THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2008-30

Connectivity in One-Dimensional Geometric Random Graphs: Poisson Approximations, Zero-One Laws and Phase Transitions

Guang Han and Armand M. Makowski

Connectivity in one-dimensional geometric random graphs: Poisson approximations, zero-one laws and phase transitions

Guang Han and Armand M. Makowski, *Fellow, IEEE*

***Abstract*— Consider n points (or nodes) distributed uniformly and independently on the unit interval $[0,1]$. Two nodes are said to be adjacent if their distance is less than some given threshold value. For the underlying random graph we derive zero-one laws for the property of graph connectivity and give the asymptotics of the transition widths for the associated phase transition. These results all flow from a single convergence statement for the probability of graph connectivity under a particular class of scalings. Given the importance of this result, we give two separate proofs; one approach relies on results concerning maximal spacings, while the other one exploits a Poisson convergence result for the number of breakpoint users.**

Keywords: Geometric random graphs, Connectivity, Critical scalings, Zero-one laws, Phase transitions.

I. Introduction

Geometric random graphs appear in settings as diverse as statistical physics, pattern recognition and biomedical research to name a few; see the monograph by Penrose [34] for a comprehensive discussion. In the context of wireless networks, Gupta and Kumar [18] have recently drawn attention to the two-dimensional version through their use of the popular disk model.

In this paper we consider the simplest of geometric random graph models, namely the ones defined over finite (one-dimensional) intervals – Consider n points (or communication nodes) which are distributed uniformly and independently on the (generic) interval $[0,1]$. Two nodes are said to be adjacent (or to communicate with each other) if their distance is less than some given transmission range $\tau > 0$. This one-dimensional model has been proposed for wireless networks constrained over "linear highways," e.g., see [7], [11], [12], [13], [16], [32].

We focus on the connectivity of the induced graph when n becomes large and the transmission range is appropriately scaled with n, i.e., the transmission range is made to depend on n through scalings $\tau : \mathbb{N}_0 \to \mathbb{R} : n \to \tau_n$. It is well known that the graph is connected (resp. not connected) with a very high probability (as n becomes large) depending on how the scaling used deviates from a *critical* scaling $\tau^\star : \mathbb{N}_0 \to \mathbb{R}$. This critical scaling, given by (9), has the following rough operational meaning: Let $P(n;\tau)$ denote the probability that the n node network is connected under transmission range τ. For n sufficiently large, a transmission range τ_n suitably larger (resp. smaller) than $\tau^\star_n$ ensures $P(n;\tau_n) \simeq 1$ (resp. $P(n;\tau_n) \simeq 0$). Such statements are known in the literature as *zero-one* laws once the precise technical meaning for suitably larger (resp. smaller) has been elucidated, see e.g., [1], [21], [32]. Phase transitions are associated with such zero-one laws [26] and sharp asymptotics are sometimes available for the corresponding transition width [19], [20].

Han Guang was with the Department of Electrical and Computer Engineering, and the Institute for Systems Research, University of Maryland, College Park, MD 20742, U.S.A. He is now with the Advanced Network Technology Group, Motorola Inc., 1441 W. Shure Drive, Arlington Heights, IL 60004, U.S.A. E-mail: guang.han@motorola.com

Armand M. Makowski is with the Department of Electrical and Computer Engineering, and the Institute for Systems Research, University of Maryland, College Park, MD 20742, U.S.A. E-mail: armand@isr.umd.edu.

The material in this paper was presented in part at the Inaugural Workshop, Information Theory and Applications, University of California, San Diego (CA), February 2006 and at the IEEE International Conference on Communications (ICC 2006), Istanbul (Turkey), June 2006.

This work was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.

The one-dimensional models are arguably the least geometric in nature, and as such occupy a somewhat singular place in the literature on geometric random graphs [34, p. 283]. This is reflected by the continuing attention they have received from research communities with various (non-geometric) perspectives: The monograph by Godehardt [14] deals with applications to cluster analysis, and the exhaustive study in [15] provides a direct combinatorial analysis of many results of interest. Appel and Russo [1, p. 352] leverage the connection with maximal spacings, while Muthukrishnan and Pandurangan [32, Thm. 2.2] make use of bin-covering techniques.

As a result of these efforts, many questions concerning graph connectivity have by now been answered, albeit in various forms of completeness. However, some of the results have been reproduced independently, are scattered in multiple literatures and are not always couched in graph-theoretic terms. Here we provide a unified presentation of these results, both old and new, in their sharpest form; the discussion emphasizes the single convergence statement

$$\lim_{n\to\infty} P\left(n; \frac{\log n + x}{n}\right) = e^{-e^{-x}}, \quad x \in \mathbb{R}. \quad (1)$$

as the source for all relevant results. This convergence points to the special role played by the scalings $n \to \frac{\log n + x}{n}$, and foreshadows the form (9) of the critical scaling.

The convergence (1) paves the way to a number of results: From it we first can derive a zero-one law for graph connectivity [Section III]. The version given here is stronger than the one usually discussed in the literature, and for this reason we refer to it as a very strong zero-one law [Section IV]. This zero-one law was already obtained by the authors by means of a different technique in [21]. The convergence (1) leads easily to precise asymptotics on the width for the phase transition inherent in the very strong zero-one law [Section V]; these asymptotics were already announced in [19], [20].

A closed-form expression for $P(n;\tau)$ is available [Section VI]. However, there does not appear a simple way to use it in order to establish (1). Given the central place occupied by this convergence, we shall present two very different approaches to its proof – Each proof makes use of a different characterization of graph connectivity in terms of the spacings induced by i.i.d. samples drawn from the uniform distribution on finite intervals [3, Chap. 7]. The contributions of the paper can be summarized as follows:

Lévy's result on maximal spacings – The first approach relies on the fact that the maximal spacing induced by the node positions provides an immediate characterization of graph connectivity. We then obtain (1) as a byproduct of classical results concerning the asymptotic theory of maximal spacings. The result goes back the work of Lévy [28], but we provide here a simple proof which appears to be new [Section VII].

