On Dominating Cliques in Random Graphs

Martin Nehéz
Department of Computer Science and Engineering,
FEE CTU Prague, Karlovo náměstí 13,
121 35 Praha 2, Czech Republic
e-mail: nehez@fel.cvut.cz

Daniel Olejár
Department of Computer Science,
FMPI, Comenius University in Bratislava, Mlynská dolina,
842 48 Bratislava, Slovak Republic

Abstract

Motivated by the communication problems in large-scale networks, we study the dominating cliques in random graphs in this paper. Our main result points out conditions for an existence of dominating cliques in random graphs $\mathbb{G}(n,p)$ in the terms of bounds on the probability p.

1 Introduction

Given a graph $G = (V, E)$, a set $S \subseteq V$ is said to be a *dominating set* of G if each node $v \in V$ is either in S or is adjacent to a node in S. The *domination number* $\gamma(G)$ is the minimum cardinality of a dominating set of G.

There are several alternative definitions of the dominating set [4]. The following one is important for the our purpose. Given two nodes $u, v \in V$, let $d_G(u, v)$ denote the distance between u and v in G. Let $\Gamma(u) = \{v \in$

$V \mid d_G(u,v) \leq 1\}$ denote a *ball* of radius 1 centered at u. For every subset $S \subseteq V$, let $\Gamma(S) = \cup_{u \in S}\Gamma(u)$. A subset S is said to be a *dominating set* of G if $\Gamma(S) = V$.

A *clique* in G is a maximal set of mutually adjacent nodes of G, i.e., it is a maximal complete subgraph of G. The *clique number*, denoted $cl(G)$, is the number of nodes of clique of G. If a subgraph induced by a dominating set is a clique in G then the induced subgraph is called a *dominating clique* in G. Dominating sets and cliques are basic structures in graphs that have been investigated very intensively. To determine whether the domination number of a graph is at most r is an NP-complete problem [3]. The maximum-clique problem is one of the first shown to be NP-hard [7]. A well-known celebrated result of B. Bollobás, P. Erdös et al. is a proof that the clique number in random graphs is bounded by a very tight bounds [1, 2, 6, 8, 9, 10].

The model of random graphs is introduced in the following way. Let p, $0 \leq p \leq 1$, be a *probability of an edge*. The *(probabilistic) model of random graphs* $\mathbb{G}(n,p)$ consists of all graphs with n-node set $V = \{1, \ldots, n\}$ such that each graph has at most $\binom{n}{2}$ edges being inserted independently with probability p. Equivalently, if G is a graph with node set V and it has $|E(G)|$ edges, then:

$$\Pr[G] = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|} ,$$

where Pr is a probability measure defined on $\mathbb{G}(n,p)$. This model is also called *Erdös-Rényi random graph model* [1, 5].

Let A be any set of graphs from $\mathbb{G}(n,p)$ with a property Q. We say that *almost all graphs* have the property Q iff:

$$\Pr[A] \rightarrow 1 \quad as \quad n \rightarrow \infty .$$

The domination number of a random graph have been studied by B. Wieland and A. P. Godbole in [11]. The expectation of a random variable D_r which stands for the number of dominating sets of size r in $\mathbb{G}(n,p)$ is given by:

$$E(D_r) = \binom{n}{r}(1-(1-p)^r)^{n-r} .$$

As was also claimed in [11], the domination number of the random graphs $G \in \mathbb{G}(n,p)$ with probability approaching to 1 is bounded as follows:

$$\lfloor \, \mathbb{L}n - \mathbb{L}[(\mathbb{L}n)(\ln n)] \, \rfloor + 1 \leq \ \gamma(G) \ \leq \lfloor \, \mathbb{L}n - \mathbb{L}[(\mathbb{L}n)(\ln n)] \, \rfloor + 2 \ , \quad (1)$$

where $\mathbb{L}x$ denotes $\log_{1/(1-p)} x$ and p is a fixed constant or a suitable function [11]. Using the logarithm with only the one base, the inequality (1) can be rewritten as follows (derived from Lemma 3 of [11], p. 6):

$$\lfloor \mathbb{L}n - 2 \cdot \mathbb{L}(\mathbb{L}n) + \mathbb{L}(\mathbb{L}e) \rfloor + 1 \leq \gamma(G) \leq \lfloor \mathbb{L}n - 2 \cdot \mathbb{L}(\mathbb{L}n) + \mathbb{L}(\mathbb{L}e) \rfloor + 2 \,. \quad (2)$$

Motivated by the communication problems in large-scale networks, we deal with the dominating cliques in random graphs in this paper. The key question posed is whether dominating cliques really exist. And if they do, how many dominating cliques of a given order are in $\mathbb{G}(n,p)$. We answer both these questions in this paper. We show that the existence of dominating cliques strongly depends on the edge probability p. Our main result points out existence conditions in terms of bounds on p.

