B$ holds.

**A.** $N_v \cap (\sigma \cup \tau) = \emptyset$

**B.** $N_v \cap \sigma \cap \tau = \emptyset$ and there are the terminal sets $I_v(\sigma)$ and $I_v(\tau)$ of size at most $7d$ such that

- $N_v \cap I_v(\sigma) = \emptyset$
- $I_v(\sigma)$ is an independent set of $G_{n,d/n}$
- $|I_v(\sigma)| = |N_v \cap \sigma|$
- $\forall w \in I_v(\sigma)$ it holds that $|N_w \cap \sigma| = 1$ and $|N_w \cap N_v \cap \sigma| = 1$

The corresponding conditions should hold for $I_v(\tau)$, too.
Figure 3 shows an example of a pair of independent sets where the vertex $v$ is an augmenting vertex.

The notion of “augmenting vertex” and the Property $\Gamma$ are closely related since we will show that for a pair $\sigma, \tau \in S_k(G_{n,d/n})$ that has an augmenting vertex $v$ we can find short chains $\sigma, \sigma', \sigma''$ and $\tau, \tau', \tau''$ as in Figure 1. Thus, if we can find an augmenting vertex for any two members of $S_k(G_{n,d/n})$, then $S_k(G_{n,d/n})$ has Property $\Gamma$.

In what follows we introduce the process called Collider which takes as an input $\sigma, \tau$ and the augmenting vertex $v$ and returns the independent sets $\sigma''$ and $\tau''$ as in Figure 1. Thus, if we can find an augmenting vertex for any two members of $S_k(G_{n,d/n})$, then $S_k(G_{n,d/n})$ has Property $\Gamma$.

Collider $(\sigma, \tau, v)$: Let $V_{\sigma}, V_{\tau}$ be the set of vertices in $\sigma$ and $\tau$, correspondingly.

Phase 1.
1. Remove from $V_{\sigma}$ all the vertices in $N_v \cap \sigma$.
2. Insert into $V_{\sigma}$ the vertex $v$ and the vertices in $I_v(\sigma)$.
3. Do the same for $V_{\tau}$.

Phase 2.
1. Remove some vertex from $V_{\sigma}$ that is nor $v$ neither some vertex in $\sigma \cap \tau$.
2. Do the same for $V_{\tau}$.
3. Return the sets $V_{\sigma}$ and $V_{\tau}$.

End

Figure 3 shows the changes that have taken place to the independent sets in Figure 2 at the end of “Phase 1” of Collider$(\sigma, \tau, v)$. The independent sets $\sigma'$ and $\tau'$ in the short chains correspond to the vertex sets $V_{\sigma}$ and $V_{\tau}$ at the end of the “Phase 1”. Both $\sigma', \tau'$ now contain the augmenting vertex $v$, moreover, it holds that $\sigma' \cap \tau' = (\sigma \cap \tau) \cup \{v\}$. The independent sets $\sigma''$ and $\tau''$ correspond to the sets $V_{\sigma}$ and $V_{\tau}$ that are returned by Collider$(\sigma, \tau, v)$. After “Phase 2”, the independent sets in Figure 3 are transformed to those in Figure 4 There the vertices $u_2$ and $u_7$ are removed from $\sigma'$ and $\tau'$, correspondingly.

In the following lemma we show all the above properties of Collider.

**Lemma 9** Let $\sigma, \tau \in S_k(G)$ with augmenting vertex $v$. Let $\sigma''$ and $\tau''$ be the two sets of vertices that are returned from Collider$(\sigma, \tau, v)$. The two sets have the following properties:

1. $\sigma'', \tau'' \in S_k(G)$.
2. $|\sigma'' \cap \tau''| = |\sigma \cap \tau| + 1$,
3. There are $\sigma', \tau' \in S_{k+1}(G)$ such that $\sigma'$ is adjacent to both $\sigma$ and $\sigma''$ while $\tau'$ is adjacent to both $\tau$ and $\tau''$.

If $N_v \cap (V_{\sigma} \cup V_{\tau}) = \emptyset$ then we set $I_v(\sigma) = \emptyset$
Proof: First we show that \( V_\sigma \) and \( V_\tau \), as returned by \textit{Collider} \( (\sigma, \tau, v) \), correspond to independent sets of \( G \). The same arguments apply for both \( V_\sigma \) and \( V_\tau \), we consider only the case of \( V_\sigma \).

Let \( v \) be an augmenting vertex for the pair \( \sigma, \tau \). If \( N_v \cap (\sigma \cup \tau) = \emptyset \), then it is direct that the set \( V_\sigma \), at the end of the process, is an independent set of \( G \). Consider, now, the case where \( N_v \cap (\sigma \cup \tau) \neq \emptyset \). Assume that \( V_\tau \), at the end of the process, does not correspond to an independent set of \( G \), i.e., there is an edge between two vertices in \( V_\tau \). Then, the edge must be either between the new vertices that are inserted in \( V_\sigma \) (i.e. \( v \) and \( I_v(\sigma) \)) or between some newly inserted vertex and an old one.

The first case cannot be true since the assumption that \( v \) is an augmenting vertex implies that there is no edge connecting \( v \) with \( I_v(\sigma) \), or vertices in \( I_v(\sigma) \) with each other. As far as the second case is considered note that both \( v \) and \( I_v(\sigma) \) have the same neighbours in \( \sigma \). During the process \textit{Collider}(\( \sigma, \tau, v \)) all the vertices in \( V_\sigma \) that are adjacent to \( v \) and \( I_v(\sigma) \) are removed. The second case cannot occur. Thus the sets \( V_\sigma \) and \( V_\tau \), as returned by the process \textit{Collider}(\( \sigma, \tau, v \)), correspond to independent sets of \( G \).

It is direct to show that after the execution of \textit{Collider}(\( \sigma, \tau, v \)) the sets \( V_\sigma \) and \( V_\tau \) are both of size \( k \). Property 1 follows.

Property 2 follows by noting that \( |\sigma'' \cap \tau''| = |\sigma \cap \tau \cup \{v\} \). As far as the property 3 is regarded it is easy to see that \( \sigma' \) corresponds to the set \( V_\sigma \) exactly at the end of “Phase 1”. We argue similarly for \( \tau' \).

We are going to use the first moment method to show that with probability \( 1 - o(1) \), the graph \( G_{n,d/n} \) has no pair of independent sets in \( S_k(G_{n,d/n}) \) with no augmenting vertex. This implies that with probability \( 1 - o(1) \) the set \( S_k(G_{n,d/n}) \) has Property \( \Gamma \) and Theorem \( \ref{thm:main} \) follows.

We compute, first, the probability for a pair in \( S_k(G_{n,d/n}) \) to have an augmenting vertex.

**Proposition 3** For \( \sigma, \tau \in S_k(G_{n,d/n}) \) and \( |\sigma \cap \tau| = i \leq k \), let \( p_{k,i} \) be the probability for this pair to have an augmenting vertex. There is \( \epsilon_d \to 0 \) such that for any \( \epsilon_d \leq \epsilon \leq 1 - \epsilon_d \) and \( k = (1 - \epsilon)\frac{\ln d}{d} \) the following is true

\[
p_{k,i} \geq 1 - 2 \exp \left( -0.8 \frac{d^i}{d} \right).
\]

The proof of Proposition \( \ref{prop:augmenting} \) appears in Section \( \ref{sec:proof_propositions} \).

**Proof of Theorem \( \ref{thm:main} \)** Let \( Z_k \) be the number of pairs of independent sets of size \( k \) in \( G_{n,p} \) that do not have augmenting vertex. Theorem follows by showing that \( E[Z_k] = \exp \left( -0.5 \frac{d^i}{d} n \right) \) and then by applying the Markov inequality. It holds that

\[
E[Z_k] = \sum_{i=0}^{k} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} (1-p)^{2(i-k)} (1-p) (1-p_{k,i})
\]

\[
\leq 2 \exp \left( -0.8 \frac{d^i}{d} n \right) \sum_{i=0}^{k} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} (1-p)^{2(i-k)} (1-p)
\]

where the last inequality follows by Proposition \( \ref{prop:augmenting} \). Let \( X_k \) be the number of independent sets of size \( k \) in \( G_{n,d/n} \). It is direct that

\[
E[Z_k] \leq 2 \exp \left( -0.8 d^i n/d \right) E[X_k^2].
\]

By the relation in (7.10) in \( \ref{levitin} \) (page 182) we have that

\[
E[X_k^2] \leq (E[X_k])^2 k \exp \left( k/2 \ln^3 d \right)
\]

\[
\leq \exp \left( 4n \ln^2 d/d \right) k \exp \left( n/d \ln^2 d \right) \leq \exp \left( 5n \ln^2 d/d \right).
\]

It is clear that for sufficiently large \( d \), and \( 3 \ln \ln d \leq \epsilon \leq 1 \) we have \( E[Z_k] \leq \exp \left( -0.5 \frac{d^i}{d} n \right) \). The theorem follows.
5.1 Proof of Proposition 3

Given an arbitrary pair \( \sigma, \tau \in S_k(G_{n,d/n}) \), we are going to construct an appropriate set \( Q_0 \) containing vertices that each of them can be used as an augmenting vertex of the pair \( \sigma, \tau \). Proposition 3 follows by deriving an appropriate lower bound for \( Pr[|Q_0| > 0] \). First, we need to define some sets of vertices:

- **Q1(σ)** \( Q_1(\sigma) \subseteq V \backslash (\sigma \cup \tau) \) contains those vertices that have with exactly one neighbour in \( \sigma \) and they do not have exactly one neighbour in \( \tau \). For the augmenting vertex \( v \in Q_0 \), we require that the set of terminal vertices \( I_u(\sigma) \) belong to \( Q_1(\sigma) \).

