On the Domination Number of a Random Graph

Ben Wieland
Department of Mathematics
University of Chicago
wieland@math.uchicago.edu

Anant P. Godbole
Department of Mathematics
East Tennessee State University
godbolea@etsu.edu

Submitted: May 2, 2001; Accepted: October 11, 2001.
MR Subject Classifications: 05C80, 05C69

Abstract

In this paper, we show that the domination number D of a random graph enjoys as sharp a concentration as does its chromatic number χ. We first prove this fact for the sequence of graphs $\{G(n, p_n\}$, $n \to \infty$, where a *two point concentration* is obtained with high probability for $p_n = p$ (fixed) or for a sequence p_n that approaches zero sufficiently slowly. We then consider the infinite graph $G(\mathbb{Z}^+, p)$, where p is fixed, and prove a *three* point concentration for the domination number with probability one. The main results are proved using the second moment method together with the Borel Cantelli lemma.

1 Introduction

A set γ of vertices of a graph $G = (V, E)$ constitutes a *dominating set* if each $v \in V$ is either in γ or is adjacent to a vertex in γ. The *domination number* D of G is the size of a dominating set of smallest cardinality. Domination has been the subject of extensive research; see for example Section 1.2 in [1], or the texts [6], [7]. In a recent Rutgers University dissertation, Dreyer [3] examines the question of domination for *random graphs*, motivated by questions in search structures for protein sequence libraries. Recall that the random graph $G(n, p)$ is an ensemble of n vertices with each of the potential $\binom{n}{2}$ edges being inserted independently with probability p, where p often approaches zero as $n \to \infty$. The treatises of Bollobás [2] and Janson et al. [8] between them cover the theory of random graphs in admirable detail. Dreyer [3] generalizes some results of Nikoletseas and Spirakis [5] and proves that with $q = 1/(1-p)$ (p fixed) and for any $\varepsilon > 0$, any fixed set of cardinality $(1+\varepsilon)\log_q n$ is a dominating set with probability approaching unity as $n \to \infty$, and that sets of size $(1-\varepsilon)\log_q n$ dominate with probability approaching zero ($n \to \infty$). The elementary proofs of these facts reveal, moreover, that rather than having ε fixed, we may instead take $\varepsilon = \varepsilon_n$ tending to zero so that $\varepsilon_n \log_q n \to \infty$. It follows from the first of these results that the domination number of $G(n, p)$ is no larger than $\lceil \log_q n + a_n \rceil$ with probability approaching unity – where a_n is *any* sequence that approaches infinity. This is because

$$\begin{aligned}
\mathbb{P}(D \le \lceil \log_q n + a_n \rceil) &= \mathbb{P}(\exists \text{ a dominating set of size } r := \lceil \log_q n + a_n \rceil) \\
&\ge \mathbb{P}(\{1, 2, \ldots, r\} \text{ is a dominating set}) \\
&= (1 - (1-p)^r)^{n-r} \\
&\ge 1 - (n-r)(1-p)^r \\
&\ge 1 - n(1-p)^r \\
&\ge 1 - n(1-p)^{\log_q n + a_n} \\
&= 1 - (1-p)^{a_n} \\
&\to 1.
\end{aligned}$$

In this paper, we sharpen this result, showing that the domination number D of a random graph enjoys as sharp a concentration as does its chromatic number χ [1]. In Section 2, we prove this fact for the sequence of graphs $\{G(n, p_n\}$, $n \to \infty$, where a *two point concentration* is obtained with high probability (*w.h.p.*) for $p_n = p$ (fixed) or for a sequence p_n that approaches zero sufficiently slowly. In Section 3, on the other hand, we consider the infinite graph $G(\mathbb{Z}^+, p)$, where p is fixed, and prove a *three* point concentration for the domination number with probability one (i.e., in the almost everywhere sense of measure theory.) The main results are proved using the so-called *second moment method* [1] together with the Borel Cantelli lemma from probability theory. We consider our results to be interesting, particularly since the problem of determining domination numbers is known to be NP-complete, and since very little appears to have been done in the area of domination for random graphs (see, e.g., [4] in addition to [3],[5].)