The rest of this paper is organized as follows. Section 2 contains the preliminary results. The number of dominating cliques in $\mathbb{G}(n,p)$ is estimated here. The main result is proved in Section 3. Several open problems are posed in Conclusions.

2 Preliminary results

For $r > 1$, let S be an r-node subset of an n-node graph G. Let A denote the event that "S is a dominating clique of $G \in \mathbb{G}(n,p)$". Let in_r be the associated 0-1 (indicator) random variable on $\mathbb{G}(n,p)$ defined as follows: $in_r = 1$ if G contains a dominating clique S and $in_r = 0$, otherwise. Let X_r be a random variable that denotes the number of r-node dominating cliques. More precisely, $X_r = \sum in_r$ where the summation ranges over all sets S. The following lemma expresses the expectation of X_r.

Lemma 1 *For the expectation $E(X_r)$ of the random variable X_r*

$$E(X_r) = \binom{n}{r} p^{\binom{r}{2}} (1 - p^r - (1-p)^r)^{n-r} \,. \quad (3)$$

Proof. The linearity of the expectation leads to

$$E(X_r) = \sum E(in_r) = \sum in_r \cdot \Pr[A] \,,$$

over all r-node sets S. The nodes of the S can be chosen in $\binom{n}{r}$ ways. Since S is a complete subgraph, every of its r nodes has to be adjacent to the remaining $r-1$ nodes of S. Hence, the probability of this fact is $p^{\binom{r}{2}}$. The last term in (3) expresses the probability that S is a clique spanning a dominating set of $G \in \mathbb{G}(n,p)$. More precisely, let v be an arbitrary but fixed node, $v \notin V(S)$; v is said to be a "good" node (i.e., the node which does not spoil the "cliqueness" and the "domination" of S), if it cannot be adjacent neither to all nodes of S nor to none of them. It follows that v has to be adjacent at least to one and at most to $r-1$ of nodes of S. Therefore,

$$\Pr[\ v \textit{ is a "good" node with respect to } S\ ] = \sum_{0<j<r} \binom{r}{j} p^j (1-p)^{r-j} =$$

$$= \left[\sum_{j=0}^{r} \binom{r}{j} p^j (1-p)^{r-j} \right] - p^r - (1-p)^r = 1 - p^r - (1-p)^r \ .$$

All of $(n-r)$ nodes from $V(G) \setminus V(S)$ must be "good". Hence, the Lemma follows. $\diamondsuit$ To obtain an upper bound on the order r of dominating cliques in random graphs we use the following property adopted from [9], pp. 501–502.

Claim 1 *Let $0 < p < 1$ and $k = O(\log n)$. Then:*

$$(1-p^k)^n = \exp(-np^k)\left(1 + O(np^{2k})\right) = 1 - np^k + O\left(np^{2k}\right) \ .$$

The upper bound on r is stated in the following lemma.

Lemma 2 *Let $b = 1/p$ and*

$$r_u = 2\log_b n - 2\log_b \log_b n + 2\log_b e + 1 - 2\log_b 2 \ . \tag{4}$$

A random graph from $\mathbb{G}(n,p)$ does not contain dominating cliques on the order greater than r_u with probability approaching to 1 as $n \to \infty$.

Proof. The proof follows from the Markov's inequality [5], p. 8:

$$\Pr[\ X \geq t\ ] \leq \frac{E(X)}{t} \ , \qquad t > 0 \ .$$

Let us denote $\alpha = \log_{1/p}\left(\frac{1}{1-p}\right) = -\log_b(1-p)$. Note that:

$$(1-p)^r = p^{r\alpha} \ . \tag{5}$$

Let $r = (2-\varepsilon)\log_b n$, where $0 \le \varepsilon < 1$. According to Claim 1,

$$(1-p^r-(1-p)^r)^{n-r} = (1-p^r-p^{r\alpha})^{n-r} =$$
$$= 1-n^{-1+\varepsilon}-n^{-2\alpha+1+\varepsilon}+O(n^{-4\alpha+1+2\alpha\varepsilon}) \ .$$

It follows that:

$$(1-p^r-(1-p)^r)^{n-r} \to 1 \qquad as \qquad n \to \infty \ ,$$

if $p \ge 1/2$ and hence, $\alpha \ge 1$ (see Figure 1), and

$$(1-p^r-(1-p)^r)^{n-r} \le 1 \ ,$$

otherwise. The Stirling's formula (e.g. [10], p. 127) yields to:

$$\binom{n}{r} p^{\binom{r}{2}} \sim \left(\frac{nep^{(r-1)/2}}{r}\right)^r \ . \tag{6}$$

Consequently,

$$\binom{n}{r_u} p^{\binom{r_u}{2}} \to 1 \qquad and \qquad \binom{n}{r_u+1} p^{\binom{r_u+1}{2}} \sim \frac{\log_b n}{n} \to 0$$

The rest follows from the Markov's inequality (2) for $t=1$. $\diamondsuit$

To obtain conditions for the existence of dominating cliques in random graphs it is necessary to estimate the variance $Var(X_r)$. However, it seems to be difficult problem. We can use the fact that the clique number in random graphs is bounded within the tight interval. This assumption leads to the simplification of the enumeration of the variance. Therefore, we introduce the following notations.