Stein-Chen approximation – Next we characterize graph connectivity through the number of breakpoint nodes which counts the number of connected components minus one. A key result is a Poisson approximation for this count variable through the Stein-Chen method [Section VIII with a proof in Section X]. This approach provides convergence to a Poisson random variable under the scaling (4), as well as an *approximation* (in the total variation norm) which can be used to glean information on the corresponding rate of convergence. As a result we are now in a position to understand the performance of *finite* node graphs [Section IX] through an explicitly computable bound on the difference

$$\left| P\left(n; \frac{\log n + x}{n}\right) - e^{-e^{-x}} \right|, \quad x \in \mathbb{R}$$

for all $n = 2, 3, \ldots$. As an added bonus, this bound can be leveraged to obtain a rate of convergence for Lévy's original result. This time, the convergence (1) is an easy consequence of the aforementioned Poisson convergence.

We close the paper [Section XI] with several pointers: The two approaches given here have analogs in geometric random graphs of dimension two and higher. When nodes are placed according to an arbitrary distribution, results have been developed by the authors in a series of recent papers [22], [23], [24].

A word on the notation and conventions in use: All limiting statements, including asymptotic equivalences, are understood with n going to infinity. The random variables (rvs) under consideration are all defined on the same probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. Probabilistic statements are made with respect to this probability measure $\mathbb{P}$, and we denote the corresponding expectation operator by $\mathbb{E}$. The notation

$\xrightarrow{P}_n$ (resp. $\Longrightarrow_n$) is used to signify convergence in probability (resp. weak convergence) with n going to infinity. Also, we use the notation $=_{st}$ to indicate distributional equality. The indicator function of an event E is denoted by $\mathbf{1}[E]$.

II. THE MODEL AND A KEY CONVERGENCE RESULT

We start with a sequence $\{X_i, \ i = 1, 2, \ldots\}$ of i.i.d. rvs which are distributed uniformly in the interval $[0, 1]$. For each $n = 2, 3, \ldots$, we think of $X_1, \ldots, X_n$ as the locations of n nodes (or users), labelled $1, \ldots, n$, in the interval $[0, 1]$. Given a fixed distance (or transmission range) $\tau > 0$, nodes i and j are said to be adjacent if $|X_i - X_j| \leq \tau$, in which case an undirected edge exists between them. This notion of adjacency amongst nodes gives rise to an undirected geometric random graph, thereafter denoted $\mathbb{G}(n; \tau)$. As usual, the graph $\mathbb{G}(n; \tau)$ is said to be connected if every pair of nodes can be linked by at least one path over the edges of the graph, and we write

$$P(n; \tau) := \mathbb{P}\left[\mathbb{G}(n; \tau) \text{ is connected}\right]. \tag{2}$$

We refer to the quantity $P(n; \tau)$ as the probability of graph connectivity. Obviously, $P(n; \tau) = 1$ if $\tau \geq 1$. We also find it convenient to set $P(n; \tau) = 0$ if $\tau \leq 0$.

We are interested in understanding how the probability of graph connectivity behaves when the number n of nodes becomes large and the transmission range τ is scaled appropriately. Thus, with range function or *scaling* $\tau : \mathbb{N}_0 \to \mathbb{R} : n \to \tau_n$, we investigate the limit

$$\lim_{n\to\infty} P(n; \tau_n) \tag{3}$$

whenever it exists. We allow scalings to take on negative values as a matter of convenience in order to simplify the presentation in a number of places (with the help of the convention following (2)).

The basic message of the paper is that the needed asymptotics pertaining to (3) all flow from a single statement which we now present: With each x in $\mathbb{R}$, we associate the range function $\sigma(x) : \mathbb{N}_0 \to \mathbb{R}$ defined by

$$\sigma_n(x) := \frac{\log n + x}{n}, \quad n = 1, 2, \ldots \tag{4}$$

Note that $\sigma_n(x) > 0$ for all n sufficiently large.

Theorem 2.1: For each x in $\mathbb{R}$, we have

$$\lim_{n\to\infty} P(n; \sigma_n(x)) = e^{-e^{-x}}. \tag{5}$$

Theorem 2.1 has several byproducts which are discussed in the next three sections. In the course of the paper two very different approaches will be presented to establish (5).

III. A ZERO-ONE LAW AND ITS CRITICAL SCALING

We start by noting that there is no loss of generality in writing any range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ in the form

$$\tau_n = \frac{1}{n}\left(\log n + \alpha_n\right), \quad n = 1, 2, \ldots. \tag{6}$$

for some deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$ – Just take

$$\alpha_n = n\tau_n - \log n, \quad n = 1, 2, \ldots.$$

Theorem 3.1: For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ written in the form (6) for some deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, we have

$$\lim_{n\to\infty} P(n; \tau_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \alpha_n = -\infty \\ 1 & \text{if } \lim_{n\to\infty} \alpha_n = +\infty. \end{cases} \tag{7}$$

In [21], the authors gave a direct derivation of Theorem 3.1 by the method of first and second moments, an approach widely used in the theory of Erdős-Renyi graphs [25, p. 55]. Here we take a different approach: The form of (6)–(7) suggests interpreting (5) (via (4)) as an *interpolation* result between $x = -\infty$ and $x = \infty$. This is indeed borne out by the proof of Theorem 3.1 given next, which exploits only the validity of (5).

Proof. Consider a range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ written in the form (6) with $\lim_{n\to\infty} \alpha_n = \infty$. Thus, for every x in $\mathbb{R}$, there exists an integer $n(x)$ such that $x < \alpha_n$ for all $n > n(x)$, in which case

$$P(n; \sigma_n(x)) \leq P(n; \tau_n), \quad n \geq n(x)$$

since the mapping $\tau \to P(n; \tau)$ is monotonically increasing for each $n = 2, 3, \ldots$. Letting n go to infinity in this last inequality, we conclude from (5) that

$$e^{-e^{-x}} \leq \liminf_{n\to\infty} P(n; \tau_n).$$

We get the one-law upon noting that x can be made arbitrarily large in this last inequality. This implies $\liminf_{n\to\infty} P(n;\tau_n) = 1$, whence $\lim_{n\to\infty} P(n;\tau_n) = 1$. The case $\lim_{n\to\infty} \alpha_n = -\infty$ can be handled *mutatis mutandis* with details left to the interested reader. ■

We note that Theorem 3.1 can be consolidated with (5) through a single statement.

Theorem 3.2: For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ *written in the form (6), we have*

$$\lim_{n\to\infty} P(n;\tau_n) = e^{-e^{-\alpha}} \tag{8}$$

whenever $\alpha = \lim_{n\to\infty} \alpha_n$ *(possibly* $\pm\infty$*).*

Theorem 3.2 is in fact equivalent to (5), and can be established from it by arguments similar to the ones used in the proof of Theorem 3.1; details are left to the interested reader.