- **Q2(σ)** \( Q_2(\sigma) \subseteq \sigma \backslash \tau \) is the set of vertices that have at least one neighbour in \( Q_1(\sigma) \). Since we have required that for any \( v \in Q_0 \) the associated terminal set belongs to \( Q_1(\sigma) \), then it is clear that only vertices in \( Q_2(\sigma) \) can be in \( N_v \cap \sigma \).

- **Q3(σ)** Let \( Q_3(\sigma) \subseteq V \backslash (\sigma \cup \tau \cup Q_1(\sigma)) \) be the set with the following three properties:
  
  - \( S_1 \)- \( \forall w \in Q_3(\sigma) \) it holds \( N_w \cap \sigma \subseteq Q_2(\sigma) \).
  - \( S_2 \)- \( |N_w \cap \sigma| \leq 7d \).
  - \( S_3 \)- For any two \( x, y \in Q_3(\sigma) \) it holds that \( |N_x \cap N_y \cap \sigma| \leq 1 \).

In an analogous manner we define \( Q_1(\tau), Q_2(\tau) \) and \( Q_3(\tau) \), w.r.t. the independent set \( \tau \).

If any \( u \in Q_3(\sigma) \cap Q_3(\tau) \) does not have any neighbours in \( \sigma \cap \tau \) and we can find appropriate terminal sets \( I_u(\sigma), I_u(\tau) \), then \( u \) is an augmenting vertex. \( Q_0 \) will consist of all the augmenting vertices in \( Q_3(\sigma) \cap Q_3(\tau) \).

So as to derive a lower bound on \( Pr[|Q_0| > 0] \), first we provide a concentration result for \( Q_3(\sigma) \cap Q_3(\tau) \). Then, we provide a concentration result on the number of vertices in \( Q_3(\sigma) \cap Q_3(\tau) \) that are augmenting.

**Claim 3** Let \( X_1 = |Q_1(\sigma)| \). It holds that \( E[X_1] = \frac{(1 - \epsilon) \ln d}{d^3} n(1 - o_d(1)) + O(1) \) and

\[
Pr \left[ |X_1 - E[X_1]| \geq \Theta(n^{3/4}) \right] \leq 2 \exp \left( -n^{1/2} \right).
\]

The proof of Claim 3 appears in Section 5.2.

**Lemma 10** For \( u \in V \backslash (\sigma \cup \tau \cup Q_1(\sigma)) \), it holds that

\[
Pr[|u \in Q_3(\sigma)|] \geq 1 - \frac{(1 - a)(1 - \epsilon) \ln d}{e^{\ln d}} - 3 d^6 \ln d - O(n^{-1}).
\]

The proof of Lemma 10 appears in Section 5.3.

For \( u \in V \backslash (\sigma \cup \tau \cup Q_1(\sigma) \cup Q_1(\tau)) \) let \( J_u \) be the indicator random variable such that \( J_u = 1 \) if \( u \in Q_3(\sigma) \cap Q_3(\tau) \) and \( J_u = 0 \) otherwise. It is easy to verify that \( E[J_u] = Pr[|u \in Q_3(\sigma)|^2 \) for every \( u \in V \backslash (\sigma \cup \tau \cup Q_1(\sigma) \cup Q_1(\tau)) \). Note that the indicator random variables \( J_u \) are negatively correlated with each other. More specifically, due to condition \( S_3 \) it holds that

\[
Pr[J_x = 1 | J_y = 1] \leq Pr[J_x = 1]
\]

since the event \( J_y = 1 \), can reduce the choices of \( x \) for neighbours in \( Q_2(\sigma) \) and \( Q_2(\tau) \). Let \( X_3 = \sum_u J_u \). Using Claim 3 and Lemma 10 we get that

\[
E[X_3] \geq (1 - o_d(1)) n |P[v \in Q_3(\sigma)]|^2 \geq (1 - o_d(1)) n,
\]

for \( v \in V \backslash (\sigma \cup \tau \cup Q_1(\sigma) \cup Q_1(\tau)) \).

\[\text{Note that it always hold that } Q_1(\sigma) \cap Q_1(\tau) = \emptyset.\]
Since $X_3$ is a sum of negatively depended random variables, Theorem $\square$ suggests that
\[ \Pr[X_3 < 0.8n] \leq \exp\left(-3.8n/100\right). \] (10)

So far we have shown that the size of $Q_3(\sigma) \cap Q_3(\tau)$ is large w.h.p. The proposition will follow by showing that many of the vertices in $Q_3(\sigma) \cap Q_3(\tau)$ are augmenting. In particular, every $u \in Q_3(\sigma) \cap Q_3(\tau)$ is augmenting, i.e. belongs to $Q_0$, if the following two conditions hold:

$Z_1$. $u$ does not have edge connecting it to $\sigma \cap \tau$.

$Z_2$. there exist $R_u(\sigma) \subseteq Q_1(\sigma)$, $R_u(\tau) \subset Q_1(\tau)$ which can be used as the terminal sets $I_u(\sigma)$ and $I_u(\tau)$, correspondingly.

**Claim 4** Each vertex in $Q_3(\sigma) \cap Q_3(\tau)$ satisfies $Z_2$, independently of the others with probability
\[ p_4 \geq 1 - 50d^3/n. \]

The proof of Claim $\square$ appears in Section 5.2.

Let $X_4$ be the number of vertices in $Q_3(\sigma) \cap Q_3(\tau)$ that satisfy condition $Z_2$. Using (10) and Chernoff bounds we get
\[ \Pr[X_4 \leq 0.7n] \leq \exp\left(-n/1000\right). \] (11)

Assume that $|\sigma \cap \tau| = \alpha \cdot k$, for some $\alpha \in [0, 1]$. It is easy to show that each vertex in $Q_3(\sigma) \cap Q_3(\tau)$ that satisfies $Z_2$, independently of the other vertices satisfies $Z_1$ with probability $d^{-a(1-\epsilon)} + O(n^{-1}).$

We let $B=\"X_4 \geq 0.7n\"$, it is direct that $E[|Q_0||B| \geq 0.7nd^{-a(1-\epsilon)} - O(1)$. Since $a \in [0, 1]$, there exists $\epsilon' = \epsilon'(\epsilon, a) > \epsilon$ such that $a(1-\epsilon) = 1 - \epsilon'$. Applying the Chernoff bounds for $Y$ we get that
\[ \Pr[|Q_0| \leq 0|B] \leq \exp\left(-0.7d^{\epsilon'} n/d\right) \leq \exp\left(-0.8d^{\epsilon'} n/d\right). \]

By the law of total probability and (11) we get $\Pr[|Q_0| \leq 0] \leq \exp\left(-0.8d^{\epsilon'} n\right)$. The proposition follows.

5.2 Proof of claims $\square$

**Proof of Claim $\square$** The probability for a vertex to be in $Q_1(\sigma)$ is
\[ p_1 = kp(1-p)^{k-1}(1-kp(1-p))^{k-1} = \frac{(1-\epsilon)\ln d}{d^{1-\epsilon}} \left(1 - \frac{(1-\epsilon)\ln d}{d^{1-\epsilon}} - O(n^{-1})\right). \]

It is direct that
\[ E[X_1] = (1-2s)np_1 = \frac{(1-\epsilon)\ln d}{d^{1-\epsilon}} n(1-o_d(1)) + O(1). \]

Applying the Chernoff bounds we get the claim. $\diamond$

**Proof of Claim $\square$** For some vertex $u \in Q_3(\sigma) \cap Q_3(\tau)$ there is at least one set $R_u(\sigma) \subseteq Q_1(\sigma)$ such that the following is true: Every vertex in $N_u \cap \sigma$ has exactly one neighbour in $R_u(\sigma)$. By definition it also holds that every vertex in $R_u(\sigma)$ has exactly one neighbour in $N_u \cap \sigma$.

Assuming that the condition $Z_1$ holds for $u$, then $R_u(\sigma)$ can be the terminal set $I_u(\sigma)$ iff no two vertices in $R_u(\sigma)$ are connected with each other and no vertex in $R_u(\sigma)$ is connected to $u$.