Recall that $b = 1/p$ and let

$$r_0 = \log_b n - 2\log_b \log_b n + \log_b 2 + \log_b \log_b e \ , \tag{7}$$

$$r_1 = 2\log_b n - 2\log_b \log_b n + 2\log_b e + 1 - 2\log_b 2 \ . \tag{8}$$

J. G. Kalbfleisch and D. W. Matula [6, 8] proved that a random graph from $\mathbb{G}(n,p)$ does not contain cliques of the order greater than $\lceil r_1 \rceil$ and less or equal than $\lfloor r_0 \rfloor$. (See also [2, 9, 10].)

Remark 1 *Note that the upper bounds r_u and r_1 are the same. The bound r_1 comes out from the fact that the expectation of the random variable which "counts" the number of r-node cliques is given by:*

$$\binom{n}{r} p^{\binom{r}{2}} (1-p^r)^{n-r} \ .$$

Claim 1 implies $(1-p^r)^{n-r} = 1 - 1/n^{1-\varepsilon} + O(n^{-3})$ for $r = (2-\varepsilon)\log_b n$, where $0 \le \varepsilon < 1$. Thus, the argument for estimation of r_1 is the same as in Lemma 2.

D. Olejár and E. Toman [9] used the bounds (7) and (8) to obtain an estimation of the number of cliques in random graphs. To obtain an estimation of the $Var(X_r)$ we will apply a similar approach. We also use the following property adopted from [9], pp. 501–502.

Claim 2 *Let $k = o(\sqrt{n})$, then:*

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1) = n^k \left(1 - \binom{k}{2}\frac{1}{n} + O\left(\frac{k^4}{n^2}\right)\right) \ .$$

The estimation of the variance $Var(X_r)$ follows.

Lemma 3 *Let p be fixed, $0 < p < 1$ and $\lfloor r_0 \rfloor \le r \le \lceil r_1 \rceil$. Let*

$$\beta = \min\{\ 2/3,\ \ -2\log_b(1-p)\ \} \ .$$

Then:

$$Var(X_r) = E(X_r)^2 \cdot O\left(\frac{(\log n)^3}{n^\beta}\right) \ . \tag{9}$$

Proof. We will use the following formula [9]:

$$Var(X) = E(X^2) - E^2(X) \ . \tag{10}$$

The symbol $E(X_r^2)$ stands for the expectation of the number of ordered pairs of dominating cliques in a random graph G. The expectation can be expressed in the following way:

$$E(X_r^2) = \sum_{j=0}^{r} \binom{n}{r}\binom{r}{j}\binom{n-r}{r-j} \cdot p^{2\binom{r}{2} - \binom{j}{2}} \times$$

$$\times(1-p^r-(1-p)^r)^{2n-4r+2j}\cdot\Pr[S_r^1,S_r^2]\ . \tag{11}$$

The equation (11) follows from the next analysis. The nodes of the first dominating clique S_r^1 can be chosen in $\binom{n}{r}$ ways. The dominating cliques S_r^1, S_r^2 can be (but need not) have j common nodes. These nodes can be chosen in $\binom{r}{j}$ ways. The remaining $(r-1)$ nodes of the second dominating clique S_r^2 have to be chosen from $(n-r)$ nodes of $V(G)\setminus V(S_r^1)$. Now we shall choose edges: both dominating cliques are r-node complete graphs and therefore they contain $2\binom{r}{2}$ edges. But S_r^1, S_r^2 can have a nonempty intersection - a complete j-node subgraph. Therefore $\binom{j}{2}$ edges were counted twice. Both subgraphs S_r^1, S_r^2 are dominating cliques and so all $n-2r+j$ nodes of the set $V(G)\setminus[V(S_r^1)\cup V(S_r^2)]$ are "good" with respect to both S_r^1, S_r^2. The last term, $\Pr[S_r^1,S_r^2]$ denotes the probability that the nodes of $V(S_r^1)\setminus V(S_r^2)$ are good with respect to S_r^2 and the nodes of $V(S_r^2)\setminus V(S_r^1)$ are good with respect to S_r^1. It is sufficient to estimate $\Pr[S_r^1,S_r^2]$ by 1.

