Atlantic Electronic
Journal of Mathematics
Volume **3**, Number **1**, Winter **2008**

http://aejm.ca
http://aejm.ca/rema
pp. **1–5**

A NOTE ON THRESHOLDS AND CONNECTIVITY IN RANDOM DIRECTED GRAPHS

Alasdair J. Graham and David A. Pike

Department of Mathematics and Statistics
Memorial University of Newfoundland
St. John's, Newfoundland
Canada. A1C 5S7

Abstract. Thresholds and phase transitions have been well studied for several properties of random undirected graphs. In this article we adapt a result of Palásti to the context of directed graphs, thereby allowing thresholds for the uniform random directed graph model $\mathbb{D}(n, M)$ to be used in determining thresholds for the binomial random directed graph model $\mathbb{D}(n, p)$. We then determine the threshold for strong connectivity in $\mathbb{D}(n, p)$.

1. **Introduction.** The study of random graphs harkens back to the 1940s when Paul Erdős introduced a model whereby for each of the $\binom{n}{2}$ possible edges in an undirected graph on n labelled vertices one tosses a fair coin to determine whether the edge is present or absent [4]. In this initial model, the likelihood of selecting a particular graph G is $2^{-\binom{n}{2}}$.

In a subsequent model Erdős and Rényi included a probability value p, which need not be $\frac{1}{2}$, to be used for each edge [5]. As one would intuitively expect, when the value of p is small (that is, near zero) there are too few edges for a randomly constructed graph to be likely to possess certain properties, yet when p is large (that is, near one) there are so many edges that the property is nearly certain to be present. Somewhere in between these two extremes, such properties have often been found to cross a threshold from being almost certainly absent to almost certainly present.

The seminal work on the topic of such phase transitions is also due to Erdős and Rényi [6]. The topic has since been widely studied and developed; an excellent overview can be found in Chapter 5 of [7], which carefully presents several aspects of the phase transition in the size of the largest connected component in random undirected graphs.

In the present paper, we focus our attention on the setting of random directed graphs, for which we consider the nature of the threshold for the property of being strongly connected. As our main result, we determine the threshold for strong connectivity in the binomial random directed graph model $\mathbb{D}(n, p)$.

2000 *Mathematics Subject Classification.* Primary: 05C80, 05C20; Secondary: 05C40.
Key words and phrases. random graphs, directed graphs, connectivity.

2. **Definitions and Notation.** For a detailed review of random graphs the reader is referred to [7], some of which we now draw upon to introduce four different random graph models. The first model, the one that readers are most likely to be familiar with, is the Binomial Random Graph Model, $\mathbb{G}(n,p)$. In this model, a graph on the vertex set $\{1,2,\ldots,n\}$ is constructed by relying on $\binom{n}{2}$ independent random events, each with probability p, to determine one at a time which of the $\binom{n}{2}$ possible edges are to be included in the graph. If $p=\frac{1}{2}$, then this becomes the model introduced by Erdős in 1947 [4].

A second model for undirected graphs is the Uniform Random Graph Model, denoted by $\mathbb{G}(n,M)$ in which a graph on n labelled vertices and M edges is selected uniformly at random from among the $\binom{\binom{n}{2}}{M}$ graphs on n vertices and M edges.

These two models for undirected graphs have natural counterparts for directed graphs. In the Binomial Random Directed Graph Model, $\mathbb{D}(n,p)$, each of the n^2 possible arcs (including possible loops) in a directed graph on n labelled vertices is selected on the basis of a random event with probability p. In the Uniform Random Directed Graph Model, $\mathbb{D}(n,M)$, a directed graph on n vertices and M arcs is selected uniformly at random from the $\binom{n^2}{M}$ directed graphs on n vertices and M arcs.

For a property $\mathcal{P}$ we will write $G\in\mathcal{P}$ to indicate that G is a member of the family of graphs that possess property $\mathcal{P}$. Both $\mathcal{P}$ and its corresponding family of graphs are said to be *increasing* if $G\in\mathcal{P}$ whenever there exists a spanning subgraph H of G such that $H\in\mathcal{P}$. So for example, containing a 3-cycle is an increasing property, whereas being bipartite is not.