2 Two Point Concentration

For $r \ge 1$, let the random variable X_r denote the number of dominating sets of size r. Note that

$$X_r = \sum_{j=1}^{\binom{n}{r}} I_j,$$

where I_j equals one or zero according as the j^{th} set of size r forms or doesn't form a dominating set, and that the expected value $\mathbb{E}(X_r)$ of X_r is given by

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

We first analyze (1) on using the easy estimates $\binom{n}{r} \le (ne/r)^r$ and $1 - x \le \exp(x)$ to get

$$\begin{aligned}
\mathbb{E}(X_r) &\le \left(\frac{ne}{r}\right)^r \exp\{-(n-r)(1-p)^r\} \\
&= \exp\{-n(1-p)^r + r(1-p)^r + r + r\log n - r\log r\}. \qquad (2)
\end{aligned}$$

Here and throughout this paper, we use log to denote the natural logarithm. Note that the right hand side of (2) makes sense even if $r \notin \mathbb{Z}^+$, and that it can be checked to be an increasing function of r by verifying that its derivative is non-negative for $r \leq n$. Keeping these facts in mind, we next denote $\log_{1/(1-p)} n$ (for fixed p) by $\mathbb{L}n$ and note that with $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))$ the exponent in (2) can be bounded above as follows:

$$\begin{aligned} & \exp\left\{-n(1-p)^r + r(1-p)^r + r + r\log n - r\log r\right\} \\ \leq\ & \exp\left\{-n(1-p)^r + 2r + r\log n - r\log r\right\} \\ \leq\ & \exp\{2\mathbb{L}n - 2\mathbb{L}((\mathbb{L}n)(\log n)) - (\log n)\mathbb{L}((\mathbb{L}n)(\log n)) \\ & \ - r\log r\} \\ \to\ & 0 \quad (n \to \infty). \end{aligned} \tag{3}$$

It follows from (3) that with $r = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) \rfloor$ and D_n denoting the domination number, we have

$$\mathbb{P}(D_n \leq r) = \mathbb{P}(X_r \geq 1) \leq \mathbb{E}(X_r) \to 0 \ \ (n \to \infty).$$

We have thus proved

Lemma 1 *The domination number D_n of the random graph $G(n,p)$ satisfies, for fixed p,*

$$\mathbb{P}(D_n \geq \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) \rfloor + 1) \to 1 \ \ (n \to \infty).$$

The values of $\mathbb{L}n$ tend to get somewhat large if $p \to 0$. For example, if $p = 1 - 1/e$, then $\mathbb{L}n = \log n$, but with $p = 1/n$, $\mathbb{L}n \approx n\log n$, where, throughout this paper, we write $a_n \approx b_n$ if $a_n/b_n \to 1$ as $n \to \infty$. In general, for $p \to 0, \mathbb{L}(\cdot) \approx \log(\cdot)/p$. If the argument leading to (3) is to be generalized, we clearly need $r := \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) \geq 1$ so that $r\log r \geq 0$; note that r may be negative if, e.g., $p = 1/n$. One may check that $r \geq 1$ if $p \geq e\log^2 n/n$. It is not too hard to see, moreover, that the argument leading to (3) is otherwise independent of the magnitude of p (since $(\log n)\mathbb{L}((\mathbb{L}n)(\log n))$ always far exceeds $2\mathbb{L}n$), so that we have

Lemma 2 *The conclusion of Lemma 1 holds for each sequence of graphs $G(n,p_n)$ with $p_n \geq e\log^2 n/n$.*

We next continue with the analysis of the expected value $\mathbb{E}(X_r)$. Throughout this paper, we will use the notation $o(1)$ to denote a *generic* function that tends to zero with n. Also, given non-negative sequences a_n and b_n, we will write $a_n \gg b_n$ (or $b_n \ll a_n$) to mean $a_n/b_n \to \infty$ as $n \to \infty$. Returning to (1), we see on using the estimate $1 - x \geq \exp\{-x/(1-x)\}$ that for $r \geq 1$,

$$\begin{aligned} \mathbb{E}(X_r) &= \binom{n}{r}\left(1-(1-p)^r\right)^{n-r} \\ &\geq \binom{n}{r}\left(1-(1-p)^r\right)^{n} \\ &\geq (1-o(1))\frac{n^r}{r!}\exp\left\{-\frac{n(1-p)^r}{1-(1-p)^r}\right\}, \end{aligned} \tag{4}$$

where the last estimates in (4) hold provided that $r^2 = o(n)$, which is a condition that is certainly satisfied if p is fixed (and in general if $p \gg \log n/\sqrt{n}$) and $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon$, where the significance of the arbitrary $\varepsilon > 0$ will become clear in a moment[1]. Assume that $p \gg \log n/\sqrt{n}$ and set $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon$, i.e., a mere ε more than the value $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))$ ensuring that "$\mathbb{E}$"$(X_r) \to 0$. We shall show that this choice forces the right hand side of (4) to tend to infinity. Stirling's approximation yields,