Since each neighbour of $u$ in $Q_3(\sigma)$ can have many neighbours in $Q_1(\sigma)$, possibly, there can by many choices for the set $R_u(\sigma)$. It is direct that the probability for $u$ to fulfil condition $Z_2$ is minimized when there is only one choice for $R_u(\sigma)$ and this set has the maximum possible size.
Since the maximum number of neighbours of \( u \in Q_3(\sigma) \cap Q_3(\tau) \) in \( Q_2(\sigma) \) is at most \( 7d \) we have that \( |R_u(\sigma)| \leq 7d \). It holds that

\[
Pr[\text{there exists edge with both ends in } R_u(\sigma)] \leq \left( \frac{7d}{2} \right)^2 \frac{d}{n} \leq 49d^3/2n,
\]

\[
Pr[u \text{ is adjacent to some vertex in } R_u(\sigma)] \leq 7d^2/n.
\]

The same holds for \( R_u(\tau) \). It is direct that \( 1 - p_u \leq 2 (49d^3/(2n) + 7d^2/n) \leq 50d^3/n \).

It remains to show that each vertex in \( Q_3(\sigma) \cap Q_3(\tau) \) satisfies \( Z_2 \) independently of the others. Note that we have the restriction that no two vertices in \( Q_3(\sigma) \) (in \( Q_3(\tau) \)) share more than one neighbour in \( Q_2(\sigma) \) (\( Q_2(\tau) \)). Thus, each edge between any two vertices in \( Q_1(\sigma) \) (or in \( Q_1(\tau) \)) eliminates at most one vertex in \( Q_3(\sigma) \cap Q_3(\tau) \). The independence follows by the fact that the edge events in \( G_{n,d/n} \) are independent.

\[\diamondsuit\]

### 5.3 Proof of Lemma \(10\)

Before the proof of Lemma \(10\) we need a concentration results on the size of the set \( Q_2(\sigma) \).

For every \( u \in \sigma \) let \( J_u \) be the indicator r.v. such that \( J_u = 1 \) if \( u \in Q_2(\sigma) \) and zero otherwise. Conditioning on \( w \in Q_1(\sigma) \) the vertices in \( \sigma \) are adjacent to \( w \), equiprobably. It holds that

\[
Pr[J_u = 1 | X_1] = 1 - \left( 1 - \frac{d}{(1 - e) \ln d n} \right)^{X_1} \quad \forall u \in \sigma.
\]

Let the event \( A = \{ X_1 \geq (1 - \Theta(n^{-1/3})) (1 - e) \ln d n \} \). We have that

\[
\gamma = Pr[J_u = 1 | A] \geq 1 - \exp (-d^6) - O(n^{-1}).
\]

Let \( X_2 = \sum_{u \in \sigma} J_u = |Q_2(\sigma)| \). It holds that \( E[X_2 | A] \geq (1 - a) k \gamma \). It is easy to see that the random variables \( J_u \) are negatively depended. Using Theorem \(7\) we get that

\[
Pr[X_2 \leq (1 - \Theta(n^{-1/6})) (1 - a) k \gamma | A] \leq \exp \left( -n^{0.45} \right).
\]

Using Claim \(3\) and the law of total probability we get

\[
Pr[X_2 \leq (1 - \Theta(n^{-1/6})) k \gamma] \leq \exp \left( -n^{0.4} \right). \quad (12)
\]

**Proof of Lemma \(10\)** Consider the set \( \hat{Q}_3(\sigma) \subset V \setminus (\sigma \cup \tau \cup Q_1(\sigma)) \) which has the same properties as \( Q_3(\sigma) \) apart from \( S_3 \). First, we derive a lower for \( Pr[u \in \hat{Q}_3(\sigma)] \).

Let \( a \in [0,1] \) be such that \( |\sigma \cap \tau| = ak \). For each \( u \in V \setminus (\sigma \cup \tau \cup Q_1(\sigma)) \) let \( d_{\sigma,\tau}(u) \) be the number of vertices in \( \sigma \setminus \tau \) which are adjacent to \( u \). If \( d_{\sigma,\tau}(u) = i \), then all the subsets of size \( i \) in \( \sigma \setminus \tau \) are equiprobably adjacent to \( u \), i.e. \( Pr[u \in \hat{Q}_3(\sigma) | d_{\sigma,\tau}(u) = i] = \binom{X_2}{i} / \binom{(1-a)k}{i} \).

For \( a \geq 1 - n^{-1/4} \) we have that

\[
Pr[u \in \hat{Q}_3(\sigma)] \geq Pr[N_a \cap (\sigma \setminus \tau) = \emptyset] = (1 - p)^{(1-a)k} \geq \exp \left( -(1-a)k p \frac{p}{1-p} \right) \geq \exp \left( -n^{-1/4} \ln d \right) \geq 1 - \frac{\ln d}{n^{1/4}}. \quad (13)
\]

Assume now that \( a < 1 - n^{-1/4} \). Let the event \( A = \{ X_2 \geq (1 - \Theta(n^{-1/6})) (1 - a) k \gamma \} \) where \( \gamma = (1 - \exp(-d^6)) \). Assuming that the event \( A \) holds, for \( i \leq 7d \) we get

\[
\frac{(X_2)^i}{(1-a)^k} = \left( \frac{X_2}{(1-a)k} \right)^i (1 - o(1)) = \gamma^i (1 - o(1)).
\]
Let the event $C = \{d_{\sigma, \tau}(u) \neq 1 \text{ and } d_{\sigma, \tau}(u) \leq 7d\}$. Note that $d_{\sigma, \tau}(u)$ is binomially distributed with parameters $(1 - a)k$ and $p = d/n$ with the restriction that the event $C$ holds. We have

$$\Pr\{u \in \hat{Q}_3|A\} \geq \sum_{i \geq 0}^{7d} \Pr\{u \in Q_3|d_{\sigma, \tau}(u) = i, A\} \Pr\{d_{\sigma, \tau}(u) = i|A\}$$

where

$$q = (1 - a)kp(1 - p)^{(1 - a)k} \leq d^{-(1 - \epsilon)} \ln d.$$ 

Note that

$$\sum_{i = 7d + 1}^{(1 - a)k} \binom{(1 - a)k}{i} p^i(1 - p)^{(1 - a)k - i}(1 - \gamma)^i \leq \sum_{i = 7d + 1}^{(1 - a)k} \binom{(1 - a)k}{i} p^i(1 - p)^{(1 - a)k - i}$$

because $\gamma \leq 1$. Also

$$\sum_{i = 7d}^{(1 - a)k} \binom{(1 - a)k}{i} p^i(1 - p)^{(1 - a)k - i} \leq \sum_{i = 7d}^{n} \binom{n}{i} p^i(1 - p)^{n - i} \leq \exp(-7d).$$

The last inequality follows from Corollary 2.4 in [22]. We get that

$$S = \sum_{i = 0}^{7d} \binom{(1 - a)k}{i} p^i(1 - p)^{(1 - a)k - i}(1 - \gamma)^i (1 - p)^i$$

$$\geq (1 - p(1 - \gamma + \gamma p))^{(1 - a)k} - \exp(-7d)$$

$$\geq \exp(-(1 - a)(1 - \epsilon)e^{-d} \ln d - O(n^{-1})) - \exp(-7d)$$

$$\geq 1 - \frac{(1 - a)(1 - \epsilon) \ln d}{e^{d}} - \exp(-7d) - O(n^{-1}).$$

By the fact that $\Pr[C] \leq 1$, and $\gamma \leq 1$ we get

$$\Pr\{u \in \hat{Q}_3|A\} \geq 1 - \frac{(1 - a)(1 - \epsilon) \ln d}{e^{d}} - \frac{d^k \ln d}{d} - \exp(-7d) - O(n^{-1}).$$

Noting that $\Pr[A] \geq 1 - \exp(-\Theta(n^{a}))$ we get that for $0 \leq a < 1 - n^{-1/4}$

$$\Pr\{u \in \hat{Q}_3(\sigma)\} \geq 1 - \frac{(1 - a)(1 - \epsilon) \ln d}{e^{d}} - \frac{3 d^k \ln d}{2d} - O(n^{-1}).$$

By (13), it is direct that the above lower bound of $\Pr\{u \in \hat{Q}_3(\sigma)\}$ holds for any $a$.

Given that $v \in Q_3(\sigma)$ it is easy to see that it does not violate $S_3$ with probability at least $1 - O(n^{-1})$.