Since we need to prove that $Var(X_r)$ is asymptotically less than $E^2(X_r)$, we extract the expression $E^2(X_r)$ in front of the sum stated by the equation (11). We have:

$$E(X_r^2)\le E^2(X_r)\cdot\sum_{j=0}^{r}\binom{n}{r}^{-1}\binom{r}{j}\binom{n-r}{r-j}\cdot p^{-\binom{j}{2}}\cdot Q(p,r,j)\ , \tag{12}$$

where $Q(p,r,j)=(1-p^r-(1-p)^r)^{-2r+2j}$.

First we estimate the expression $Q(p,r,j)$. Let us denote $\alpha=-\log_b(1-p)$, as before. Recall that $(1-p)^r=p^{r\alpha}$. Let us also denote:

$$\nu=\min\{1,-\log_b(1-p)\}\ . \tag{13}$$

Therefore, from $\lfloor r_0\rfloor\le r\le\lceil r_1\rceil$ (cf. [9]), Claim 1 and (5), it follows:

$$Q(p,r,j)<[1-p^{r_0}-p^{\alpha r_0}]^{-2r}\le$$

$$\le\left[1-\frac{(\log_b n)^2}{2n\cdot\log_b e}-\left(\frac{(\log_b n)^2}{2n\cdot\log_b e}\right)^{\alpha}\right]^{-4\log_b n}=$$

$$=\exp\left\{4\log_b n\cdot\left[\frac{(\log_b n)^2}{2n\cdot\log_b e}+\left(\frac{(\log_b n)^2}{2n\cdot\log_b e}\right)^{\alpha}\right]\right\}\times$$

$$\times\left(1+O\left(\frac{(\log n)^{1+2\nu}}{n^{2\nu}}\right)\right)=$$

$$= \exp\left(\frac{2(\log_b n)^3}{n\cdot\log_b e}\right)\cdot\exp\left(\frac{4(\log_b n)^{2\alpha+1}}{(2n\cdot\log_b e)^{\alpha}}\right)\cdot\left(1+O\left(\frac{(\log n)^{1+2\nu}}{n^{2\nu}}\right)\right),$$

where $\nu = \min\{1,\alpha\}$. Since

$$\frac{2(\log_b n)^3}{n\cdot\log_b e}\to 0 \qquad \textit{and} \qquad \frac{4(\log_b n)^{2\alpha+1}}{(2n\cdot\log_b e)^{\alpha}}\to 0$$

as $n\to\infty$, the value of $Q(p,r,j)$ is $1+o(1)$ or, more precisely:

$$Q(p,r,j) = 1+O\left(\frac{(\log n)^{2\nu+1}}{n^{2\nu}}\right)\ . \tag{14}$$

Now we can concentrate our effort on the estimation of the sum

$$\sum_{j=0}^{r}\binom{n}{r}^{-1}\binom{r}{j}\binom{n-r}{r-j}\cdot p^{-\binom{j}{2}}\ , \tag{15}$$

where:

$$\lfloor r_0\rfloor \le r \le \lceil r_1\rceil\ .$$

We use the similar approach as D. Olejár and E. Toman in [9], pp. 504–506. This sum was also estimated in Subsection 5.3. of [10] (pp. 77–80), but we need more accurate calculation here. First we introduce the following notation:

$$S(n,r,c,d) = \sum_{j=c}^{d}\binom{n}{r}^{-1}\binom{r}{j}\binom{n-r}{r-j}\cdot b^{\binom{j}{2}}\ .$$

Our solution is based on the idea to divide the sum $S(n,r,a,b)$ into three parts by the following way:

$$S(n,r,0,r) \le S(n,r,0,1)+S(n,r,2,r_2)+S(n,r,r_2,r)\ , \tag{16}$$

where:

$$r_2 = (1+\lambda)\log_b n \qquad \textit{for} \qquad 0<\lambda<1\ .$$

A strict value of λ will be determined later. All these three parts will be estimated separately. Using Claim 2, the first part is estimated as follows:

$$S(n,r,0,1) = \binom{n-r}{r}\binom{n}{r}^{-1} + r\cdot\binom{n-r}{r-1}\binom{n}{r}^{-1} =$$

$$= \left(1-\frac{r^2}{n}\right)\left[1+O\left(\frac{(\log n)^4}{n^2}\right)\right]+\frac{r^2}{n}+O\left(\frac{(\log n)^3}{n^2}\right)=$$

$$= 1+O\left(\frac{(\log n)^4}{n^2}\right) \ . \tag{17}$$

To estimate the second part, it is necessary to analyze the binomial coefficients. (See also [10], pp. 79–80.)