$$\begin{aligned}
& (1-o(1))\frac{n^r}{r!}\exp\left\{-\frac{n(1-p)^r}{1-(1-p)^r}\right\} \\
\geq\ & (1-o(1))\left(\frac{ne}{r}\right)^r\frac{1}{\sqrt{2\pi r}}\exp\left\{-\frac{n(1-p)^r}{1-(1-p)^r}\right\} \\
\geq\ & (1-o(1))\exp\left\{A-B\right\}, \qquad (5)
\end{aligned}$$

where

$$A = (\log n)(\mathbb{L}n)\left\{1 - \frac{(1-p)^\varepsilon}{1-\frac{(1-p)^\varepsilon \mathbb{L}n \log n}{n}}\right\} + \mathbb{L}n$$

and

$$B = \mathbb{L}(\mathbb{L}n\log n) + (\log n)\mathbb{L}(\mathbb{L}n\log n) + \mathbb{L}n\log(\mathbb{L}n) + K + \log(\mathbb{L}n)/2,$$

where $K = \log\sqrt{2\pi}$. We assert that the right side of (5) tends to infinity for all positive values of ε provided that p is fixed or else tends to zero at an appropriately slow rate. Some numerical values may be useful at this point. Using $p = 1-(1/e)$ and $\mathbb{E}(X_r) \approx (ne/r)^r\exp\{-ne^{-r}\}$, Rick Norwood has computed that with $n = 100{,}000$, $\mathbb{E}(X_7) = 3.26\cdot 10^{-8}$, while $\mathbb{E}(X_8) = 4.8\cdot 10^{21}$. Since $p \gg \log n/\sqrt{n}$ and $\mathbb{L}n \approx \log n/p$, we see that $p \gg \mathbb{L}n\log n/n$ and thus that for large n,

$$A \geq \log n\mathbb{L}n\left\{1-\frac{(1-p)^\varepsilon}{1-\varepsilon p(1-p)^\varepsilon}\right\} + \mathbb{L}n.$$

For specificity, we now set $\varepsilon = 1/2$ and use the estimate $1-\sqrt{1-x} \geq x/2$, which implies that for large n

$$\begin{aligned}
A \ &\geq\ (\log n)(\mathbb{L}n)\left\{1-\frac{(1-p)^\varepsilon}{1-\varepsilon p(1-p)^\varepsilon}\right\} + \mathbb{L}n \\
&=\ (\log n)(\mathbb{L}n)\left\{\frac{1}{1-\varepsilon p(1-p)^\varepsilon} - \frac{(1-p)^\varepsilon}{1-\varepsilon p(1-p)^\varepsilon} - \frac{\varepsilon p(1-p)^\varepsilon}{1-\varepsilon p(1-p)^\varepsilon}\right\} + \mathbb{L}n \\
&\geq\ (\log n)(\mathbb{L}n)\frac{\varepsilon p[1-(1-p)^\varepsilon]}{1-\varepsilon p(1-p)^\varepsilon} + \mathbb{L}n \\
&\geq\ (\log n)(\mathbb{L}n)\frac{p^2\varepsilon^2}{1-\varepsilon p(1-p)^\varepsilon} + \mathbb{L}n
\end{aligned}$$

[1]Recall that we will find it beneficial to continue to plug in a non-integer value for r on the right side of an equation such as (4), fully realizing that $\mathbb{E}(X_r)$ makes no sense. In such cases, the notation "$\mathbb{E}$"(X_r), "$\mathbb{V}$"(X_r) etc. will be used

$$\geq \ (\log n)(\mathbb{L}n)p^2\varepsilon^2 + \mathbb{L}n = \frac{(\log n)(\mathbb{L}n)p^2}{4} + \mathbb{L}n := C.$$

The choice of $\varepsilon = 1/2$ has its drawbacks as we shall see; it is the main reason why a two point concentration (rather than a far more desirable one point concentration) will be obtained at the end of this section. The problem is that $\mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))$ may be arbitrarily close to an integer, so that we might, in our quest to have

$$\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon\rfloor,$$

be forced to deal with a sequence of ε's that tend to zero with n. From now on, we shall take $\varepsilon = 1/2$ unless it is explicitly specified to be different. We shall show that $C/10$ exceeds each of the five quantities that constitute B, so that