The law of total probability implies that

$$\Pr\{u \in Q_3\} \geq 1 - \frac{(1 - a)(1 - \epsilon) \ln d}{e^{d}} - \frac{3 d^k \ln d}{2d} - O(n^{-1}).$$

The lemma follows. \(\diamondsuit\)
6 Proof of Theorem \[2\]

To measure how close \(\sigma, \tau \in \mathcal{S}_k(G_{n,m})\) are, we define

\[
f_\sigma(\tau) = \left( \frac{|\sigma \cap \tau|}{n} \right).
\]  

(14)

The function \(f_\sigma\) is a map from the set \(\binom{[n]}{k}\) to the interval \([0, (k/n)]\). There is a direct correspondence between \(f_\sigma(\tau)\) and \(\text{dist}(\sigma, \tau)\), the Hamming distance between \(\sigma, \tau\). The greater \(f_\sigma(\tau)\) the smaller the Hamming distance between \(\sigma, \tau\). For a fixed \(\sigma \in \mathcal{S}(G)\) and the real \(\lambda > 0\) let

\[
g_{\sigma,G,\lambda}(x) = |\{\tau \in [n]_k : (f_\sigma(\tau) = x) \land H(\tau) \leq \lambda n\}|.
\]

Theorem \[2\] follows as a corollary of the following lemma.

**Lemma 11** There is \(\epsilon_d \to 0\) such that for any \(\epsilon_d \leq \epsilon \leq 1 - \epsilon_d\) and \(k = (1 + \epsilon)\frac{\ln d}{d} n\) the following is true. There are 0 < \(y_1 < y_2 < (k/n)^2\) and \(\lambda, \gamma > 0\) such that w.h.p. a pair \((G, \sigma)\) distributed as in \(\mathcal{U}_k(n, m)\) has the following two properties:

1. For all \(x \in [y_1, y_2]\) we have \(g_{\sigma,G,\lambda}(x) = 0\).

2. The number of independent sets \(\tau \in \mathcal{S}_k(G)\) such that \(f_\sigma(\tau) > y_2\) is at most \(\exp(-\gamma n)|\mathcal{S}_k(G)|\).

In the above lemma we take \(\epsilon_d > \ln^{-1/2} d\). Then, using Lemma \[8\] Lemma \[2\] and Proposition \[1\] suggest that w.h.p. \(|\mathcal{S}_k(G)| \geq \exp(2^{\ln^d d} - n)\).

The above lemma states that w.h.p. a 1 – \(o(1)\) fraction of independent sets in \(\mathcal{S}_k(G_{n,m})\) satisfy 1. and 2. We partition the space \(\mathcal{S}_k(G_{n,m})\) into subsets as follows: For each \(\sigma \in \mathcal{S}_k(G_{n,m})\) we define

\[
C_\sigma = \{\tau \in \mathcal{S}_k(G_{n,m}) : f_\sigma(\tau) > y_2\}.
\]

Start at \(S = \mathcal{S}_k(G_{nm})\) and remove iteratively some \(C_\sigma\) for \(\sigma \in S\) satisfying 1. and 2. of Lemma \[11\] It direct that we remove exponentially many subsets of \(\mathcal{S}_k(G_{nm})\). Furthermore, each such subset \(C_\sigma\) is separated by linear Hamming distance from \(\mathcal{S}_k(G_{nm}) \backslash C_\sigma\). Thus, Theorem \[2\] follows from Lemma \[11\]

To establish Lemma \[11\] we use the planted model.

**Lemma 12** There is \(\epsilon_d \to 0\) such that for any \(\epsilon_d \leq \epsilon \leq 1 - \epsilon_d\) and \(k = (1 + \epsilon)\frac{\ln d}{d} n\) the following is true. There are 0 < \(y_1 < y_2 < (k/n)^2\) and \(\lambda, \gamma > 0\) such that with probability at least \(1 - \exp\left(-\frac{\ln d}{d} n\right)\) a pair \((G, \sigma)\) distributed as in \(\mathcal{P}_k(n, m)\) has the following two properties

1. For all \(x \in [y_1, y_2]\) we have \(g_{\sigma,G,\lambda}(x) = 0\).

2. The number of independent sets \(\tau \in \mathcal{S}_k(G)\) such that \(f_\sigma(\tau) > y_2\) is at most \(\exp\left(\frac{2\log d}{d} n\right)\).

**Proof:** Consider the pair \((G, \sigma) \in \Lambda_k(n, m)\) chosen according with distribution \(\mathcal{P}_k(n, m)\). It direct that for \(\tau \in \binom{[n]}{k}\) the probability that \(H(\tau) \leq \lambda n\), w.r.t. to \(G\), is at most

\[
\left(\frac{1 - 2(k/n) + f_\sigma(\tau)}{1 - (k/n)^2}\right)^m \exp((y(\lambda) + o(1)) n)
\]

where \(\lim_{\lambda \to 0} y(\lambda) = 0\). For the rest of the proof assume that we take \(\lambda\) sufficiently small such that \(y(\lambda) \leq \frac{1}{d}\).
We prove the lemma by computing $A(x)$, the expected number of subsets $\tau \in [n]_k$ such that $H(\tau) \leq \lambda n$ and $f_\sigma(\tau) = x$, for suitable $y_1 < x < y_2$. Let, also, $s = (k/n)$ and assume that $x = s(1 - b)$, for $b \in [0, 1]$. It holds that

$$E[A(x)] \leq \left( \frac{k}{xn} \right) \left( \frac{n - k}{k - xn} \right) \left( \frac{1 - 2(k/n)^2 + f_\bar{\sigma}^2(\bar{\tau})}{1 - (k/n)^2} \right)^m \exp((y(\lambda) + o(1))n)$$

$$\leq \left( \frac{\epsilon^2 n^2 s}{(k - xn)^2} \right) \left( \frac{1 - 2s^2 + x^2}{1 - s^2} \right)^m \exp((y(\lambda) + o(1))n)$$

$$\leq \left( \frac{\epsilon^2}{bsn} \right) \left( 1 - s^2(1 - (1 - b)^2) \right)^m \exp((y(\lambda) + o(1))n)$$

Using the fact that $-\log s \leq \log d$ and $-q \ln q < e^{-1}$ for any real $q \geq 0$, we get

$$\frac{1}{n} \ln E[A(x)] \leq s \left[ 2 - 2e^{-1} + b \ln d(1 - (1 + \epsilon)(1 - b/2)) \right] + y(\lambda) + o(1).$$

Taking sufficiently large $d$ and sufficiently small $\lambda$ we get that $\frac{1}{n} \ln E[A(x)] < 0$ for $10 \ln^{-1} d < b < \frac{2e}{1 + \epsilon}$. In particular, for such $b$ we get that $E[A(x)] \leq \exp \left(-\frac{n}{d} \right)$.

The independent sets that belong to the same cluster as $\sigma$ are those for which $b$ is very small (the overlap $x$ is big). We should have $b$ so small that $2 - 2e^{-1} > -b \ln d(1 - (1 + \epsilon)(1 - b/2))$. Then, it is easy to see that $E[A(x)] < \exp \left(\frac{2\ln b}{d}n \right)$. The lemma follows by the Markov inequality.

\[\Box\]

7 Proof of Theorem

Definition 6 For the graph $G = (V, E)$ and an independent set $\sigma \in V$ in $G$, a vertex $v \in V \setminus \sigma$ will be called $\sigma$-pure if it is not adjacent to any vertex in $\sigma$.

Let $(G, \sigma)$ be a pair chosen from the distribution $\mathcal{P}_k(n, m)$. If we condition on all vertices in $\mathcal{P} \subset V \setminus \sigma$ being pure, then this implies that there are no $\sigma$-$\mathcal{P}$-edges, but it does not impose any conditioning on the subgraph $G[\mathcal{P}]$ induced on $\mathcal{P}$. Hence, $G[\mathcal{P}]$ is just a uniformly random graph, given its number of vertices and edges. More specifically we have the following:

Lemma 13 Consider the pair $(G, \sigma)$ distributed as in $\mathcal{P}_k(n, m)$ for $k = (1 - \epsilon)\ln_d n$ and $\epsilon \in [d^{-1/4}, 1]$. With probability at least $1 - 5 \exp(-\frac{d^3}{2\epsilon})$ the “$\sigma$-pure” vertices in $(G, \sigma)$ induce a graph which is distributed as in $G_{N,M}$ with $N \geq (1 - d^{-\epsilon/4})\frac{d^4}{\alpha}n$ and $M \leq (1 + d^{-\epsilon/5})\frac{d^5}{\alpha^2}n$.