From Theorem 3.1 we see that the range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ given by

$$\tau^\star_n = \frac{\log n}{n}, \quad n = 1,2,\ldots \tag{9}$$

acts as a *critical* scaling for graph connectivity in that it defines a *boundary* in the space of range functions. Roughly speaking, for n large, a transmission range τ_n suitably larger (resp. smaller) than $\tau^\star_n$ ensures that the graph $\mathbb{G}(n;\tau_n)$ is connected (resp. disconnected) with very high probability. The precise technical meaning for suitably larger (resp. smaller) is found at (7): For any range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ written in the form (6) with some deviation function $\alpha : \mathbb{N}_0 \to \mathbb{R}$, we can write

$$\tau_n = \tau^\star_n + \frac{\alpha_n}{n}, \quad n = 1,2,\ldots. \tag{10}$$

By Theorem 3.1 the perturbations $\frac{\alpha_n}{n}$ from the critical scaling yield the one-law in the form $P(n;\tau_n) \simeq 1$ (resp. the zero-law in the form $P(n;\tau_n) \simeq 0$) provided $\lim_{n\to\infty} \alpha_n = \infty$ (resp. $\lim_{n\to\infty} \alpha_n = -\infty$) with *no* further constraint on the deviation function α. This issue will be revisited in Section IV.

IV. How strong is the critical scaling $\tau^\star$?

A number of definitions have been given in the literature regarding critical scalings [31, p. 376]. In this section we explore their relevance for the scaling $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ in the context of the random graph $\mathbb{G}(n;\tau)$.

With range function $\tau : \mathbb{N}_0 \to \mathbb{R}$, it is a simple matter to check from (37) (with the help of (36)) that

$$\lim_{n\to\infty} P(n;\tau_n) = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \frac{\tau_n}{\tau^\star_n} = 0 \\ \\ 1 & \text{if } \lim_{n\to\infty} \frac{\tau_n}{\tau^\star_n} = \infty. \end{cases} \tag{11}$$

It is customary to summarize (11) by saying that graph connectivity in $\mathbb{G}(n;\tau)$ admits a *weak* zero-one law, and the range function $\tau^\star$ is the corresponding (weak) critical scaling [31, p. 376]. This terminology reflects the fact that the one law (resp. zero law) emerges with range functions $\tau : \mathbb{N}_0 \to \mathbb{R}$ which are *at least* an order of magnitude larger (resp. smaller) than $\tau^\star$.

However, a much stronger conclusion than (11) has been obtained by several authors, namely

$$\lim_{n\to\infty} P(n;c\tau^\star_n) = \begin{cases} 0 & \text{if } 0 < c < 1 \\ \\ 1 & \text{if } 1 < c. \end{cases} \tag{12}$$

It is easy to see that this last result still holds for any range function $\tau : \mathbb{N}_0 \to \mathbb{R}$ such that $\tau_n \sim c\tau^\star_n$ for some $c > 0$ with

$$\lim_{n\to\infty} P(n;\tau_n) = \begin{cases} 0 & \text{if } 0 < c < 1 \\ \\ 1 & \text{if } 1 < c. \end{cases} \tag{13}$$

The two zero-one laws (12) and (13) are equivalent, and are already contained in Theorem 1 by Appel and Russo [1, p. 352]. Muthukrishnan and Pandurangan [32, Thm. 2.2] have also derived (12) by a bin-covering technique.

We characterize (12)–(13) by saying that graph connectivity in $\mathbb{G}(n;\tau)$ admits a *strong* zero-one law, and we refer to the range function $\tau^\star$ as a *strong* critical scaling [31]. Indeed, for n sufficiently large, a transmission range τ_n suitably larger (resp. smaller) than $\tau^\star_n$ ensures $P(n;\tau_n) \simeq 1$ (resp. $P(n;\tau_n) \simeq 0$) provided $\tau_n \sim c\tau^\star_n$ with $c > 1$ (resp. $0 < c < 1$). This is in sharp contrast with (11) in that the one law (resp. zero law) still emerges with range functions $\tau : \mathbb{N}_0 \to \mathbb{R}$ which are larger (resp. smaller) than $\tau^\star$ but of the *same* order of magnitude as $\tau^\star$!

It is easy to see that either (12) or (13) implies (11), as would be expected. Moreover, it is also straightforward to see that (12) is a mere byproduct of Theorem 3.1: Indeed, for each $c > 0$ we have

$$c\tau^\star_n = \frac{1}{n}\left(\log n + \alpha_n\right) \text{ with } \alpha_n = (c-1)\frac{\log n}{n}$$

for all $n = 1, 2, \ldots$ so that $\lim_{n\to\infty} \alpha_n = \infty$ (resp. $\lim_{n\to\infty} \alpha_n = -\infty$) if $c > 1$ (resp. $0 < c < 1$).

Theorem 3.1 represents a considerable strengthening of (12)–(13) in that the zero-one law still holds if we allow perturbations from the critical scaling $\tau_n^\star$ which are much smaller than $(c-1)\tau_n^\star$ – For instance, the "small" deviations of the form $\alpha_n = \pm \log\log n$, which were not covered under the zero-one law (12) (since $\alpha_n = o(\log n)$), are now covered by Theorem 3.1. In some sense, Theorem 3.1 corresponds to the case $c = 1$ in (12)–(13). Therefore it seems appropriate to call the zero-one law (7) a *very strong* zero-one law, and the critical scaling $\tau^\star$ a *very strong* (and not merely a strong) critical scaling for the property of graph connectivity. This captures the extreme sensitivity to perturbations from the critical scaling $\tau^\star$, as is already apparent from graphs available in several papers, e.g., see [13], [16]. This sharp phase transition is discussed formally in Section V.

V. Transition widths

For each $n = 2, 3, \ldots$, the mapping $\tau \to P(n;\tau)$ is continuous and strictly monotone increasing on $(0,1)$. Given a in $(0,1)$, this guarantees the existence and uniqueness of solutions to the equation

$$P(n;\tau) = a, \quad \tau \in (0,1) \tag{14}$$

and let $\tau_n(a)$ denote this unique solution, The main result concerning the behavior of $\tau_n(a)$ for large n is also a consequence of (5).