$$\exp\{A - B\} \geq \exp\{C - B\} \geq \exp\{C/2\} \to \infty.$$

It is clear that we only need focus on the case $p \to 0$. Also, it is evident that for large n, $C/10 \geq K = \log\sqrt{2\pi}$ and $C/10 \geq \log(\mathbb{L}n)/2$. Next, note that the second term in B dominates the first, so that we need to exhibit the fact that

$$C/10 \geq (\log n)\mathbb{L}(\mathbb{L}n \log n). \tag{6}$$

Since $\mathbb{L}(\cdot) \approx \log(\cdot)/p$, (6) reduces to

$$\frac{p\log^2 n}{40} + \frac{\log n}{10p} \geq \log n \mathbb{L}(\frac{\log^2 n}{p}),$$

and thus to

$$\frac{p\log n}{40} + \frac{1}{10p} \geq \frac{1}{p}\log(\frac{\log^2 n}{p}).$$

(6) will thus hold provided that

$$\frac{p\log n}{40} \geq \frac{1}{p}\log(\frac{\log^2 n}{p}),$$

or if

$$\frac{p^2}{40} \geq \frac{\log\left(\frac{\log^2 n}{p}\right)}{\log n},$$

a condition that is satisfied if p is not too small, e.g., if $p = 1/\log\log n$. Finally, the condition $C/10 \geq \mathbb{L}n\log(\mathbb{L}n)$ may be checked to hold for large n provided that

$$\frac{p^2\log n}{40} \geq \log\left(\frac{\log n}{p}\right),$$

or if

$$\frac{p^2}{40} \geq \frac{\log\left(\frac{\log n}{p}\right)}{\log n},$$

and is thus satisfied if (6) is.

It is easy to check that the derivative (with respect to r) of the right hand side of (5) is non-negative if r is not too close to n, e.g., if $r^2 \ll n$, so that

$$\begin{aligned}\mathbb{E}(X_{\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor+2}) &\geq \text{ right side of } (5)|_{r=\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor+2} \\ &\geq \text{ right side of } (5)|_{r=\mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))+\varepsilon} \\ &\to \infty.\end{aligned}$$

The above analysis clearly needs that the condition $r^2 \ll n$ be satisfied. This holds for $p \gg \log n/\sqrt{n}$ and $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + K$, where K is any constant. Now the condition

$$\frac{p^2}{40} \geq \frac{\log\left(\frac{\log^2 n}{p}\right)}{\log n},$$

ensuring the validity of (6) is certainly weaker than the condition $p \gg \log n/\sqrt{n}$. We have thus proved:

Lemma 3 *The expected number $\mathbb{E}(X_r)$ of dominating sets of size r of the random graph $G(n,p)$ tends to infinity if p is either fixed or tends to zero sufficiently slowly so that $p^2/40 \geq [\log\left((\log^2 n)/p\right)]/\log n$, and if $r \geq \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 2$.*

It would be most interesting to see how rapidly the expected value of X_r changes from zero to infinity if p is smaller than required in Lemma 3. A related set of results, to form the subject of another paper, can be obtained on using a more careful analysis than that leading to Lemma 3 – with the focus being on allowing ε to get as large as needed to yield $\mathbb{E}(X_r) \to \infty$.

We next need to obtain careful estimates on the variance $\mathbb{V}(X_r)$ of the number of r-dominating sets. We have

$$\begin{aligned}\mathbb{V}(X_r) &= \sum_{j=1}^{\binom{n}{r}} \mathbb{E}(I_j)\left\{1 - \mathbb{E}(I_j)\right\} + 2\sum_{j=1}^{\binom{n}{r}}\sum_{j<i}\left\{\mathbb{E}(I_iI_j) - \mathbb{E}(I_i)\mathbb{E}(I_j)\right\} \\ &= \binom{n}{r}\rho + \binom{n}{r}\sum_{s=0}^{r-1}\binom{r}{s}\binom{n-r}{r-s}\mathbb{E}(I_1I_s) - \binom{n}{r}^2\rho^2, \qquad (7)\end{aligned}$$

where $\rho = \mathbb{E}(I_1) = (1-(1-p)^r)^{n-r}$ and I_s is any generic r-set that intersects the 1^{st} r-set in s elements. Now, on denoting the 1^{st} and s^{th} r-sets by A and B respectively, we have

$$\begin{aligned}\mathbb{E}(I_1I_s) &= \mathbb{P}(A \text{ dominates and } B \text{ dominates}) \\ &\leq \mathbb{P}(A \text{ dominates } (\widetilde{A\cup B}) \text{ and } B \text{ dominates } (\widetilde{A\cup B})) \\ &= \mathbb{P}(\text{each } x \in \widetilde{A\cup B} \text{ has a neighbour in } A \text{ and in } B) \\ &= \left(1 - 2(1-p)^r + (1-p)^{2r-s}\right)^{n-2r+s}. \qquad (8)\end{aligned}$$