Proof: Let $G_{n,p}(\sigma)$ be the graph $G_{n,p}$ conditional that $\sigma \in \mathcal{S}_k(G_{n,p})$. Assume that $p = \frac{D}{n}$ and $D = \frac{d}{(1 - s^2)(1 - d^{-\epsilon/3})} > d$. For every $v \in V \setminus \sigma$ let

$$I_v \begin{cases} 1 & \text{if } v \text{ is } \sigma\text{-pure in } G_{n,p}(\sigma), \\ 0 & \text{otherwise.} \end{cases}$$

It holds that $E[I_v] = (1 - p)^k = (1 - O(n^{-1}))\frac{d^4}{\alpha}$. Also, let $X_{pr} = \sum_{v \in V \setminus \sigma} I_v$. By the linearity of expectation we have $E[X_{pr}] = (n - k)(1 - p)^k = (1 - (1 - \epsilon)\ln d/d - O(n^{-1}))\frac{d^4}{\alpha}n$. Applying the Chernoff bounds we get that

$$\Pr \left[ |X_{pr} - E[X_{pr}]| \geq nd^{3\epsilon/4}/d \right] \leq 2 \exp \left(-\frac{n}{8d} \right). \quad (15)$$
Let the event $A = \{ |X_{pr} - E[X_{pr}]| \leq \frac{d^{3/4}}{d} \cdot n \}$. Let $Y_{pr}$ be the number of edges which are incident to $\sigma$-pure vertices only. It is direct that $E[Y_{pr}] \leq (1 + 3d^{-\epsilon/3}) \frac{d^2 \epsilon}{2d} n$. Applying the Chernoff bounds we get

$$Pr \left[ Y_{pr} \geq (1 + d^{-\epsilon/5}) \frac{d^2 \epsilon}{2d} n \right] \leq \exp \left( -\frac{n}{8d} \right).$$

(16)

Let the event $B = \{ Y_{pr} \leq (1 + d^{-\epsilon/5}) \frac{d^2 \epsilon}{2d} n \}$. By (15) and (16) we get that $Pr[A, B] \geq 1 - 4 \exp \left( -\frac{n}{8d} \right)$.

The lemma follows by bounding appropriately the number of edges in $G_{np}(\sigma)$. Let $Y$ be the number of edges in $G_{n,p}(\sigma)$. Clearly, $E[Y] = \frac{d^2 \epsilon / 3}{2(1-\epsilon^2)(1-d^{-\epsilon})} n$. Using Chernoff bounds it easy to show that

$$Pr[Y \leq m] \leq \exp \left( -nd^{2\epsilon/3} / 4 \right).$$

The lemma follows.

Combining Lemma 6 and Lemma 1 we get the following.

**Corollary 3** For the graph $G_{n,m}$ of expected degree $d$ it holds that

$$Pr \{ \alpha(G_{n,m}) \geq 2n(1 - o_d(1))\ln d / d \} \geq 1 - \exp \left( -\frac{n}{d \ln d} \right).$$

**Proof:** Consider $G^*_{n,m}$ of expected degree $d_0 = d(1 + 4d^{-1/4})$. Using Lemma 5 and Lemma 6 we get that

$$Pr \left[ \alpha(G^*_{n,m}) \geq \frac{2n}{d_0} \left( \ln d_0 - \ln \ln d_0 + 1 - \ln 2 - 10 \right) \right] \geq 1 - \exp \left( -\frac{2n}{d_0 \ln d_0} \right).$$

Substituting $d_0 = d(1 + 4d^{-1/4})$

$$Pr \left[ \alpha(G^*_{n,m}) \geq 2n(1 - o_d(1))\ln d / d \right] \geq 1 - \exp \left( -n/(d \ln d) \right).$$

Then, noting that $d_0(1 - 2d_0^{-1/3}) \geq d$ we use Lemma 1 to get the corollary.

If $I \subset P$ is an independent set, then so is $\sigma \cup I$. Hence, Lemma 13 and Corollary 3 imply that a random pair $(G, \sigma)$ chosen from the distribution $P_k(n, m)$ is $(2\epsilon/(1-\epsilon), 0)$-expandable with probability, at least, $1 - \exp(-n/(d \ln d))$. Finally, invoking Theorem 5 we see that a pair $(G, \sigma)$ chosen from the distribution $U_k(n, m)$ enjoys the same property w.h.p., whence we obtain Theorem 5.

**8 Proof of Theorem 4**

Now, assume that $k > (1 + \epsilon)n \ln(d) / d$ for $\epsilon > 0$. In this case it is direct to show the for the pair $(G, \sigma)$ distributed as in $P_k(n, m)$, the (expected) number of $\sigma$-pure vertices is bounded by $d^{-1-\epsilon} n \ll k$. Indeed, it is not very difficult to show that all but, say, $d^{-1-\epsilon/2} n$ vertices $v \in V \setminus \sigma$ have at least $0.1 \ln d$ neighbours in $\sigma$ w.h.p., provided that $d \geq d_0$ is not too small. Let us call these vertices blocking. Thus, in order to create an independent set of size bigger than $(1 + c)k$, for small fixed $c > 0$, we need to include $l \geq ck - d^{-1-\epsilon/2} n \geq ck/2$ blocking vertices w.h.p. Furthermore, since the bipartite graph spanned by $\sigma$ and the blocking vertices is random, it has excellent expansion properties w.h.p. Indeed, via a first moment argument one can show that these expansion properties imply that there is no independent set $\tau$ that contains $l \geq ck/2$ blocking vertices and $ck$ vertices from $\sigma$. More precisely, the first moment argument yields the following bounds.
Lemma 14 There is $\epsilon_d \to 0$ such that for any $\epsilon_d \leq \epsilon \leq 1 - \epsilon_d$ and $k = (1 + \epsilon)\ln \frac{n}{d} n$ the following is true. For $(G, \sigma)$ distributed as in $\mathcal{P}_k(n, m)$ and for $\delta, \gamma \in [0, 1]$ such that either $\delta < \gamma + 2\frac{\epsilon - 2\sqrt{\ln d}}{1 + \epsilon}$ or $(1 + \gamma)k > \alpha(G)$ we have that

$$P_{\mathcal{P}_k(n, m)}[\ln (G, \sigma), \sigma \text{ is not } (\gamma, \delta)-\text{expandable}] \geq 1 - \exp\left(-\frac{d}{\ln d} n\right).$$

Proof: Let $(G, \sigma) \in \Lambda_k(n, m)$ distributed as in $\mathcal{P}_k(n, m)$, where $k = (1 + \epsilon)\ln \frac{n}{d} n$ and $\epsilon > 0$. Let $E$ be the event that $\sigma$ is $(\gamma, \delta)$-expandable where $\delta < \gamma + 2\frac{\epsilon - 2\sqrt{\ln d}}{1 + \epsilon}$. Let $A_{\sigma}(\gamma, \delta)$ be the expected number of independent sets in $G$ which have size $(1 + \gamma)k$ and share $(1 - \delta)k$ vertices with $\sigma$. For $s = |\sigma|/n$ we have

$$A_{\sigma}(\gamma, \delta) = \left(\frac{k}{(1 - \delta)k}\right) \cdot \left(1 - \frac{n-k}{(1+\gamma)k - (1-\delta)k}\right) \cdot \left(\frac{1 - s^2 - (1 + \gamma)^2 s^2 + (1 - \delta)^2 s^2}{1 - s^2}\right)^m$$

$$\leq \left(\frac{k}{k/2}\right) \cdot \left(1 - \frac{n}{(1+\gamma)k}\right) \cdot \left(1 - 2\frac{(1 + \gamma - \delta)}{2} (\gamma + \delta) s^2\right)^m$$

$$\leq (2e)^{k/2} \cdot \exp\left(- (1 + \epsilon) (1 + \frac{\gamma - \delta}{2}) (\gamma + \delta) \ln dk\right).$$

By the fact that $-x \ln x \leq e^{-1}$ for $x \geq 0$ and $-\ln s \leq \ln d$ we get that

$$\frac{1}{k} \ln A_{\sigma}(\gamma, \delta) \leq \frac{1 + \ln 2}{2} + (\gamma + \delta) + e^{-1} - (\gamma + \delta) \ln d \left(1 + \epsilon\right) \left(1 + \frac{\gamma - \delta}{2}\right) - 1.$$

For $\delta = \delta(q) = \gamma + 2\frac{\epsilon - 2}{1 + \epsilon}$, for some $q > 0$, we get that $(1 + (\gamma - \delta)/2)(1 + \epsilon) - 1 = q$. Thus

$$\frac{1}{|\sigma|} \ln A_{\sigma}(\gamma, \delta) \leq 4 - 2 \left(\frac{\gamma}{1 + \epsilon}\right) q \ln d.$$

For $q > 2/\sqrt{\ln d}$ and $\epsilon > 10/\sqrt{\ln d}$ we get $\frac{1}{|\sigma|} \ln A_{\sigma}(\gamma, \delta) \leq -1$.

It is easy to show that for given $\gamma$ the expected $A_{\sigma}(\gamma, \delta)$ decreases as $\delta$ decreases. Thus, the lemma follows by noting that for $\delta = \gamma + 2\frac{\epsilon - 2\sqrt{\ln d}}{1 + \epsilon}$ it holds that $P_{\mathcal{P}}[E] \leq n A_{\sigma}(\gamma, \delta)$. \hfill \Box

Theorem 4 follows by using Theorem 5.

9 Proof of Corollary 1

For the Metropolis process described in Section 1.2, it is easy to see that every state in $\Omega = \bigcup_k \mathcal{S}_k(G(n, m))$ communicates with every other. Thus the process is ergodic and possesses a unique stationary distribution. Let $\pi : \Omega \to [0, 1]$ denote the stationary distribution of the Metropolis process with parameter $\lambda$, for some $\lambda > 0$. It is easy to show that $\pi(\sigma) = \lambda^{\sigma}/Z$ where $Z = \sum_{\sigma \in \Omega} \lambda^{\sigma}$.