$$\binom{n}{r}^{-1}\binom{r}{j}\binom{n-r}{r-j} = \frac{r!}{n^{\underline{r}}}\cdot\frac{r^{\underline{j}}}{j!}\cdot\frac{(n-r)^{\underline{r-j}}}{(r-j)!} =$$

$$= \frac{r^{\underline{j}}\cdot(r-j)!}{(r-j)!}\cdot\frac{r^{\underline{j}}}{j!}\cdot\frac{(n-r)^{\underline{r-j}}}{n^{\underline{j}}\cdot(n-j)^{\underline{r-j}}} \ \le\ \frac{r^{\underline{j}}\cdot r^{\underline{j}}}{j!\cdot n^{\underline{j}}} \ \le\ \frac{r^{2j}}{j!\cdot n^{\underline{j}}} \ \sim\ \frac{r^{2j}}{j!\cdot n^{j}}$$

We use the Stirling's formula in the following form:

$$j! \ \sim\ \left(\frac{j}{e}\right)^j \ .$$

Consequently,

$$\binom{n}{r}^{-1}\binom{r}{j}\binom{n-r}{r-j}\cdot b^{\binom{j}{2}} \ \sim\ \left(\frac{r^2\cdot b^{j/2}\cdot e}{j\cdot n\cdot\sqrt{b}}\right)^j \ . \tag{18}$$

The members of the sum $S(n,r,2,r_2)$ attain their asymptotic maximum for $j=r_2$. More precisely, letting $j=r_2=(1+\lambda)\log_b n$ we have:

$$\frac{r^2\cdot b^{j/2}\cdot e}{j\cdot n\cdot\sqrt{b}} = O\left(\frac{\log n}{n^{1/2-\lambda/2}}\right) \ .$$

Thus,

$$S(n,r,2,r_2) \le \left(\frac{c_1\cdot\log n}{n^{1/2-\lambda/2}}\right)^2+\left(\frac{c_1\cdot\log n}{n^{1/2-\lambda/2}}\right)^3+\ldots+\left(\frac{c_1\cdot\log n}{n^{1/2-\lambda/2}}\right)^{r_2}$$

for a suitable constant c_1. It yields:

$$S(n,r,2,r_2) = O\left(\frac{(\log n)^2}{n^{1-\lambda}}\right) \ . \tag{19}$$

To estimate the sum $S(n,r,r_2,r)$ we extract the term $\binom{n}{r}^{-1} \cdot b^{\binom{r}{2}}$:

$$S(n,r,r_2,r) = \binom{n}{r}^{-1} \cdot b^{\binom{r}{2}} \cdot \sum_{j=r_2}^{r} \binom{r}{r-j}\binom{n-r}{r-j} \cdot p^{\binom{r}{2}-\binom{j}{2}} \ .$$

To obtain the upper bound on the right-hand side sum, we substitute $\lceil r_1 \rceil$ for r in its upper border and $\lceil r_1 \rceil + 1$ for r in all the summands. The reasoning of such a substitution is the assertion of Lemma 2 and Remark 1. We have:

$$S(n,r,r_2,r) \le \binom{n}{r}^{-1} \cdot b^{\binom{r}{2}} \cdot \sum_{j=r_2}^{\lceil r_1 \rceil} \binom{\lceil r_1 \rceil + 1}{\lceil r_1 \rceil + 1 - j}\binom{n - \lceil r_1 \rceil - 1}{\lceil r_1 \rceil + 1 - j} \cdot p^{\binom{\lceil r_1 \rceil + 1}{2}-\binom{j}{2}} \ .$$

Let us put $k = \lceil r_1 \rceil + 1 - j$. Consequently,

$$S(n,r,r_2,r) \le \tag{20}$$

$$\le \binom{n}{r}^{-1} \cdot b^{\binom{r}{2}} \cdot \sum_{k=1}^{\lceil r_1 \rceil - r_2 + 1} \binom{\lceil r_1 \rceil + 1}{k}\binom{n - \lceil r_1 \rceil - 1}{k} \cdot p^{k[\lceil r_1 \rceil - (k-1)/2]} \ .$$

Note that

$$\binom{\lceil r_1 \rceil + 1}{k}\binom{n - \lceil r_1 \rceil - 1}{k} \cdot p^{k[\lceil r_1 \rceil - (k-1)/2]} \le \left((\lceil r_1 \rceil + 1) \cdot np^{\lceil r_1 \rceil - (k-1)/2} \right)^k ,$$

and

$$\lceil r_1 \rceil - (k-1)/2 \ge \lceil r_1 \rceil/2 + r_2/2 =$$
$$= (3/2 + \lambda/2)\log_b n - \log_b \log_b n + O(1) \ .$$

It yields:

$$(\lceil r_1 \rceil + 1) \cdot np^{\lceil r_1 \rceil - (k-1)/2} = O\left(\frac{(\log n)^2}{n^{1/2+\lambda/2}} \right) \ . \tag{21}$$