In view of (7) and (8), we have

$$\begin{aligned}\mathbb{V}(X_r) &= \binom{n}{r}\rho - \binom{n}{r}^2\rho^2 \\ &\quad + \binom{n}{r}\sum_{s=0}^{r-1}\binom{r}{s}\binom{n-r}{r-s}\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}. \qquad (9)\end{aligned}$$

We claim that the $s=0$ term in (9) is the one that dominates the sum. Towards this end, note that the difference between this term and the quantity $\binom{n}{r}^2\rho^2$ may be bounded as follows:

$$\begin{aligned}&\binom{n}{r}\binom{n-r}{r}\left(1-(1-p)^r\right)^{2(n-2r)} - \binom{n}{r}^2\left(1-(1-p)^r\right)^{2n-2r} \\ &= \binom{n}{r}^2\rho^2\left\{\frac{\binom{n-r}{r}}{\binom{n}{r}}\left(1-(1-p)^r\right)^{-2r}-1\right\} \\ &\le \binom{n}{r}^2\rho^2\left(e^{-r^2/n}\exp\left(\frac{2r(1-p)^r}{1-(1-p)^r}\right)-1\right) \\ &= \binom{n}{r}^2\rho^2\left(\exp\left(-\frac{r^2}{n}+2r(1-p)^r(1+o(1))\right)-1\right), \qquad (10)\end{aligned}$$

where the last estimate in (10) holds due to the fact that $(1-p)^r \to 0$ if $r = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon$ and $p \gg \log^2 n/n$ – which are both facts that have been assumed. Note also that

$$2r(1-p)^r(1+o(1)) > \frac{r^2}{n}$$

holds if

$$2(\mathbb{L}n)\log n \gg \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon$$

is true; the latter condition may be checked to hold for all reasonable choices of p. It follows that the exponent in (10) is non-negative. Furthermore, $r(1-p)^r \to 0$ since $p \gg \log^{3/2} n/\sqrt{n}$. We thus have from (10)

$$\binom{n}{r}\binom{n-r}{r}\left(1-(1-p)^r\right)^{2(n-2r)} - \binom{n}{r}^2\left(1-(1-p)^r\right)^{2n-2r} = o([\mathbb{E}(X_r)]^2). \qquad (11)$$

Next define

$$f(s) = \binom{r}{s}\binom{n-r}{r-s}\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s};$$

we need to estimate $\sum_{s=1}^{r-1} f(s)$. We have

$$f(s) \le \binom{r}{s}\frac{n^{r-s}}{(r-s)!}\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}$$

$$
\begin{aligned}
&\leq\ 2\binom{r}{s}\frac{n^{r-s}}{(r-s)!}\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^n \\
&\leq\ 2\binom{r}{s}\frac{n^{r-s}}{(r-s)!}\exp\left\{n\left((1-p)^{2r-s}-2(1-p)^r\right)\right\} =: g(s), \qquad (12)
\end{aligned}
$$

where the next to last inequality above holds due to the assumption that $p \gg \log^{3/2} n/\sqrt{n}$. Consider the rate of growth of g as manifested in the ratio of consecutive terms. By (12),

$$
\frac{g(s+1)}{g(s)} = \frac{(r-s)^2}{n(s+1)}\exp\left\{np(1-p)^{2r-s-1}\right\} =: h(s). \qquad (13)
$$

We claim that $h(s) \geq 1$ iff $s \geq s_0$ for some $s_0 = s_0(n) \to \infty$, so that g is first decreasing and then increasing. We shall also show that $g(1) \geq g(r-1)$, which implies that $\sum_{s=1}^{r-1} f(s) \leq rg(1)$. First note that

$$
\begin{aligned}
h(1) \ &\leq\ \frac{r^2}{2n}\exp\left\{\frac{np}{(1-p)^2}(1-p)^{2r}\right\} \\
&=\ \frac{r^2}{2n}\exp\left\{\frac{p}{n(1-p)^{2-2\varepsilon}}(\mathbb{L}n\log n)^2\right\} \to 0
\end{aligned}
$$