For an increasing property $\mathcal{P}$, a sequence $\hat{p}=\hat{p}(n)$ is called a *threshold* for the binomial random graph model if

$$\lim_{n\to\infty}\mathbb{P}\big(\mathbb{G}(n,p)\in\mathcal{P}\big)=\begin{cases}0 & \text{if } p\ll\hat{p},\\ 1 & \text{if } p\gg\hat{p},\end{cases}$$

where $p\ll\hat{p}$ if and only if $\lim_{n\to\infty}\dfrac{p}{\hat{p}}=0$. It was shown in 1987 by Bollobás and Thomason that every increasing property has a threshold [3]. As an example, the threshold for $\mathbb{G}(n,p)$ to have a 3-cycle is $\hat{p}=\frac{1}{n}$; this follows from a result of Bollobás [1].

In the setting of the uniform random graph model, thresholds also exist and are based on the number of edges (rather than on a probability parameter as in the binomial model). To continue with the example of the property of containing a 3-cycle, the threshold for this property in $\mathbb{G}(n,M)$ is $\hat{M}=n$.

3. **Main Results.** Consider now the property $\mathcal{C}$ of an undirected graph being connected. The following theorem of Erdős and Rényi, stated in the form that is presented in [2, pages 150–151], is useful to us.

Theorem 1 (Erdős and Rényi [5])**.** *Let $c\in\mathbb{R}$ be fixed and let $M=\frac{n}{2}\big\lfloor\ln n+c+o(1)\big\rfloor\in\mathbb{N}$ and $p=\frac{1}{n}\big(\ln n+c+o(1)\big)$. Then, as $n\to\infty$, $\mathbb{P}\big(\mathbb{G}(n,M)\in\mathcal{C}\big)\to e^{-e^{-c}}$ and $\mathbb{P}\big(\mathbb{G}(n,p)\in\mathcal{C}\big)\to e^{-e^{-c}}$.*

Theorem 1 establishes the thresholds for $\mathcal{C}$ in both the uniform and binomial random graph models to be $\hat{M}=\frac{n}{2}\big\lfloor\ln n+c+o(1)\big\rfloor$ and $\hat{p}=\frac{1}{n}\big(\ln n+c+o(1)\big)$, respectively. That these are in fact thresholds follows from the fact that thresholds are known to exist (by Bollobás and Thomason [3]) and so if $\hat{M}$ and $\hat{p}$ were not thresholds then $\mathbb{P}\big(\mathbb{G}(n,M)\in\mathcal{C}\big)$ and $\mathbb{P}\big(\mathbb{G}(n,p)\in\mathcal{C}\big)$ would be elements of the set $\{0,1\}$. But $e^{-e^{-c}}\notin\{0,1\}$.

We now turn our interest to directed graphs. A directed graph D is said to be *strongly connected* if, for each pair of vertices u and v of D, there exists a directed path from u to v (and hence there also exists a directed path from v to u). If we let $\mathcal{S}$ denote the property of a directed graph being strongly connected, then the following result of Palásti establishes a threshold for $\mathcal{S}$ in the context of the uniform random directed graph model.

Theorem 2 (Palásti [9]). *Let $c\in\mathbb{R}$ be a fixed real number. If $M=\lfloor n\ln n+cn\rfloor$, then $\lim_{n\to\infty}\mathbb{P}\big(\mathbb{D}(n,M)\in\mathcal{S}\big)=e^{-2e^{-c}}$.*

Our goal is to now determine the threshold for $\mathcal{S}$ in $\mathbb{D}(n,p)$. In order to do so, we shall first prove an analogue of the following theorem of Łuczak which converts thresholds for the uniform random graph model to thresholds for the binomial random graph model.

Theorem 3 (Łuczak [8]). *Let $\mathcal{P}$ be an arbitrary property of subsets of the family of all graphs on n vertices, $p=p(n)\in[0,1]$, $0\leqslant a\leqslant 1$, and $N=\binom{n}{2}$. If for every sequence $M=M(n)$ such that $M=Np+O\big(\sqrt{Np(1-p)}\big)$ it holds that $\mathbb{P}\big(\mathbb{G}(n,M)\in\mathcal{P}\big)\to a$ as $n\to\infty$, then also $\mathbb{P}\big(\mathbb{G}(n,p)\in\mathcal{P}\big)\to a$ as $n\to\infty$.*

By adapting the proof of Theorem 3 to directed graphs, we establish the following theorem.