Here, we are interested in finding the rate that the Metropolis process converges to the equilibrium. There are a number of ways of quantifying the closeness to stationarity. Let $P^t(\sigma, \cdot) : \Omega \to [0, 1]$ denote the distribution of the state at time $t$ given that $\sigma$ was the initial state. The total variation distance at time $t$ with respect to the initial state $\sigma$ is defined as

$$\Delta_{\sigma}(t) = \max_{S \subseteq \Omega} |P^t(\sigma, S) - \pi(S)| = \frac{1}{2} \sum_{\tau \in \Omega} |P^t(\sigma, \tau) - \pi(\tau)|.$$
Starting from $\sigma$, the rate of convergence to stationarity may then be measured by the function

$$
\tau_\sigma = \min_t \{ \Delta_\sigma(t') < e^{-1} \text{ for all } t' > t \}.
$$

We define the **mixing time** of the Metropolis process as

$$
T = \max_{\sigma \in \Omega} \tau_\sigma.
$$

Before showing Corollary 1 we provide some auxiliary results. In the following proposition we show that for a given parameter $\lambda$ the stationary distribution of the metropolis process concentrates on a small range of sizes of independent sets.

**Proposition 4** Let $I$ be an independent set chosen from the stationary distribution of the Metropolis process on the independent sets of $G(n, m)$ with parameter $\lambda$ as in (2). Letting

$$
K = \{ k : |\mu(G_{n,m}, \lambda) - k| \leq 2n/d \},
$$

we have that $\Pr[|I|/\in K] \leq \exp\left(-\frac{\ln^2 d n}{2d^2}\right)$. (17)

The proof of Proposition 4 appears in Section 9.1.

**Lemma 15**

$G(n, m)$ has the following property w.h.p. Let $K$ be as in (17), the set $\bigcup_{k \in K} S_k(G)$ admits a partition into classes $C_1, \ldots, C_N$ such that

1. $Pr[I \in C_i] \leq \exp(-n \ln d/d)$ for each $i \leq i \leq N$.
2. We have that $Pr[I \in \bigcup_{1 \leq i \leq N} C_i] \geq 1 - \exp(-n \ln d/d)$.
3. The distance between any two independent sets in different classes is at least 2.

The proof of Lemma 15 appears in Section 9.2. Corollary 1 follows from the following lemma.

**Lemma 16**

Consider the Metropolis process on the independent sets of $G(n, m)$ with parameter $\lambda$ as in (2). W.h.p. the mixing time of the process is greater than $e^{n/d^2}$.

**Proof:** Let $K$ be as in (17) and assume that $G_{n,m}$ is such that $\bigcup_{k \in K} S_k(G)$ has a partition $C_1, \ldots, C_N$ satisfying 1, 2 and 3 of Lemma 15.

The proof is going to be made by contradiction. Assume that our Metropolis process has mixing time $T \leq e^{n/d^2}$.

Let $I_t$ be the state of the Metropolis process at time $t > 0$.

Consider, now, the time interval $[t_1, t_2]$ such that $t_1 = n^2 T$ and $t_2 = 2n^2 T$. It is direct to show that the distribution of $I_t$, for $t \in [t_1, t_2]$ is extremely close to the stationary distribution, i.e. for any $t \in [t_1, t_2]$ the total variation distance between the distribution of $I_t$ and $\pi$ is at most $e^{-n^2}$. By Proposition 4, Lemma 15 and standard arguments about mixing time, it follows that for any $t \in [t_1, t_2]$

$$
Pr \left[ I_t \notin \bigcup_{1 \leq i \leq N} C_i \right] \leq 2 \exp\left(-\frac{n \ln^2 d(1-o(1))}{d^2}\right).
$$

Applying the union bound we get that

$$
Pr \left[ \exists t_1 \leq t \leq t_2 : I_t \notin \bigcup_{1 \leq i \leq N} C_i \right] \leq 2 \exp\left(-\frac{(1-o(1)) \ln^2 d n}{2d^2} + \frac{n}{d^2}\right) \leq \exp\left(-\frac{\ln^2 d n}{2d^2}\right).
$$

(18)
The above relation implies that so as the Metropolis to get from \(I_{t_1}\) to \(I_{t_2}\) w.h.p. it uses only independent sets which are in \(\bigcup_{1 \leq i \leq N} C_i\).

Let us exam now the states \(I_{t_1}\) to \(I_{t_2}\). We should observe that since the interval \([t_1, t_2]\) is so long, i.e. \(n^2 T\), the two states are (almost) independent of each other. Using Proposition 4 and (1) from Lemma 15 we get the that

\[
Pr[\exists i, j \in [N], i \neq j : I_{t_1} \in C_i \cap I_{t_2} \in C_j] \geq 1 - 4 \exp(-n \ln d/d^2).
\]  

(19)

Now (3) of Lemma 15 causes the problem. For assume that there are two distinct \(i, j \in [N]\) such that \(I_{t_1} \in C_i\) and \(I_{t_2} \in C_j\). Let \(t > t_1\) be the first time that \(I_t \notin C_i\). Then by definition we should have that \(\text{dist}(I_t, I_{t-1}) \leq 1\). This implies that \(I_t \notin \bigcup_{1 \leq i \leq N} C_i\) because otherwise there would be two independent sets in different classes at distance one. Thus we conclude that

\[
Pr[\exists i, j \in [N], i \neq j : I_{t_1} \in C_i \cap I_{t_2} \in C_j] \leq Pr \left[ \exists t_1 \leq t \leq t_2 : I_t \notin \bigcup_{1 \leq i \leq N} C_i \right],
\]

contradicting (18) and (19). The lemma follows. \(\diamondsuit\)

9.1 Proof of Proposition 4

For the Metropolis process on the independent sets of \(G_{n,m}\) with parameter \(\lambda\), we let

\[
R(k, \lambda) = |S_k(G_{n,m})|\lambda^k.
\]

In essence, \(R(k, \lambda)\) is the total weight carried by the independent sets of size \(k\) in \(G_{n,m}\). For \(c \in [0, 1]\), we denote with \(\lambda_c\) the value of the parameter of the Metropolis process that maximizes \(E[R(k, \lambda)]\) for \(k = (1 + c)\frac{\ln d}{d^2}n\). First, we show the following lemma.

Lemma 17 For \(c \in [0, 1]\) and \(k = (1 + c)\frac{\ln d}{d^2}n\), it holds that

\[
Pr \left[ R(k, \lambda_c) \leq \exp \left( -14n \sqrt{\ln^5 d/d^2} \right) E[R(k, \lambda_c)] \right] \leq \exp \left( -n \ln^2 d/d^2 \right).
\]

Proof: Note that \(R(k, \lambda_c) = (\lambda_c)^k X(k)\), where \(X(k)\) is the number of independent sets of size \(k\) in \(G_{n,m}\). It is easy to see that using Proposition 1 we can get a lower bound on \(R(k, \lambda_c)\), since we only need a lower bound on \(X(k)\). The lemma follows. \(\diamondsuit\)

Proposition 4 follows as a corollary from the following lemma.

Lemma 18 For some constant \(c \in [0, 1]\), let and \(k = (1 + c)\frac{\ln d}{d^2}n\) and

\[
R_0(\lambda_c) = \exp \left( -14n \sqrt{\ln^5 d/d^2} \right) E[R(k, \lambda_c)].
\]

It holds that

\[
Pr \left[ \sum_{k' : |k - k'| \geq \frac{2n}{d}} R(k', \lambda_c) \geq \exp (-n/d) R_0(\lambda) \right] \leq \exp \left( -3n/d \right).
\]
Proof: For convenience we write \( k \), the size of the independent sets, in the form

\[
k_{-\epsilon} = \frac{2n}{d} \left( \ln d - \ln \ln d + 1 - \ln 2 - \epsilon \right).
\]

where \( \epsilon \) can vary so as \( k_{-\epsilon} \leq \alpha(G_{n,m}) \). Assume that for \( \epsilon_0 \) it holds that \( k_{-\epsilon_0} = (1 + \epsilon) \frac{\ln d}{d} n \). For convenience we define the following functions:

- \( y(\epsilon) = (\ln d - \ln \ln d + 1 - \ln 2 - \epsilon) \)
- \( h(\epsilon) = \frac{2}{\lambda} y(\epsilon) \ln \lambda_c + \epsilon - \ln (y(\epsilon)/\ln d) \).