since $p \gg \log n/\sqrt{n}$, and that

$$
\begin{aligned}
h(r-1) \ &\approx\ \frac{1}{nr}\exp\left\{(1-p)^{\varepsilon}\log^2 n\right\} \\
&\approx\ \frac{p}{n\log n}\exp\left\{(1-p)^{\varepsilon}\log^2 n\right\} \\
&\geq\ \frac{1}{n^{3/2}}\exp\left\{(1-p)^{\varepsilon}\log^2 n\right\} \geq 1
\end{aligned}
$$

provided that p is not of the form $1-o(1)$. Now,

$$
h(s) = \frac{(r-s)^2}{n(s+1)}\exp\left\{np(1-p)^{2r-s-1}\right\} \geq 1
$$

iff

$$
\exp\left\{\frac{p(1-p)^{-s-1+2\varepsilon}(\mathbb{L}n\log n)^2}{n}\right\} \geq \frac{n(s+1)}{(r-s)^2}
$$

iff

$$
(1-p)^{s+1}\log\frac{n(s+1)}{(r-s)^2} \leq p(1-p)^{2\varepsilon}\frac{(\mathbb{L}n\log n)^2}{n}
$$

iff

$$
(1-p)^{s+1-2\varepsilon}(\log n)(1+\delta(s)) \leq p\frac{(\mathbb{L}n\log n)^2}{n},
$$

$$
(\text{where } \delta(s) = \Theta(\log r/\log n))
$$

iff

$$
(s+1-2\varepsilon) = s \geq \frac{\log p + 2\log(\mathbb{L}n) + \log\log n - \log n - \log(1+\delta(s))}{\log(1-p)}. \qquad (14)
$$

First note that

$$\left|\frac{\log(1+\delta(s))}{\log(1-p)}\right| \approx \frac{\delta(s)}{p} \le \frac{2\log r}{p\log n} \le \frac{2\log(\mathbb{L}n)}{p\log n} \to 0$$

if $p \gg \log((\log n)/p)/\log n$, which is a weaker condition than (6). Also, since $\log n \gg \log\log n + 2\log(\mathbb{L}n)$, it follows that the right hand side of (14) is of the form $a_n + o(1)$, $a_n \to \infty$, so that $h(s) \ge 1$ iff $s \ge s_0$, as claimed. Note next that $g(1) \ge g(r-1)$ iff

$$\begin{aligned} &2r\frac{n^{r-1}}{(r-1)!}\exp\left\{n\left((1-p)^{2r-1} - 2(1-p)^r\right)\right\} \\ &\quad \ge 2nr\exp\left\{n\left((1-p)^{r+1} - 2(1-p)^r\right)\right\}, \end{aligned}$$

i.e., if

$$\frac{n^{r-1}}{(r-1)!}\exp\left\{n\left((1-p)^{2r-1} - (1-p)^{r+1}\right)\right\} \ge n,$$

which in turn is satisfied provided that

$$\frac{n^{r-1}}{(r-1)!}\left(1-(1-p)^r\right)^n \ge n,$$

or if

$$\mathbb{E}(X_r) \ge \frac{n^2}{r}(1+o(1)).$$

The last condition above holds since $\mathbb{E}(X_r) \ge \exp\{C/2\}$, where $C = ((\log n)(\mathbb{L}n)p^2)/4 + \mathbb{L}n$ is certainly larger than (say) $6\log n$ if p is not too small, e.g., if $p \ge 24/\log n$. In conjunction with the fact that $h(1) < 1$ and $h(r-1) > 1$, (9) and (10) and the above discussion show that

$$\frac{\mathbb{V}(X_r)}{\mathbb{E}^2(X_r)} \le \frac{1}{\mathbb{E}(X_r)} + \left(2r(1-p)^r - \frac{r^2}{n}\right)(1+o(1)) + \frac{rg(1)\binom{n}{r}}{\mathbb{E}^2(X_r)}; \tag{15}$$

we will thus have $\mathbb{V}(X_r) = o(\mathbb{E}^2(X_r))$ if $\mathbb{E}(X_r) \to \infty$ provided that we can show that the last term on the right hand side of (15) tends to zero. We have

$$\begin{aligned} \frac{rg(1)\binom{n}{r}}{\mathbb{E}^2(X_r)} &\le \frac{2r^2n^{r-1}\exp\left\{n\left((1-p)^{2r-1} - 2(1-p)^r\right)\right\}}{(r-1)!\binom{n}{r}\rho^2} \\ &\le 3\frac{r^3}{n}\frac{\left(1-2(1-p)^r + (1-p)^{2r-1}\right)^n}{\left(1-2(1-p)^r + (1-p)^{2r}\right)^n} \\ &\le 3\frac{r^3}{n}\left(1 + \frac{(1-p)^{2r-1} - (1-p)^{2r}}{\left(1-(1-p)^r\right)^2}\right)^n \\ &\le 3\frac{r^3}{n}\exp\left\{\frac{np(1-p)^{2r-1}}{\left(1-(1-p)^r\right)^2}\right\} \\ &\le 3\frac{r^3}{n}\exp\left\{p\frac{\left((\mathbb{L}n)(\log n)\right)^2}{n}(1+o(1))\right\} \to 0, \end{aligned}$$

since $p \gg \log n/\sqrt[3]{n}$, establishing what is required. We are now ready to state our main result.