Theorem 4. *Let $\mathcal{P}$ be an arbitrary property of subsets of the family of all directed graphs on n vertices, $p=p(n)\in[0,1]$, $0\leqslant a\leqslant 1$, and $N=n^2$. If for every sequence $M=M(n)$ such that $M=Np+O\big(\sqrt{Np(1-p)}\big)$ it holds that $\mathbb{P}\big(\mathbb{D}(n,M)\in\mathcal{P}\big)\to a$ as $n\to\infty$, then also $\mathbb{P}\big(\mathbb{D}(n,p)\in\mathcal{P}\big)\to a$ as $n\to\infty$.*

Proof. Let $C\in\mathbb{R}$, $C\geqslant 0$, and for each n define

$$\mathcal{M}(C)=\left\{M:|M-Np|\leqslant C\sqrt{Np(1-p)}\right\}.$$

Let $M_{\inf}\in\mathcal{M}(C)$ be such that $\mathbb{P}\big(\mathbb{D}(n,M_{\inf})\in\mathcal{P}\big)\leqslant\mathbb{P}\big(\mathbb{D}(n,M)\in\mathcal{P}\big)$ for each $M\in\mathcal{M}(C)$. Now, if $E_p=\big|E\big(\mathbb{D}(n,p)\big)\big|$, then by the law of total probability, $\mathbb{P}\big(\mathbb{D}(n,p)\in\mathcal{P}\big)=\sum_{M=0}^{N}\mathbb{P}\big(\mathbb{D}(n,p)\in\mathcal{P}\ :\ E_p=M\big)\cdot\mathbb{P}\big(E_p=M\big)$.

Now consider a fixed directed graph D with n vertices and M arcs. The probability of selecting D when using the binomial model $\mathbb{D}(n,p)$ is $p^M(1-p)^{n^2-M}$, which we note is dependent only on n and M. Hence each directed graph on n vertices and M arcs is equally likely to be selected, implying that $\mathbb{P}\big(\mathbb{D}(n,p)\in\mathcal{P}\ :\ E_p=M\big)=\mathbb{P}\big(\mathbb{D}(n,M)\in\mathcal{P}\big)$ for any property $\mathcal{P}$. It now follows that

$$\begin{aligned}
\mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{P}\big) &= \sum_{M=0}^{N} \mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{P} \;:\; E_p = M\big) \cdot \mathbb{P}\big(E_p = M\big) \\
&= \sum_{M=0}^{N} \mathbb{P}\big(\mathbb{D}(n,M) \in \mathcal{P}\big) \cdot \mathbb{P}\big(E_p = M\big) \\
&\geqslant \sum_{M \in \mathcal{M}(C)} \mathbb{P}\big(\mathbb{D}(n,M_{\text{inf}}) \in \mathcal{P}\big) \cdot \mathbb{P}\big(E_p = M\big) \\
&= \mathbb{P}\big(\mathbb{D}(n,M_{\text{inf}}) \in \mathcal{P}\big) \cdot \mathbb{P}\big(E_p \in \mathcal{M}(C)\big).
\end{aligned}$$

Since $\mathbb{D}(n,p)$ obeys a binomial distribution, we have $\mathbb{E}(E_p) = Np$ and $\text{Var}(E_p) = Np(1-p)$. Therefore, using Chebyshev's Inequality and the assumption that $\mathbb{P}\big(\mathbb{D}(n,M_{\text{inf}}) \in \mathcal{P}\big) \to a$, we have that for $t = C\sqrt{Np(1-p)}$,

$$\mathbb{P}\big(|E_p - \mathbb{E}(E_p)| \geqslant t\big) \quad \leqslant \quad \frac{\text{Var}(E_p)}{t^2}$$

and thus

$$\mathbb{P}\big(E_p \notin \mathcal{M}(C)\big) \quad \leqslant \quad \frac{\text{Var}(E_p)}{\big(C\sqrt{Np(1-p)}\big)^2} \;=\; \frac{1}{C^2}.$$

Hence,

$$\liminf_{n\to\infty} \mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{P}\big) \geqslant a \liminf_{n\to\infty} \mathbb{P}\big(E_p \in \mathcal{M}(C)\big) \geqslant a\left(1 - \frac{1}{C^2}\right).$$