According to (20) and (21),

$$S(n,r,r_2,r) \le \binom{n}{r}^{-1} \cdot b^{\binom{r}{2}} \cdot O\left(\frac{(\log n)^2}{n^{1/2+\lambda/2}} \right) \ .$$

The term $\binom{n}{r}^{-1} \cdot b^{\binom{r}{2}}$ can be estimated using the Stirling's formula. The estimation is the same as in the proof of Lemma 2, see (6). Thus,

$$\binom{n}{r_1}^{-1} b^{\binom{r_1}{2}} \to 1 \ ,$$

$$\binom{n}{r}^{-1} b^{\binom{r}{2}} \sim \frac{(\log_b n)^c}{n^c} \leq 1 \ ,$$

if $r = \lceil r_1 \rceil - c$, where $c \geq 1$. Hence,

$$S(n, r, r_2, r) = O\left(\frac{(\log n)^2}{n^{1/2+\lambda/2}}\right) \ . \tag{22}$$

Let us choose λ in such a way that denominators of the expressions (19) and (22) will be asymptotically equivalent. Namely, for $\lambda = 1/3$ it holds $1 - \lambda = 1/2 + \lambda/2$ and we have:

$$S(n, r, 2, r_2) = O\left(\frac{(\log n)^2}{n^{2/3}}\right) \ , \tag{23}$$

$$S(n, r, r_2, r) = O\left(\frac{(\log n)^2}{n^{2/3}}\right) \ . \tag{24}$$

Consequently, (16), (17), (23) and (24) imply:

$$S(n, r, 0, r) = 1 + O\left(\frac{(\log n)^2}{n^{2/3}}\right) \ . \tag{25}$$

Formulae (12),(14) and (25) lead to:

$$\begin{aligned} E(X_r^2) &= E^2(X_r) \cdot \left[1 + O\left(\frac{(\log n)^2}{n^{2/3}}\right)\right] \cdot \left[1 + O\left(\frac{(\log n)^{2\nu+1}}{n^{2\nu}}\right)\right] \\ &= E^2(X_r) \cdot \left[1 + O\left(\frac{(\log n)^3}{n^{\beta}}\right)\right] \ , \end{aligned}$$

where $\nu = \min\{\ 1, -\log_b(1-p)\ \}$ and $\beta = \min\{\ 2/3, \ -2\log_b(1-p)\ \}$.

Substituting into (10) we obtain the estimation of $Var(X_r)$. $\diamondsuit$

The following claim expresses the number of the dominating cliques in random graphs.

Lemma 4 *Let p, r and β be as before, and*

$$X_r = \binom{n}{r} p^{\binom{r}{2}} (1 - p^r - (1-p)^r)^{n-r} \times \left\{1 + O\left(\frac{(\log n)^3}{n^{\beta/2}}\right)\right\} \ . \tag{26}$$

With probability $1 - O\left((\log n)^{-3}\right)$, a random graph from $\mathbb{G}(n, p)$ contains X_r dominating cliques on r nodes.

Proof. It follows from the Chebyshev's inequality [5]: if $Var(X)$ exists, then:

$$\Pr[|X - E(X)| \geq t] \geq \frac{Var(X)}{t^2} , \qquad t > 0 .$$

Letting $t = E(X_r) \cdot (\log n)^3 \cdot n^{-\beta/2}$ and using Lemma 3, we obtain the assertion of Lemma 4. $\diamondsuit$

3 Statement of the main result

For $r > 1$, let Y_r be the random variable on $\mathbb{G}(n,p)$ which denotes the number of r-node cliques. According to [9],

$$Y_r = \binom{n}{r} p^{\binom{r}{2}} (1-p^r)^{n-r} \times \left\{ 1 + O\left(\frac{(\log n)^3}{\sqrt{n}}\right) \right\} . \tag{27}$$

The ratio X_r/Y_r expresses the relative number of dominating cliques to all cliques in $\mathbb{G}(n,p)$ and it attains the value within the interval [0, 1]. By analysis of cases whether X_r/Y_r tends to 0 or 1, we obtain the main result of this paper. The term "almost surely" stands for "with the probability approaching to 1 as $n \to \infty$".