Using Lemma 3 and Lemma 2, we get that for any \( \epsilon \), \( \epsilon_0 \) it holds that

\[
\text{the above imply that } h(\epsilon) \leq \frac{1}{n} \ln E[R(k_{-\epsilon}, \lambda_c)] \leq h(\epsilon) + \frac{3}{d^2} \frac{\ln^2 d}{d^2}.
\]

First we need to compute a lower bound of \( \ln \lambda_c \) with respect to \( \epsilon_0 \). It is easy to get that

\[
\frac{d}{d\epsilon} h(\epsilon) = \frac{2}{d} \left( y(\epsilon) + 1 - \epsilon + \ln \left( \frac{y(\epsilon)}{\ln d} \right) - \ln \lambda_c \right).
\]

Letting \( \epsilon' \) be such that \( h'(\epsilon') = 0 \) we get that

\[
\ln \lambda_c = y(\epsilon') + 1 - \epsilon' + \ln \left( \frac{y(\epsilon')}{\ln d} \right).
\]

It is direct to check that \( \epsilon' \) is the maximum of the function \( h \). For \( \epsilon = \epsilon' + q \) where \( |q| < (\ln d)^{1/3} \) we get that \( h(\epsilon) \leq h(\epsilon') - \frac{3}{d^2} \frac{q^2}{d^2} \).

We let \( f(\epsilon) = \frac{1}{n} \ln E[R(k_{-\epsilon}, \lambda_c)] \). Since \( f(\epsilon) \leq h(\epsilon) + 2 \frac{\ln^2 d}{d^2} \) it is easy to see that \( |\epsilon_0 - \epsilon'| \leq d^{-2/5} \).

Otherwise, i.e. if \( |\epsilon_0 - \epsilon'| \geq d^{-2/5} \) then we would get that \( f(\epsilon_0) < h(\epsilon') \), which is a contradiction. All the above imply that

\[
\log \lambda_c \geq y(\epsilon_0) + 1 - \epsilon_0 + \ln \left( \frac{y(\epsilon_0)}{\ln d} \right) - 5d^{-2/5}. \tag{20}
\]

It holds that

\[
f(\epsilon) - f(\epsilon_0) \leq \frac{2}{d} \left( \ln \lambda_c (y(\epsilon) - y(\epsilon_0)) + y(\epsilon) \left( \epsilon - \ln \left( \frac{y(\epsilon)}{\ln d} \right) \right) - y(\epsilon_0) \left( \epsilon_0 - \ln \left( \frac{y(\epsilon_0)}{\ln d} \right) \right) + 3 \frac{\ln^2 d}{d} \right).
\]

Setting \( t = \epsilon - \epsilon_0 \), then for \( |t| \leq 5 \) we get that

\[
f(\epsilon) - f(\epsilon_0) \leq \frac{2}{d} \left( -t \ln \lambda_c + ty(\epsilon) - t\epsilon_0 - y(\epsilon) \ln \left( \frac{y(\epsilon)}{\ln d} \right) + y(\epsilon_0) \ln \left( \frac{y(\epsilon_0)}{\ln d} \right) + 3 \frac{\ln^2 d}{d} \right).
\]

Using \( y(\epsilon) = y(\epsilon_0) - t \), we get \( \ln (y(\epsilon)/\ln d) = \ln (y(\epsilon_0)/\ln d) - t/y(\epsilon_0) - t^2/(2y^2(\epsilon_0)) - O(t^3/y^3(\epsilon_0)) \).

Then we get that

\[
f(\epsilon) - f(\epsilon_0) \leq \frac{2}{d} \left[ t \left( -\ln \lambda_c + y(\epsilon_0) + 1 - \epsilon_0 + \ln \left( \frac{y(\epsilon_0)}{\ln d} \right) - O(t^2/y^2(\epsilon_0)) \right) - t^2 \left( 1 + \frac{1}{2y(\epsilon_0)} \right) + 3 \frac{\ln^2 d}{d} \right].
\]

Using (20) and the fact that \( 1/2 < \frac{y(\epsilon_0)}{\log d} < 1 \) we get

\[
f(\epsilon) - f(\epsilon_0) \leq \frac{2}{d} \left[ 5/\ln d - t^2 \right]
\]

\footnote{The reader should note that this parametrization of the size of independent sets is different than the one in Definition.}

27
For sufficiently large $d$ and $\epsilon$ such that $|\epsilon - \epsilon_0| = 1$ we get that $E[R(k_{-\epsilon}, \lambda_c)] \leq E[R(k_{-\epsilon_0}, \lambda_c)] \exp(-1.9n/d)$. By the Markov inequality we get that

$$
Pr[R(k_{-\epsilon}, \lambda_c) \geq \exp(-n/d) R_0(\lambda_c)] \leq Pr \left[ R(k_{-\epsilon}, \lambda_c) \geq \exp \left( -15n \sqrt{ \frac{\ln 5}{d^3} } E[R(k_{-\epsilon_0}, \lambda_c)] \right) \right]
$$

$$
\leq Pr \left[ R(k_{-\epsilon}, \lambda_c) \geq \exp \left( 3n/(2d) \right) E[R(k_{-\epsilon}, \lambda_c)] \right]
\leq \exp(-3n/(2d)).
$$

It is easy to see that for $|t| > 1$ the above probability is even smaller. The lemma follows. \hfill \Box

9.2 Proof of Lemma 15

Lemma 15 follows from the following lemma.

\textbf{Lemma 19} Let $(G, \sigma) \in \Lambda_k(n, m)$ be distributed as in $U_k(n, m)$, for $k \in K$, where $K$ and $\mu(G, \lambda)$ is as in (17) and (2), respectively. The set $\bigcup_{k \in K} S_k(G)$ admits a partition into classes $C_1, \ldots, C_N$ such that

1. $Pr[\sigma \in C_i] \leq \exp(-n \ln d/d)$ for any $i \in [N]$
2. $Pr[\sigma \notin \bigcup_{i \in [N]} C_i] \leq \exp(-n \ln d/d)$
3. The distance between two independent sets in different classes is at least 2.

To see that Lemma 15 follows from Lemma 19, first note that for the pair $(G, \sigma)$ distributed as in $U_k(n, m)$ and any $A \subset 2^n$, it holds that

$$
Pr[\sigma \in A] \leq \max_k Pr[\mathcal{I} \in A | |\mathcal{I}| = k]
$$

for any $k$. Given Lemma 19 we have the following: The statement 1 in Lemma 15 follows by noting $Pr[\mathcal{I} \in C_i] \leq \max_{k \in K} Pr[\mathcal{I} \in C_i | |\mathcal{I}| = k]$ and using the statement 1 in Lemma 19. Similarly, the statement 2 of Lemma 15 follows by noting $Pr[\mathcal{I} \notin \bigcup_{i \in [N]} C_i] \leq \max_{k \in K} Pr[\mathcal{I} \notin \bigcup_{i \in [N]} C_i | |\mathcal{I}| = k]$ and using statement 2 in Lemma 19. The statement 3 in Lemma 15 and statement 3 from Lemma 19 are identical.

So as to show Lemma 19 it is sufficient to prove the following lemma and invoke the transfer theorem.

\textbf{Lemma 20} Let $(G, \sigma)$ be distributed as in $P_k(n, m)$, with $k \in K$ and $K$ and $\mu(G, \lambda)$ is as in (17) and (2), correspondingly. Let $S_{\sigma}(x, t) \subseteq S_t(G)$ contain the independent sets having $xk$ common vertices with $\sigma$, where $x \in [0, 1]$ and integer $t > 0$.

1. There are constants $0 < y_1 < y_2 < 1$ such that for any $y_1 < x < y_2$ it holds that $Pr \left[ \bigcup_{t \in K} S_{\sigma}(x, t) = \emptyset \right] \geq 1 - \exp(-n \ln d/d)$.
2. $Pr \left[ \sum_{t \in K} | \bigcup_{x \geq y_2} S_{\sigma}(x, t) | \geq \exp(8n \ln d/d) \right] \leq \exp(-5n \ln d/d)$.

It is easy to see that given the statement 1 of Lemma 20 the transfer theorem implies the statements 2 and 3 of Lemma 19. The statement 2 of Lemma 20 combined with the transfer theorem and the concentration result of the number of independent sets of size $k$ in $G_{n,m}$ we get the statement 1 of Lemma 19. It remains to prove Lemma 20.
Proof of Lemma 20. As far as the first statement of the lemma is concerned, consider a pair \((G, \sigma) \in \Lambda_k(n, m)\) distributed as in \(P_k(n, m)\), with \(k \in K\). We are going to show that there is a range \([y_1, y_2]\) such that for any \(x \in [y_1, y_2]\) and any \(t \in K\) it holds that \(E[S(t, x)] < \exp(-2n \ln d/d)\). Then statement 1 of the lemma follows by noting that \(\sum_{t \in K} E[S_\sigma(x)] < \exp(-3/2n \ln d/d)\) and by invoking the Markov inequality.