Theorem 4 *The domination number of the random graph $G(n,p)$; $p = p_n \geq p_0(n)$ is, with probability approaching unity, equal to $\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 1$ or $\lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 2$, where $p_0(n)$ is the smallest p for which*

$$p^2/40 \geq [\log\big((\log^2 n)/p\big)]/\log n$$

holds.

Proof By Chebychev's inequality, Lemma 3, and the fact that $\mathbb{V}(X_r) = o(\mathbb{E}^2(X_r))$ whenever $\mathbb{E}(X_r) \to \infty$,

$$\mathbb{P}(D_n > r) = \mathbb{P}(X_r = 0) \leq \mathbb{P}(|X_r - \mathbb{E}(X_r)|) \geq \mathbb{E}(X_r)) \leq \frac{\mathbb{V}(X_r)}{\mathbb{E}^2(X_r)} \to 0$$

if $r = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 2$. This fact, together with Lemmas 1 and 2, prove the required result. (Note: strictly speaking, we had shown above that "$\mathbb{V}$"$(X_s) \to \infty$ if $s = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + \varepsilon = \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n)) + 1/2$. The fact that $\mathbb{V}(X_r) \to \infty$ $(r = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 2)$ follows, however, since we could have taken $\varepsilon = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor + 2 - \mathbb{L}n + \mathbb{L}((\mathbb{L}n)(\log n))$ in the analysis above, and bounded all terms involving ε by noting that $1 \leq \varepsilon \leq 2$.)

3 Almost Sure Results

In this section, we show that one may, with little effort, derive a three point concentration for the domination number D_n of the subgraph $G(n,p)$ of $G(\mathbb{Z}^+,p)$, p fixed. Specifically, we shall prove

Theorem 5 *Consider the infinite random graph $G(\mathbb{Z}^+,p)$, where p is fixed. Let $\mathbb{P}$ be the measure induced on $\{0,1\}^\infty$ by an infinite sequence $\{X_n\}_{n=1}^\infty$ of Bernoulli (p) random variables, and denote the domination number of the induced subgraph $G(\{1,2,\ldots,n\},p)$ by D_n. Then, with $R_n = \lfloor \mathbb{L}n - \mathbb{L}((\mathbb{L}n)(\log n))\rfloor$,*

$$\mathbb{P}\left(1 \leq \liminf_{n\to\infty}(D_n - R_n) \leq \limsup_{n\to\infty}(D_n - R_n) \leq 3\right) = 1.$$

In other words, for almost all infinite sequences $\omega = \{X_n\}_{n=1}^\infty$ of p-coin flips, i.e., for all $\omega \in \Omega$; $\mathbb{P}(\Omega) = 1$, there exists an integer $N_0 = N_0(\omega)$ such that $n \geq N_0 \Rightarrow R_n + 1 \leq D_n \leq R_n + 3$, where D_n is the domination number of the induced subgraph $G(\{1,2,\ldots,n\},p)$.

Proof Equation (3) reveals that for fixed p,

$$\begin{aligned}\mathbb{P}(D_n \leq R_n) &\leq \mathbb{E}(X_{R_n})\\ &\leq \exp\{2\mathbb{L}n - 2\mathbb{L}((\mathbb{L}n)(\log n)) - (\log n)\mathbb{L}((\mathbb{L}n)(\log n))\\ &\quad - (1-o(1))\mathbb{L}n\log\mathbb{L}n\}. \end{aligned} \tag{16}$$