Theorem 5.1: For every a in the interval $(0,1)$, we have

$$\begin{aligned} \tau_n(a) &= \frac{\log n}{n} - \frac{1}{n}\log\left(\log\left(\frac{1}{a}\right)\right) + o\left(n^{-1}\right) \\ &= \tau_n^\star - \frac{1}{n}\log\left(\log\left(\frac{1}{a}\right)\right) + o\left(n^{-1}\right). \end{aligned} \tag{15}$$

Next, with $0 < a < \frac{1}{2}$, we define

$$\delta_n(a) := \tau_n(1-a) - \tau_n(a).$$

The transition width $\delta_n(a)$ measures the increase in transmission range needed in the n node network to drive the probability of connectivity from level a to level $1-a$. The more rapidly $\tau_n(a)$ decays as a function of n, the sharper the phase transition. The following result is an easy corollary to Theorem 5.1.

Corollary 5.2: For every a in the interval $(0, \frac{1}{2})$, we have

$$\delta_n(a) = \frac{C(a)}{n} + o\left(n^{-1}\right) \tag{16}$$

with constant $C(a)$ given by

$$C(a) := \log\left(\frac{\log a}{\log(1-a)}\right). \tag{17}$$

It is a simple matter to check that $a \to C(a)$ is decreasing on the interval $(0, \frac{1}{2})$ with $\lim_{a\downarrow 0} C(a) = \infty$ and $\lim_{a\uparrow \frac{1}{2}} C(a) = 0$. These qualitative features are in line with one's intuition.

Recently Goel et al. [17, Thm. 1.1] have shown that

$$\delta_n(a) = O\left(\sqrt{\frac{-\log a}{n}}\right). \tag{18}$$

In fact these asymptotic bounds were established for every monotone graph property. Corollary 5.2 markedly improves on (18) in that *exact* asymptotics are now provided and the rate of decay n^{-1} is much faster than the rough asymptotic bound given by (18). However, these conclusions hold only for graph connectivity.

We now turn to the proof of Theorem 5.1 where we shall use the notation

$$g(x) := e^{-e^{-x}}, \quad x \in \mathbb{R}. \tag{19}$$

Proof. We begin by restating (5) as follows: For each x in $\mathbb{R}$ and each $\varepsilon > 0$, there exists a finite integer $n^\star(\varepsilon, x)$ such that

$$g(x) - \varepsilon < P(n;\sigma_n(x)) < g(x) + \varepsilon, \quad n \geq n^\star(\varepsilon, x). \tag{20}$$

Now fix a in the interval $(0,1)$. The mapping $g : \mathbb{R} \to (0,1) : x \to g(x)$ defined at (32) is strictly monotone and continuous with $\lim_{x\to-\infty} g(x) = 0$ and $\lim_{x\to\infty} g(x) = 1$. Therefore, there exists a unique scalar x_a such that $g(x_a) = a$, namely

$$x_a := -\log\left(-\log a\right). \tag{21}$$

Next, pick ε sufficiently small such that $0 < 2\varepsilon < a$ and $a + 2\varepsilon < 1$. Repeatedly applying (20) with $x = x_{a+\varepsilon}$ and $x = x_{a-\varepsilon}$, we get

$$g(x_{a+\varepsilon}) - \varepsilon < P(n;\sigma_n(x_{a+\varepsilon})) < g(x_{a+\varepsilon}) + \varepsilon \tag{22}$$

whenever $n \geq n^\star(\varepsilon, x_{a+\varepsilon})$, and

$$g(x_{a-\varepsilon}) - \varepsilon < P(n;\sigma_n(x_{a-\varepsilon})) < g(x_{a-\varepsilon}) + \varepsilon \tag{23}$$

whenever $n \geq n^\star(\varepsilon, x_{a-\varepsilon})$. In the remainder of this proof, all inequalities are understood to hold for $n \geq n^\star(a;\varepsilon)$ where we have set $n^\star(a;\varepsilon) := \max\left(n^\star(\varepsilon, x_{a+\varepsilon}), n^\star(\varepsilon, x_{a-\varepsilon})\right)$. Under these choices, the inequalities (22) and (23) ensure $\sigma_n(x_{a\pm\varepsilon}) > 0$ for $n \geq n^\star(a;\varepsilon)$.

Since $g(x_{a\pm\varepsilon}) = a \pm \varepsilon$, the two chains of inequalities at (22) and (23) become

$$a < P(n; \sigma_n(x_{a+\varepsilon})) < a + 2\varepsilon$$

and

$$a - 2\varepsilon < P(n; \sigma_n(x_{a-\varepsilon})) < a.$$

But the definitions of $\tau_n(a)$ and $\tau_n(a \pm 2\varepsilon)$ also give

$$P(n; \tau_n(a)) < P(n; \sigma_n(x_{a+\varepsilon})) < P(n; \tau_n(a + 2\varepsilon))$$

and

$$P(n; \tau_n(a - 2\varepsilon)) < P(n; \sigma_n(x_{a-\varepsilon})) < P(n; \tau_n(a)).$$

The strict monotonicity of $\tau \to P(n;\tau)$ then implies

$$\tau_n(a) < \sigma_n(x_{a+\varepsilon}) < \tau_n(a + 2\varepsilon)$$

and

$$\tau_n(a - 2\varepsilon) < \sigma_n(x_{a-\varepsilon}) < \tau_n(a),$$

and we get

$$\sigma_n(x_{a-\varepsilon}) < \tau_n(a) < \sigma_n(x_{a+\varepsilon}). \qquad (24)$$

Next, with

$$\xi_n(a) := \tau_n(a) - \sigma_n(x_a), \quad n = 2, 3, \ldots \qquad (25)$$

we obtain from (24) that

$$\sigma_n(x_{a-\varepsilon}) - \sigma_n(x_a) < \xi_n(a) < \sigma_n(x_{a+\varepsilon}) - \sigma_n(x_a).$$

It is now plain from the definition (4) that

$$x_{a-\varepsilon} - x_a \leq n\xi_n(a) \leq x_{a+\varepsilon} - x_a \qquad (26)$$

since

$$\sigma_n(x_{a\pm\varepsilon}) - \sigma_n(x_a) = n^{-1}\left(x_{a\pm\varepsilon} - x_a\right).$$