Similarly, if M_{sup} maximizes $\mathbb{P}\big(\mathbb{D}(n,M) \in \mathcal{P}\big)$ for $M \in \mathcal{M}(C)$, we have

$$\begin{aligned}
\mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{P}\big) \quad &\leqslant \quad \mathbb{P}\big(\mathbb{D}(n,M_{\text{sup}}) \in \mathcal{P}\big) \cdot \mathbb{P}\big(E_p \in \mathcal{M}(C)\big) \\
&\qquad + \mathbb{P}\big(\mathbb{D}(n,M_{\text{sup}}) \in \mathcal{P}\big) \cdot \mathbb{P}\big(E_p \notin \mathcal{M}(C)\big)
\end{aligned}$$

which implies that

$$\limsup_{n\to\infty} \mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{P}\big) \leqslant a\left(1 - \frac{1}{C^2}\right) + \frac{a}{C^2}.$$

The result follows when we let $C \to \infty$. □

We now incorporate Palásti's result for strong connectivity in the uniform random directed graph model into the binomial random directed graph model.

Theorem 5. *Let c be an arbitrary, fixed real number. If $p = p(n) = \frac{\ln n + c}{n}$, then* $\lim\limits_{n\to\infty} \mathbb{P}\big(\mathbb{D}(n,p) \in \mathcal{S}\big) = e^{-2e^{-c}}$.

Proof. We apply Theorem 4 to Theorem 2. For this to work, we need $M = Np + O\big(\sqrt{Np(1-p)}\big)$ to hold, which is the case if we let $p = p(n)$ be such that

$$M = \lfloor n \ln n + cn \rfloor = Np = n^2 p.$$

Moreover, we can relax the equation so that $n \ln n + cn = n^2 p$, or equivalently

$$p = \frac{\ln n + c}{n},$$

thereby completing the proof. □

Determining the threshold for $\mathcal{S}$ in $\mathbb{D}(n,p)$ now follows as an easy corollary.

Corollary 1. *For any fixed $c \in \mathbb{R}$, if $p = p(n)$ is a function of n, then $\hat{p} = \frac{\ln n + c}{n}$ is a threshold for $\mathcal{S}$.*

Proof. Either $\hat{p} = \frac{\ln n + c}{n}$ is a threshold for $\mathcal{S}$ or it is not. If it is not, then either $\mathbb{P}\big(\mathbb{D}(n, \hat{p}) \in \mathcal{S}\big) \to 1$ asymptotically almost surely as $n \to \infty$, or $\mathbb{P}\big(\mathbb{D}(n, \hat{p}) \in \mathcal{S}\big) \to 0$ asymptotically almost surely as $n \to \infty$.

However, from Theorem 5, we know that $\mathbb{P}\big(\mathbb{D}(n, \hat{p}) \in \mathcal{S}\big) \to e^{-2e^{-c}}$ as $n \to \infty$, and for each $c \in \mathbb{R}$, it is true that $0 < e^{-2e^{-c}} < 1$. Since $e^{-2e^{-c}}$ is neither 0 nor 1, $\hat{p} = \frac{\ln n + c}{n}$ must be a threshold for $\mathcal{S}$. $\square$

4. **Acknowledgements.** The authors gratefully acknowledge research support from MITACS and NSERC. The authors also wish to thank Tomasz Łuczak for helpful comments.

REFERENCES

[1] B. Bollobás. Random Graphs. In *Combinatorics*, Proceedings, Swansea 1981, London Math. Soc. Lecture Note Ser. 52, Cambridge Univ. Press, Cambridge, 80–102.

[2] B. Bollobás. *Random graphs*, Academic Press, London, England (1985).

[3] B. Bollobás and A. Thomason. Threshold functions, *Combinatorica* **7** (1987) 35–38.

[4] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* **53** (1947) 292–294.

[5] P. Erdős and A. Rényi. On random graphs I, *Publ. Math. Debrecen* **6** (1959) 290–297.

[6] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960) 17–61.

[7] S. Janson, T. Łuczak and A. Ruciński. *Random graphs*, John Wiley & Sons, New York, New York (2000).

[8] T. Łuczak. On the equivalence of two basic models of random graphs, *Proceedings of Random Graphs '87*, eds. M. Karoński, J. Jaworski & A. Ruciński. John Wiley & Sons, Chichester (1990) 151–158.

[9] I. Palásti. On the strong connectedness of directed random graphs, *Studia Scientiarum Mathematicarum Hungarica* **1** (1966) 205–214.

E-mail address: alasdair_john_graham@yahoo.ca; dapike@math.mun.ca