Since $\mathbb{L}n = K \log n$, the right hand side of (16) is asymptotic to

$$\exp\{-3K(1+o(1))\log n \log\log n\} = \frac{1}{n^{3K(1+o(1))\log\log n}}.$$

Thus

$$\sum_{n=1}^{\infty} \mathbb{P}(D_n \leq R_n) < \infty,$$

which proves, via the Borel-Cantelli lemma, that

$$\mathbb{P}(D_n \leq R_n \text{ infinitely often}) = 0. \tag{17}$$

Unfortunately, however, the analysis in Section 2 only gives

$$\mathbb{P}(D_n \geq R_n + 3) = O\left(\frac{\log^3 n}{n}\right),$$

so that we may only conclude (here we are launching the standard "subsequence" argument for proving almost sure results in probability theory) that

$$\mathbb{P}(D_{n^2} \geq R_{n^2} + 3 \text{ infinitely often}) = 0. \tag{18}$$

Using (18), we take any S with $|S| = R_{n^2} + 2$ that dominates $G(n^2, p)$. Let S' consist of all vertices of $G(n^2 + 2n, p) := G(1, 2, \ldots, n^2, \ldots, n^2 + 2n, p)$ that are not dominated by S; clearly we have

$$|S| + |S'| \geq D_{n^2+j} \ \forall\ 1 \leq j \leq 2n,$$

and, in particular, the set $S \cup S'$ dominates $G(n^2 + 2n, p)$. But

$$|S'| = \sum_{j=1}^{2n} F_j,$$

where the F_j are independent Bernoulli variables with parameter $(1-p)^{R_{n^2}+2}$, so that the well-known estimate

$$\mathbb{P}(\mathrm{Bin}(n,p) \geq k) \leq \frac{(np)^k}{k!}$$

yields

$$\begin{aligned}
\mathbb{P}(|S'| \geq 2) &\leq 2n^2(1-p)^{2R_{n^2}+4} \\
&\leq 2n^2(1-p)^2\,(1-p)^{2(\mathbb{L}n^2 - \mathbb{L}((\mathbb{L}n^2)(\log n^2)))} \\
&= 32(1-p)^2 \frac{(\mathbb{L}n)^2(\log n)^2}{n^2}.
\end{aligned} \tag{19}$$

We could have, in (19), used a more exact computation, but the end result would have been the same (up to a constant). In any case, (19) and the Borel-Cantelli lemma reveal that

$$\mathbb{P}(|S'| \geq 2 \text{ infinitely often}) = 0,$$

so that we have, on using equation (18) and the notation "i.o." for "infinitely often,"

$$\begin{aligned}\mathbb{P}(D_n \geq R_n + 4 \text{ i.o.}) &= \mathbb{P}(D_{n^2} \geq R_{n^2} + 3 \text{ i.o.}, D_n \geq R_n + 4 \text{ i.o.}) \\ &\quad + \mathbb{P}(D_{n^2} \leq R_{n^2} + 2\ (n \geq n_0), D_n \geq R_n + 4 \text{ i.o.}) \\ &\leq 0 + \mathbb{P}(|S'| \geq 2 \text{ i.o.}) \\ &= 0. \end{aligned} \tag{20}$$

The result follows on combining (17) and (20).

4 Open Questions

(1) Noga Alon and David Wilson both commented, after listening to Godbole's talk at the 2001 Poznań Random Structures and Algorithms conference, that it is likely that the two-point concentration result can be extended to a wider range of ps. The delicate analysis needed to show this remains to be conducted.
(2) Can the results in this paper, which have obvious connections to the so-called "tournaments with property S_k" [1], be used to improve the bounds in Section 1.2 of [1]?

Acknowledgment The research of both authors was supported by NSF Grant DMS-9619889, and was conducted at Michigan Technological University in the Summer of 1999, when Ben Wieland was an undergraduate student at the Massachusetts Institute of Technology.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, John Wiley, New York, 1992.

[2] B. Bollobás, *Random Graphs*, Academic Press, New York, 1985.

[3] P. Dryer, Ph.D. Dissertation, Department of Mathematics, Rutgers University, 2000.

[4] C. Kaiser and K. Weber (1985), "Degrees and domination number of random graphs in the n-cube," *Rostock. Math. Kolloq.* **28**, 18–32.

[5] S. Nikoletseas and P. Spirakis, "Near optimal dominating sets in dense random graphs with polynomial expected time," in J. van Leeuwen, ed., *Graph Theoretic Concepts in Computer Science*, *Lecture Notes in Computer Science* **790**, pp. 1–10, Springer Verlag, Berlin, 1994.

[6] T. Haynes, S. Hedetniemi, and P. Slater, *Fundamentals of Domination in Graphs*, Marcel Dekker, Inc., New York, 1998.

[7] T. Haynes, S. Hedetniemi, and P. Slater, eds., *Domination in Graphs: Advanced Topics*, Marcel Dekker, Inc., New York, 1998.

[8] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.