Finally, letting n go to infinity in (26) yields $x_{a-\varepsilon} - x_a \leq \liminf_{n\to\infty}(n\xi_n(a))$ and $\limsup_{n\to\infty}(n\xi_n(a)) \leq x_{a+\varepsilon} - x_a$. Given that $\varepsilon > 0$ can be taken arbitrarily small under the required conditions, it follows that $\liminf_{n\to\infty}(n\xi_n(a)) = \limsup_{n\to\infty}(n\xi_n(a)) = 0$ since $\lim_{\varepsilon\downarrow 0}(x_{a\pm\varepsilon} - x_a) = 0$. Thus, $\lim_{n\to\infty}(n\xi_n(a)) = 0$, whence $\xi_n(a) = o(n^{-1})$. Reporting this fact into (25) leads to

$$\tau_n(a) = \sigma_n(x_a) + o(n^{-1}), \quad n = 2, 3, \ldots$$

and the desired result readily follows from (4) and (21). ■

VI. Spacings and closed-form expressions

A natural approach to the derivation of (5) consists in developing expressions for (2). To that end, fix $n = 2, 3, \ldots$ and $\tau > 0$. With the node locations $X_1, \ldots, X_n$, we associate rvs $X_{n,1}, \ldots, X_{n,n}$ which are the locations of these n users arranged in increasing order, i.e., $X_{n,1} \leq \ldots \leq X_{n,n}$ with the convention $X_{n,0} = 0$ and $X_{n,n+1} = 1$. The rvs $X_{n,1}, \ldots, X_{n,n}$ are the *order statistics* associated with the n i.i.d. rvs $X_1, \ldots, X_n$. We also define the *spacing* rvs

$$L_{n,k} := X_{n,k} - X_{n,k-1}, \quad k = 1, \ldots, n+1. \qquad (27)$$

The obvious relation $L_{n,1} + \ldots + L_{n,n+1} = 1$ already suggests that the spacings $L_{n,1}, \ldots, L_{n,n+1}$ should exhibit some form of negative correlation; this will be formalized in Lemma 10.2.

Interest in these spacings derives from the observation that the graph $\mathbb{G}(n;\tau)$ is connected if and only if $L_{n,k} \leq \tau$ for *all* $k = 2, \ldots, n$, so that

$$P(n;\tau) = \mathbb{P}\left[L_{n,k} \leq \tau,\ k = 2, \ldots, n\right]. \qquad (28)$$

The following fact is well known [6, Eq. (6.4.3), p. 135], and will turn out to be useful in a number of places: For any subset $I \subseteq \{1, \ldots, n\}$, we have

$$\begin{aligned} &\mathbb{P}\left[L_{n,k} > t_k,\ k \in I\right] \\ &= \left(1 - \sum_{k\in I} t_k\right)_+^n, \quad t_k \in [0,1],\ k \in I \end{aligned} \qquad (29)$$

where we have used the notation $x_+^n = x^n$ if $x \geq 0$ and $x_+^n = 0$ if $x \leq 0$. With the help of (29), the inclusion-exclusion formula (applied to (28)) easily yields the closed form expression

$$P(n;\tau) = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} (1 - k\tau)_+^n. \qquad (30)$$

This expression has been rediscovered by several authors, e.g., Godehardt and Jaworski [15, Cor. 1, p. 146], Desai and Manjunath [7] (as Eqn (8) with $z = 1$ and $r = \tau$), Ghasemi and Nader-Esfahani [13] and Gore [16]. See also Devroye's paper [9] for pointers to an older literature.

While certainly pleasing, the expression (30) is not well suited for the purpose of establishing (5). This has prompted us to introduce other representations for the probability of graph connectivity (2) in Sections VII and VIII.

VII. A RESULT BY LÉVY ON MAXIMAL SPACINGS

The special role of the critical range function $\tau^\star : \mathbb{N}_0 \to \mathbb{R}_+$ can already be surmised from a number of earlier limit results on maximal spacings. For each $n = 1, 2, \ldots$, the *maximal spacing* associated with the rvs $X_1, \ldots, X_n$ is the rv $M_n^\star$ given by

$$M_n^\star := \max\left(L_{n,k},\ k = 1, \ldots, n+1\right). \quad (31)$$

The convergence results given next are by now classical; they were originally given by Lévy [28] via geometric arguments, but have been rederived by Darling [5], and others; see Devroye's paper [9] for additional references. In order to state these results compactly, let Λ denote any $\mathbb{R}$-valued rv with probability distribution given by

$$\mathbb{P}\left[\Lambda \le x\right] = g(x) := e^{-e^{-x}}, \quad x \in \mathbb{R}. \quad (32)$$

Any rv distributed according to (32) is called a Gumbel rv.

Theorem 7.1: We have

$$\frac{M_n^\star}{\tau_n^\star} \xrightarrow{P}{}_n 1 \quad (33)$$

and

$$nM_n^\star - \log n \Longrightarrow_n \Lambda. \quad (34)$$

In the formulation used here, the boundary points of the interval $[0, 1]$ are not automatically nodes for the random graph (as is the case in the models discussed in [11], [12], [16]). Consequently, we modify the definition (31) by considering instead the rvs

$$M_n := \max\left(L_{n,k},\ k = 2, \ldots, n\right), \quad n = 2, 3, \ldots \quad (35)$$

Given $n = 2, 3, \ldots$, for each $\tau > 0$, the graph $\mathbb{G}(n; \tau)$ is connected if and only if $M_n \le \tau$, so that (28) becomes

$$P(n; \tau) = \mathbb{P}\left[M_n \le \tau\right]. \quad (36)$$

Thus, insights into (3) are likely to be gained through limit results for the sequence $\{M_n,\ n = 1, 2, \ldots\}$. In analogy with Theorem 7.1, we have the following convergence results.

Theorem 7.2: We have

$$\frac{M_n}{\tau_n^\star} \xrightarrow{P}{}_n 1 \quad (37)$$

and

$$nM_n - \log n \Longrightarrow_n \Lambda. \quad (38)$$

For future use we find it convenient to write

$$\Lambda_n := nM_n - \log n, \quad n = 2, 3, \ldots.$$

With the help of (36) we see that the convergence (38) is a simple rewriting of (5) since

$$P(n; \sigma_n(x)) = \mathbb{P}\left[M_n \le \sigma_n(x)\right] = \mathbb{P}\left[\Lambda_n \le x\right] \quad (39)$$

for each x in $\mathbb{R}$ and all $n = 2, 3, \ldots$. Given the pivotal role played by the convergence (38), we provide below a direct proof of this result by elementary arguments.