Theorem 1 *Let $0 < p < 1$ be fixed, let r be an order of the clique such that $\lfloor r_0 \rfloor \leq r \leq \lceil r_1 \rceil$ and recall that $\mathbb{L}n$ denotes $\log_{1/(1-p)} n$. Let $\delta(n) : \mathbb{N} \to \mathbb{N}$ be an arbitrary slowly increasing function such that $\delta(n) = o(\log n)$ and let $G \in \mathbb{G}(n,p)$ be a random graph. Then it holds:*

1. *If $p > 1/2$, then an r-node clique is dominating in G almost surely;*
2. *If $p \leq (3-\sqrt{5})/2$, then an r-node clique is not dominating in G almost surely;*
3. *If $(3-\sqrt{5})/2 < p \leq 1/2$, then an r-node clique:*
   - *is dominating in G almost surely, if $r \geq \mathbb{L}n + \delta(n)$,*
   - *is not dominating in G almost surely, if $r \leq \mathbb{L}n - \delta(n)$,*
   - *is dominating with probability $\exp(-p^c)$ and it is nondominating with probability $1 - \exp(-p^c)$, where $r = \mathbb{L}n + O(1)$ and c is a suitable constant.*

Proof. Let us examine the limit value of the ratio X_r/Y_r:

$$\frac{X_r}{Y_r} = \left(\frac{1-p^r-(1-p)^r}{1-p^r}\right)^{n-r} \times$$

$$\times\left\{1+O\left(\frac{(\log n)^3}{\sqrt{n}}\right)\right\} \times \left\{1+O\left(\frac{(\log n)^3}{n^{\beta/2}}\right)\right\} \ . \qquad (28)$$

The most important term of the expression (28) is the first one, since the last two terms both tend to 1 as $n \to \infty$. Recall that $(1-p)^r = p^{r\alpha}$, where $\alpha = -\log_{1/p}(1-p)$. According to Claim 1 and (5) we have:

$$\left(\frac{1-p^r-(1-p)^r}{1-p^r}\right)^{n-r} = \left(1-\frac{p^{r\alpha}}{1-p^r}\right)^{n-r} =$$

$$= \exp\left(\frac{-np^{r\alpha}}{1-p^r}\right)\cdot\left\{1+O\left(np^{2r\alpha}\right)\cdot\left[1+O\left(\frac{(\log n)^{2+\alpha}}{n}\right)\right]\right\} =$$

$$= \exp\left(\frac{-np^{r\alpha}}{1-p^r}\right)\cdot\left[1+O\left(np^{2r\alpha}\right)\right] \ .$$

The ratio X_r/Y_r expresses the relative number of dominating cliques to all cliques in $\mathbb{G}(n,p)$ and it attains the value within the interval [0, 1]. By analysis of whether X_r/Y_r tends to 0 or 1, we obtain the following two cases:

1. $np^{r\alpha} \to \infty$, or
2. $np^{r\alpha} \to 0$.

The inequality $\lfloor r_0 \rfloor \le r \le \lceil r_1 \rceil$ yields:

$$\left(\frac{(\log_b n)^2}{n^2}\right)^{\alpha} \le p^{r\alpha} \le \left(\frac{(\log_b n)^2}{n}\right)^{\alpha} \ .$$

Let us consider $\alpha = \alpha(p) = -\log_{1/p}(1-p)$ as a function on p, where $0 < p < 1$. (See Figure 1.) We have:

- $X_r/Y_r \to 1$ if $p > 1/2$;
  (In this case $\alpha > 1$ and $np^{r\alpha} \to 0$.)

- $X_r/Y_r \to 0$ if $p \le (3-\sqrt{5})/2$.
  (In this case $\alpha \le 1/2$, hence $p \le (1-p)^2$ and $np^{r\alpha} \to \infty$.)

Finally, we examine the case $(3-\sqrt{5})/2 < p \le 1/2$ separately. Letting $r = \mathbb{L}n \pm \lambda$ for an arbitrary constant $\lambda \ge 0$, we have:

$$p^{r\alpha} = \left(\frac{1}{p}\right)^{\log_b n^{-1}} \cdot p^{\pm\alpha\cdot\lambda} = n^{-1} \cdot p^{\pm\alpha\cdot\lambda} ,$$

since $\alpha \cdot \mathbb{L}n = \log_b n$ for $b = 1/p$. Thus, $X_r/Y_r \to \exp(-p^{\pm\alpha\cdot\lambda})$, since $np^{r\alpha} \to p^{\pm\alpha\cdot\lambda}$. This idea can be extended by letting $r \sim \mathbb{L}n \pm \delta(n)$, where $\delta(n) = o(\log n)$ such that $\delta(n) \to \infty$ as $n \to \infty$. Accordingly, we obtain the following cases:

- $X_r/Y_r \to 1$ if $r \ge \mathbb{L}n + \delta(n)$,
- $X_r/Y_r \to 0$ if $r \le \mathbb{L}n - \delta(n)$,
- $X_r/Y_r \to \exp(-p^c)$ if $r = \mathbb{L}n + O(1)$ and c is a suitable constant.