For convenience we set \(s = (k/n)\) and we write \(t = ak\) for appropriate \(a\). So as to have \(t \in K\) there exists a constant \(q\), such that \(a \in (1 - q/\ln d, 1 + q/\ln d)\). Clearly, it holds that

\[
E[A(x, t)] \leq \left( \frac{k}{xk} \right) \left( \frac{n-k}{t-xk} \right) \left( \frac{1 - (k/n)^2 - (t/n)^2 + (sx)^2}{1 - (k/n)^2} \right)^m (1 + o(1)).
\]

Substituting \(t = ak\) in the above inequality we get

\[
E[A(x, t)] \leq \left( \frac{ke}{xk} \right)^x \left( \frac{ne}{(a-x)k} \right)^{(a-x)k} \left( 1 - s^2 \frac{a^2 - x^2}{1 - s^2} \right)^m.
\]

Thus we get

\[
\frac{1}{k} \ln E[A(x, t)] \leq x[1 - \ln x] + (a-x)[1 - \ln s - \ln(a-x)] - (1 + \epsilon)(a-x) \ln d^{a+x/2}
\]

\[
\leq 3 + (a-x) \ln d \left[ 1 - (1 + \epsilon)^{a+x/2} \right]
\]

where in the second derivation we used the fact that \(- \ln s < \ln d, -q \ln q < e^{-1}\) for any real \(q > 0\) and for sufficiently large \(d\) \(0 < a < 1.2\).

It is clear that for given \(\epsilon\) and \(a\) such that \(t \in K\) in the r.h.s. of the last inequality, above, we can find an appropriate range for \(x\) such that \(\frac{1}{k} \ln E[A(x, t)] < -2\). The statement 1 of the lemma follows.

We now give the proof of the statement 2 of the lemma. Using the above inequalities, it is easy to see that we can take sufficiently large \(x\) so as to have \(\frac{1}{k} \ln E[A(x, t)] > 0\). In particular, it should hold that \((a-x) \ln d (1 - (1 - \epsilon)(a+x)/2) < 3\). In this case it is easy to show that \(\frac{1}{k} \ln E[A(x, t)] < 3\). Using the Markov inequality we get that

\[
Pr \left[ \sum_{x \geq y_2} A(x, k) > \exp(8k) \right] \leq \exp(-5k).
\]

The statement 2 of the lemma follows.

\(\diamond\)

10 Large deviations results

Corollary 4 There is a constant \(d_0 > 0\) such that for all \(d \geq d_0\) there is a number \(\xi = \xi(d) > 0\) such that for any \(\ln d/d \leq \alpha \leq 2 \ln d/d\) the following is true. Let \(I\) be the number of independent sets of size \([\alpha n]\) in \(G(n, m)\). Then \(I \leq \exp(-\xi n) \cdot E I\) w.h.p.

Proof: Let \(G\) be a graph chosen from the following ‘planted’ distribution: any edge that does not connect two vertices from \(S = \{1, \ldots, \lceil \alpha n \rceil\}\) is present in \(G\) with probability \(q\) independently, where \(q\) is chosen so that the expected number of edges equals \(m\). That is,

\[
q = \left( \frac{m}{\binom{n}{2}} - \binom{\lceil \alpha n \rceil}{2} \right) \sim \frac{m}{\binom{n}{2}(1 - \alpha^2)} \sim \frac{d}{n(1 - \alpha^2)}.
\]

Let \(X\) be the number of vertices in \(v \in S\) that are isolated in \(G\). The expected degree of each \(v \in S\) equals

\[
\gamma = q(1 - \alpha)n = d \cdot \frac{1 - \alpha}{1 - \alpha^2} = \frac{d}{1 + \alpha}.
\]

29
Moreover, the degree of each \( v \in S \) is binomially distributed. Therefore, each \( v \in S \) is isolated with probability \((1 - \delta_d) \exp(-d/(1 + \alpha))\), where \( \delta_d \to 0 \) for large \( d \). As the degrees of the vertices \( v \in S \) are mutually independent random variables, we thus see that \( X \) is binomially distributed as well and

\[
\mathbb{E}X \geq an \cdot (1 - \delta_d) \exp(-d/(1 + \alpha)) \geq an \exp(-d(1 - \alpha + \alpha^2)) \geq 100n \exp(-d),
\]

provided that \( d \geq d_0 \) is sufficiently large. Furthermore, as \( X \) is binomially distributed, Chernoff bounds yield a number \( \xi = \xi(d) > 0 \) such that

\[
P[X \leq 2n \exp(-d)] \leq \exp(-3\xi n).
\]

On the other hand, let \( Y \) be the total number of isolated vertices in \( G(n, m) \). Since the degree of each vertex in \( G(n, m) \) is binomially distributed with mean \( d \), we have \( \mathbb{E}Y \leq n \exp(-d) \). Furthermore, a standard second moment computation shows that \( \text{Var}(Y) = O(n) \). Therefore, by Chebyshev’s inequality \( Y \leq 1.9n \exp(-d) \) w.h.p.

Thus, let \( E \subset \Lambda_k(n, m) \) be the set of all pairs \((G, \sigma)\) such that \( G \) has at least \( 100n \exp(-d) \) isolated vertices. We have shown that \( P_p[E] \geq 1 - \exp(-3\xi n) \), while \( P_{(d)}[E] = o(1) \). However, if \( P[I \geq \exp(-\xi n)EZ] \geq \varepsilon \) for some fixed \( \varepsilon > 0 \), then by the same argument as in Theorem 5 we would obtain \( P[\neg E] \geq (1 - o(1))\varepsilon \exp(-\xi n) \), which is a contradiction. Hence, we have shown that \( P[I \geq \exp(-\xi n)EZ] = o(1) \). \( \diamond \)

References

[1] D. Achlioptas and A. Coja-Oghlan. Algorithmic Barriers from Phase Transitions. In proc. of FOCS 2008, pp 793-802.

[2] D. Achlioptas and M. Molloy. The Analysis of a List-Coloring Algorithm on a Random Graph. In proc. of FOCS 1997, pp 204-212.

[3] N. Bhatnagar, A. Sly and P. Tetali. Reconstruction Threshold for the Hardcore Model. In proc. of APPROX-RANDOM 2010, pp 434-447.

[4] B. Bollobás. Random graphs, 2nd edition Cambridge (2001).

[5] B. Bollobás and P. Erdős. Cliques in random graphs. Math. Proc. Comb. Phil. Soc. 80, (1976) pp 419-427.

[6] V. Dani and C. Moore. Independent sets in random graphs from the weighted second moment method. Preprint.

[7] P. Erdős. Some remarks on the theory of graphs. Bull. Amer. Math. Soc. 53 (1947) pp 292294

[8] A. M. Frieze. On the independence number of random graphs. Discrete Mathematics 81 (1990) pp 171-175.

[9] A. M. Frieze and C. McDiarmid. Algorithmic theory of random graphs. Random Struct. Algorithms 10, (1997) pp 5-42.

[10] A. Frieze and S. Suen. Analysis of two simple heuristics on a random instance of k-SAT. Journal of Algorithms 20 (1996) 312--355.

[11] Y. Fu, P. W. Anderson. Applications of statistical mechanics to NP-complete problems in combinatorial optimization. J. Phys. A 19 (1986) 1605.
[12] D. Gamarnik, T. Nowicki and G. Swirscsz. Maximum Weight Independent Sets and Matchings in Sparse Random Graphs. Random Structures and Algorithms 28 (2005) pp. 76-106.

[13] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. Math. Proc. Cambridge Philos. Soc., 77 (1975), pp 313 - 324.

[14] M. R. Jerrum. Large cliques elude the Metropolis process. Random Structures and Algorithms 3 (1992) pp 347-359.

[15] M. Mezard, G. Parisi and R. Zecchina. Analytic and Algorithmic Solution of Random Satisfiability Problems. In Science 297 (2002), pp 812-815.

[16] R. M. Karp. The probabilistic analysis of some combinatorial search algorithms, in Algorithms and Complexity: New Directions and Recent Results, J.F. Traub, ed., Academic Press (1976) 1-19.

[17] R. Karp and M. Sipser. Maximum matchings in sparse random graphs. In proc. of FOCS 1981 pp. 364-375.

[18] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjianc, L. Zdeborova. Gibbs states and the set of solutions of random constraint satisfaction problems. In Proc. National Academy of Sciences 104 (2007) pp 10318-10323.

[19] L. Kučera. Expected Behavior of Graph Coloring Algorithms. In Fundamentals of Computation Theory, LNCS Vol. 56, Springer Verlag (1977), pp 447-451.

[20] D. Matula. The largest clique size in a random graph. Tech. Rep., Dept. Comp. Sci., Southern Methodist University, Dallas, Texas, 1976.

[21] A. Montanari, R. Restrepo and P. Tetali. Reconstruction and clustering thresholds in random CSPs. SIAM J. Discrete Math. (to appear).

[22] S. Janson, T. Luczak and A. Ruciński. Random graphs, Wiley and Sons, Inc. 2000.

[23] B. Rossman. The Monotone Complexity of k-Clique on Random Graphs. To appear in FOCS 2010.

[24] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. Ins. Hautes Études Sci. Publ. Math. 81 (1995) pp 73-205.