Proof. We note that (38) implies (37) by elementary properties of weak convergence: Indeed,

$$\frac{\Lambda_n}{\log n} = \frac{M_n}{\tau_n^\star} - 1, \quad n = 2, 3, \ldots$$

so that $\frac{\Lambda_n}{\log n} \Longrightarrow_n 0$ by virtue of (38). The desired conclusion (37) follows from the fact that weak convergence to a constant is equivalent to convergence in probability to that constant [4, p. 25].

The starting point for proving the convergence (38) is the following representation of the order statistics $X_{n,1}, \ldots, X_{n,n}$: Consider a collection of $\{\xi_j,\ j = 1, 2, \ldots\}$ of i.i.d. rvs which are exponentially distributed with unit parameter, and set

$$T_0 = 0,\ T_k = \xi_1 + \ldots + \xi_k, \quad k = 1, 2, \ldots.$$

For each $n = 1, 2, \ldots$, the stochastic equivalence

$$(X_{n,1}, \ldots, X_{n,n}) =_{st} \left(\frac{T_1}{T_{n+1}}, \ldots, \frac{T_n}{T_{n+1}}\right) \quad (40)$$

is known to hold [35, p. 403] (and references therein), whence

$$(L_{n,1}, \ldots, L_{n,n}) =_{st} \left(\frac{\xi_1}{T_{n+1}}, \ldots, \frac{\xi_n}{T_{n+1}}\right). \quad (41)$$

With the help of (41), simple algebra shows that

$$\begin{aligned} \Lambda_n \quad &=_{st} \quad \frac{n}{T_{n+1}} \cdot \max\left(\xi_k,\ k = 2, \ldots, n\right) - \log n \\ &= \quad \frac{n}{T_{n+1}} \cdot \left(\max\left(\xi_k,\ k = 2, \ldots, n\right) - \log n\right) \\ &\quad - \left(1 - \frac{n}{T_{n+1}}\right) \cdot \log n \end{aligned} \quad (42)$$

where we can write

$$\begin{aligned} &\left(1 - \frac{n}{T_{n+1}}\right) \cdot \log n \\ = \quad & \frac{n}{T_{n+1}} \left(\frac{T_{n+1}}{n} - 1\right) \cdot \log n \\ = \quad & \frac{n}{T_{n+1}} \cdot \sqrt{n} \left(\frac{T_{n+1}}{n} - 1\right) \cdot \frac{\log n}{\sqrt{n}}. \end{aligned} \quad (43)$$

Next let n go to infinity in (42) and (43). By the Strong law of large Numbers, we have

$$\lim_{n\to\infty} \frac{T_{n+1}}{n} = 1 \quad a.s.$$

whereas the Central Limit Theorem yields

$$\sqrt{n}\left(\frac{T_{n+1}}{n} - 1\right) \Longrightarrow_n U$$

with U denoting a zero-mean unit variance Gaussian rv. As a result, it is easy to see that

$$\frac{n}{T_{n+1}} \cdot \sqrt{n}\left(1 - \frac{T_{n+1}}{n}\right) \cdot \frac{\log n}{\sqrt{n}} \xrightarrow{P}_n 0 \quad (44)$$

by elementary properties of the convergence modes for rvs [4]. It is also well known [10, Example 3.2.7, p. 125] that

$$\max\left(\xi_k,\ k = 2, \ldots, n\right) - \log n \Longrightarrow_n \Lambda$$

so that

$$\frac{n}{T_{n+1}} \cdot \left(\max\left(\xi_k,\ k = 2, \ldots, n\right) - \log n\right) \Longrightarrow_n \Lambda. \quad (45)$$

Combining (44) and (45) leads to the desired result (38). ∎

To the best of the authors's knowledge, the method of proof used given here appears to be new. Theorem 7.1 can be established along the same lines; in fact it is not too difficult to see that Theorem 7.1 and Theorem 7.2 are equivalent.

VIII. A Poisson approximation

Fix $n = 2, 3, \ldots$ and τ in $(0, 1)$. For each $i = 1, \ldots, n$, node i is said to be a *breakpoint* node in the random graph $\mathbb{G}(n; \tau)$ whenever (i) it is not the leftmost node in $[0, 1]$ *and* (ii) there is no node in the random interval $[X_i - \tau, X_i]$. The number $C_n(\tau)$ of breakpoint nodes in $\mathbb{G}(n; \tau)$ is given by

$$C_n(\tau) := \sum_{k=2}^{n} \chi_{n,k}(\tau)$$

where the $\{0, 1\}$-valued rvs $\chi_{n,1}(\tau), \ldots, \chi_{n,n+1}(\tau)$ are the indicator functions defined by

$$\chi_{n,k}(\tau) := \mathbf{1}\left[L_{n,k} > \tau\right], \quad k = 1, \ldots, n + 1.$$

The graph $\mathbb{G}(n; \tau)$ is connected if and only if $C_n(\tau) = 0$, and the representation

$$P(n; \tau) = \mathbb{P}\left[C_n(\tau) = 0\right] \quad (46)$$

follows. This points to the possibility that the convergence (3) can be studied by developing a convergence theory for the rvs $\{C_n(\tau_n),\ n = 1, 2, \ldots\}$ under appropriate conditions on the range function $\tau : \mathbb{N}_0 \to \mathbb{R}$.

The main result along these lines takes the form of a Poisson approximation. Before presenting it, we need some additional notation: For any pair of probability mass functions (pmfs) $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ on $\mathbb{N}$, we define the total variation distance between $\boldsymbol{\mu} = (\mu(x),\ x \in \mathbb{N})$ and $\boldsymbol{\nu} = (\nu(x),\ x \in \mathbb{N})$ by

$$d_{TV}(\boldsymbol{\mu}; \boldsymbol{\nu}) := \frac{1}{2} \sum_{x=0}^{\infty} |\mu(x) - \nu(x)|.$$

It is easy to check that

$$|\mu(x) - \nu(x)| \le d_{TV}(\boldsymbol{\mu}; \boldsymbol{\nu}), \quad x \in \mathbb{N}. \quad (47)$$

Also, if X and Y are $\mathbb{N}$-valued rvs distributed according to the pmfs $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, respectively, we shall write $d_{TV}(X, Y)$ for $d_{TV}(\boldsymbol{\mu}, \boldsymbol{\nu})$. Throughout let $\Pi(\mu)$ denote a Poisson rv with parameter μ.