The proof is complete. $\diamondsuit$

4 Conclusions and open problems

We have claimed the conditions for the existence of dominating cliques in random graphs. In particular, we have showed that for every fixed $p > 1/2$, a random graph from $\mathbb{G}(n,p)$ contains a dominating clique almost surely. We have also calculated the number of dominating cliques in random graphs.

For the further works, we pose the following open questions:

- whether the assertion of the Theorem 1 can be generalized also for each $p > (3-\sqrt{5})/2 \approx 0.382$, and
- whether a random graph from $\mathbb{G}(n,p)$ contains at least one dominating clique of the order equal to its domination number.

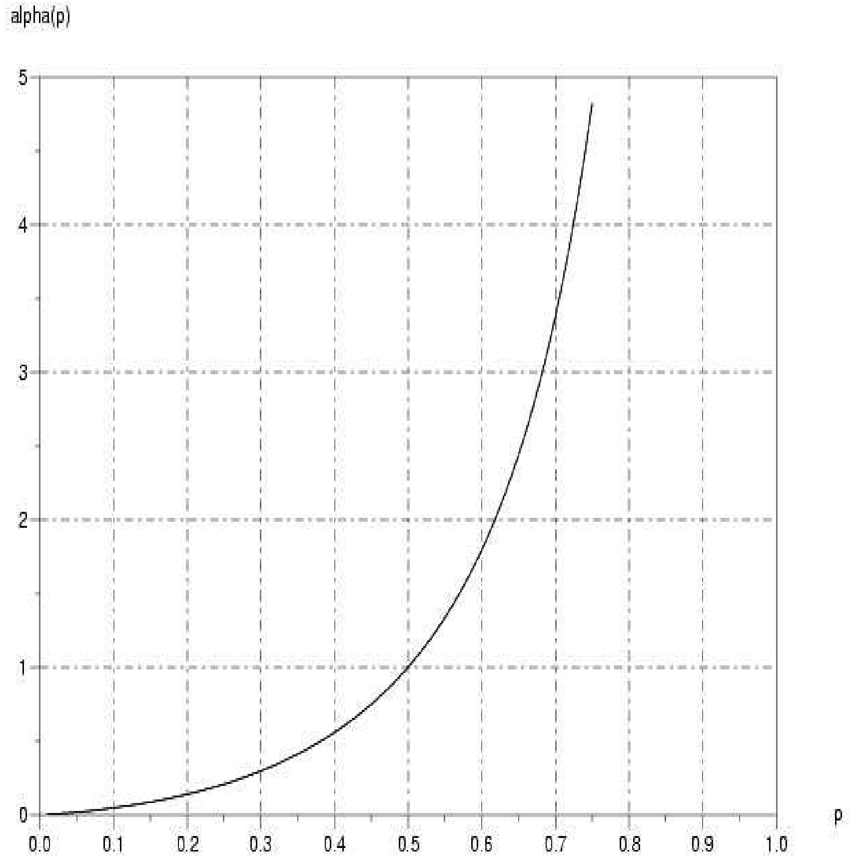

Figure 1: Graph of the function $\alpha(p) = -\log_{1/p}(1-p)$.

References

[1] B. Bollobás: *Random Graphs (2nd edition)*, Cambridge Studies in Advanced Mathmatics 73, 2001.

[2] B. Bollobás, P. Erdös: *Cliques in random graphs*, Math. Proc. Cam. Phil. Soc. (1976), 80, 419–427.

[3] M. R. Garey, D.S. Johnson: *Computers and Intractability*, Freeman, New York, 1979.

[4] J. L. Gross, J. Yellen: *Handbook of Graph Theory*, CRC Press, 2003.

[5] S. Janson, T. Luczak, A. Rucinski: *Random Graphs*, John Wiley & Sons, New York, 2000.

[6] J. G. Kalbfleisch: *Complete subgraphs of random hypergraphs and bipartite graphs*, In Proc. 3rd Southeastern Conf. of Combinatorics, Graph Theory and Computing, Florida Atlantic University, 1972, 297–304.

[7] R. M. Karp: *Reducibility among combinatorial problems*, In Complexity of Computer Computation, (R. E. Miller and J. W. Thatcher, eds.), Plenum Press, 1972. 24, 85–103.

[8] D. W. Matula: The largest clique size in a random graph, Technical report CS 7608, Dept. of Comp. Sci. Southern Methodist University, Dallas, 1976.

[9] D. Olejár, E. Toman: *On the Order and the Number of Cliques in a Random Graph*, Math. Slovaca, **47** (1997), No. 5, 499–510.

[10] E. M. Palmer: *Graphical Evolution*, John Wiley & Sons, Inc., New York, 1985.

[11] B. Wieland, A. P. Godbole: *On the Domination Number of a Random Graph*, Electronic Journal of Combinatorics, **8**, No. 1, #R37, 2001.