Theorem 8.1: For each $n = 2, 3, \ldots$ *and* τ *in the interval* $(0, 1)$*, we have*

$$d_{TV}(C_n(\tau); \Pi(\lambda_n(\tau))) \le B_n(\tau) \quad (48)$$

where the quantities $\lambda_n(\tau)$ *and* $B_n(\tau)$ *are given by*

$$\lambda_n(\tau) = \mathbb{E}\left[C_n(\tau)\right] = (n - 1)\left(1 - \tau\right)^n \quad (49)$$

and

$$B_n(\tau) = (n-1)\left(1 - \tau\right)^n - (n-2)\frac{(1 - 2\tau)_+^n}{(1 - \tau)^n}. \quad (50)$$

This result is established in Section X with the help of the Stein-Chen method [3] – In fact it is already given as Theorem 6.1 in [2, p. 83] with a slightly different proof.

IX. Poisson convergence and finite node models

An immediate, and important, consequence of Theorem 8.1 is the Poisson convergence discussed below.

Theorem 9.1: For each x in $\mathbb{R}$, we have

$$\lim_{n\to\infty} d_{TV}(C_n(\sigma_n(x)); \Pi(e^{-x})) = 0 \quad (51)$$

with the range function $\sigma(x) : \mathbb{N}_0 \to \mathbb{R}$ given by (4).

Proof. Fix $n = 2, 3, \ldots$ and τ in the interval $(0, 1)$; the notation is that of Theorem 8.1. The triangular inequality for the total variation distance yields

$$\begin{aligned} d_{TV}(C_n(\tau); \Pi(e^{-x})) &\leq d_{TV}(C_n(\tau); \Pi(\lambda_n(\tau))) \\ &\quad + d_{TV}(\Pi(\lambda_n(\tau)); \Pi(e^{-x})) \end{aligned}$$

and the estimate

$$d_{TV}(\Pi(\lambda_n(\tau)); \Pi(e^{-x})) \leq |\lambda_n(\tau) - e^{-x}|$$

is well known [29, p. 58]. Combining these two facts with the bound (48) we find

$$d_{TV}(C_n(\tau); \Pi(e^{-x})) \leq B_n(\tau) + |\lambda_n(\tau) - e^{-x}|. \quad (52)$$

Finally, replace τ by $\sigma_n(x)$ in (52), and let n go to infinity in the resulting inequality. It is a simple matter to check that $\lim_{n\to\infty} B_n(\sigma_n(x)) = 0$ and $\lim_{n\to\infty} |\lambda_n(\sigma_n(x)) - e^{-x}| = 0$. This establishes the convergence (51). ∎

Given the discrete nature of $\mathbb{N}$, the convergence (51) in the total variation distance is equivalent [3, p. 254] to the convergence $C_n(\sigma_n(x)) \Longrightarrow_n \Pi(e^{-x})$, so that

$$\lim_{n\to\infty} \mathbb{P}\left[C_n(\sigma_n(x)) = k\right] = \frac{(e^{-x})^k}{k!} e^{-e^{-x}} \quad (53)$$

for each $k = 0, 1, \ldots$. An equivalent result was already given by Godehardt and Jaworski [15, Thm. 12, p. 157] in terms of the number of connected components in $\mathbb{G}(n; \tau)$, which is given by $C_n(\tau)+1$. See also a similar result for the corresponding model on the unit circle [30, Thm. 8, p. 172].

The proof of Theorem 9.1 yields quite a bit more than the convergence (53): Since $\mathbb{P}\left[C_n(\sigma_n(x)) = 0\right] = P(n; \sigma_n(x))$, the bound (52) (with the help of (47)) provides a means to control the convergence (3) as it implies

$$\begin{aligned} & |P(n; \sigma_n(x)) - e^{-e^{-x}}| \\ \leq\ & B_n(\sigma_n(x)) + |\lambda_n(\sigma_n(x)) - e^{-x}| \end{aligned} \quad (54)$$

for *all* $n = 2, 3, \ldots$ and arbitrary x in $\mathbb{R}$. This approximation can be leveraged to deal with situations involving a finite number of nodes. For instance, the requirement

$$|P(n; \sigma_n(x)) - e^{-e^{-x}}| \leq \delta$$

for some $\delta > 0$ can be guaranteed by requiring

$$B_n(\sigma_n(x)) + |\lambda_n(\sigma_n(x)) - e^{-x}| \leq \delta. \quad (55)$$

In particular, with a in $(0, 1)$ and x_a as given by (21) we find

$$|P(n; \sigma_n(x_a)) - a| \leq \delta$$

provided (55) holds with x replaced by x_a.

Before moving on to the proof Theorem 8.1 in the next section, we note that the Poisson approximation also yields a *rate* of convergence in (38): Indeed, using (39) in the bound (54), we find

$$\begin{aligned} & \left|\mathbb{P}\left[nM_n - \log n \leq x\right] - e^{-e^{-x}}\right| \\ \leq\ & B_n(\sigma_n(x)) + |\lambda_n(\sigma_n(x)) - e^{-x}| \end{aligned} \quad (56)$$

for all $n = 2, 3, \ldots$ and x in $\mathbb{R}$. To the best of the authors's knowledge, this results appears to be new.

X. A proof of Theorem 8.1

We begin with a simple technical fact concerning binary valued rvs. For some $n = 2, 3, \ldots$, consider a collection of $\{0, 1\}$-valued rvs $\xi_1, \ldots, \xi_n$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Next, with $\mathcal{P}_n^\star$ denoting the collection of all non-empty subsets of $\{1, \ldots, n\}$, we define

$$P(K) := \mathbb{P}\left[\xi_k = 1,\ k \in K\right], \quad K \in \mathcal{P}_n^\star.$$

Lemma 10.1: The probabilities $\{P(K),\ K \in \mathcal{P}_n^\star\}$ collectively determine the joint pmf of the $\{0, 1\}^n$-valued rv $(\xi_1, \ldots, \xi_n)$.

Proof. Pick an arbitrary non-zero element $\boldsymbol{a} = (a_1, \ldots, a_n)$ in $\{0, 1\}^n$, and write $K(\boldsymbol{a}) = \{k = 1, \ldots, n : a_k = 1\}$ and $K(\boldsymbol{a})^c = \{k = 1, \ldots, n : a_k = 0\}$. Direct inspection yields

$$\mathbb{P}\left[\xi_k = a_k,\ k = 1, \ldots, n\right] = \mathbb{E}\left[\prod_{k \in K(\boldsymbol{a})} \xi_k \prod_{k \in K(\boldsymbol{a})^c} (1 - \xi_k)\right]$$

so that

$$\mathbb{P}\left[\xi_k = a_k,\ k=1,\ldots,n\right] = \sum_{K \in \mathcal{P}_n(\boldsymbol{a})} c(K)P(K)$$

for some appropriate collection $\mathcal{P}_n(\boldsymbol{a})$ of subsets of $\{1,\ldots,n\}$, and coefficients $\{c(K),\ K \in \mathcal{P}_n(\boldsymbol{a})\}$ taking values ± 1. We have used the convention $\prod_{k\in K(\boldsymbol{a})^c}(1-\xi_k) = 1$ when $K(\boldsymbol{a})$ is empty. The proof is completed upon noting that

$$\begin{aligned} & \mathbb{P}\left[\xi_1 = \ldots = \xi_n = 0\right] \\ = \ & 1 - \sum\nolimits_{\boldsymbol{a}\neq\mathbf{0}} \mathbb{P}\left[\xi_k = a_k,\ k=1,\ldots,n\right] \end{aligned}$$

where $\sum_{\boldsymbol{a}\neq\mathbf{0}}$ denotes summation over all non-zero elements $\boldsymbol{a}$ in $\{0,1\}^n$. ■

For each $n = 2,3,\ldots$ and $i = 1,\ldots,n$, let $K_{n,i}$ denote the set of integers $\{1,\ldots,n\}$ from which we have deleted i.

Lemma 10.2: Fix $n = 2,3,\ldots$ and τ in the interval $(0,1)$. For each $i = 1,\ldots,n$, we have the stochastic equivalence

$$\begin{aligned} & \left[(\chi_{n,k}(\tau),\ k \in K_{n,i}) | \chi_{n,i}(\tau) = 1\right] \\ =_{st} \ & \left(\chi_{n,k}\left(\frac{\tau}{1-\tau}\right),\ k \in K_{n,i}\right). \end{aligned} \tag{57}$$

Proof. Fix $i = 1,\ldots,n$. By Lemma 10.1, it suffices to show that

$$\begin{aligned} & \mathbb{P}\left[\chi_{n,k}(\tau) = 1,\ k \in K | \chi_{n,i}(\tau) = 1\right] \\ = \ & \mathbb{P}\left[\chi_{n,k}\left(\frac{\tau}{1-\tau}\right) = 1,\ k \in K\right] \end{aligned} \tag{58}$$

for any K in $\mathcal{P}_n^\star$ which does not contain i.

For any such K in $\mathcal{P}_n^\star$ which does not contain i, let $|K|$ denote its cardinality. Repeatedly making use of (29) we get

$$\begin{aligned} & \mathbb{P}\left[\chi_{n,k}(\tau) = 1,\ k \in K | \chi_{n,i}(\tau) = 1\right] \\ = \ & \mathbb{P}\left[L_{n,k} > \tau,\ k \in K | L_{n,i} > \tau\right] \\ = \ & \frac{\mathbb{P}\left[L_{n,k} > \tau,\ k \in K; L_{n,i} > \tau\right]}{\mathbb{P}\left[L_{n,i} > \tau\right]} \\ = \ & \frac{(1-(|K|+1)\tau)_+^n}{(1-\tau)^n} \end{aligned}$$

and

$$\begin{aligned} & \mathbb{P}\left[\chi_{n,k}\left(\frac{\tau}{1-\tau}\right) = 1,\ k \in K\right] \\ = \ & \mathbb{P}\left[L_{n,k} > \frac{\tau}{1-\tau},\ k \in K\right] \\ = \ & \left(1 - |K|\frac{\tau}{1-\tau}\right)_+^n \\ = \ & \frac{(1-(|K|+1)\tau)_+^n}{(1-\tau)^n}. \end{aligned}$$

The validity of (58) follows by direct comparison. ■

The coupling used in the proof of Theorem 6.1 in [2, p. 83] is given directly in terms of the spacings $L_{n,1},\ldots,L_{n,n}$ but is in the same spirit as the one given here. We now turn to the proof of Theorem 8.1. Fix $n = 2,3,\ldots$ and τ in the interval $(0,1)$: From (29) it is plain that the rvs $L_{n,1},\ldots,L_{n,n}$ are exchangeable; see also (41) for an additional confirmation. Leveraging this fact we see that

$$\mathbb{E}\left[C_n(\tau)\right] = (n-1)\mathbb{P}\left[L_{n,2} > \tau\right]$$

and

$$\begin{aligned} \mathbb{E}\left[C_n(\tau)^2\right] &= \sum_{k=2}^n \sum_{\ell=2}^n \mathbb{P}\left[L_{n,k} > \tau, L_{n,\ell} > \tau\right] \\ &= (n-1)\mathbb{P}\left[L_{n,2} > \tau\right] \\ &\quad + (n-1)(n-2)\mathbb{P}\left[L_{n,2} > \tau, L_{n,3} > \tau\right]. \end{aligned}$$

Repeatedly using (29), we find

$$\lambda_n(\tau) = \mathbb{E}\left[C_n(\tau)\right] = (n-1)\left(1-\tau\right)^n \tag{59}$$

and

$$\mathbb{E}\left[C_n(\tau)^2\right] = \lambda_n(\tau) + (n-1)(n-2)\left(1-2\tau\right)_+^n. \tag{60}$$

We then observe that the rvs $\chi_{n,1}(\tau),\ldots,\chi_{n,n}(\tau)$ are *negatively related* (in the technical sense given in [3, p. Defn. 2.1.1, p. 24]). This follows from (57) for all $i = 2,\ldots,n$ in view of the obvious coupling inequalities

$$\chi_{n,k}\left(\frac{\tau}{1-\tau}\right) \leq \chi_{n,k}(\tau), \quad k = 1,2,\ldots,n.$$

As a result, the basic Stein-Chen inequality [3, Cor. 2.C.2, p. 26] takes on the simpler form

$$\begin{aligned} & d_{TV}(C_n(\tau); \Pi(\lambda_n(\tau))) \\ \leq \ & \left(\frac{1-e^{-\lambda_n(\tau)}}{\lambda_n(\tau)}\right)\left(\lambda_n(\tau) - \mathrm{Var}[C_n(\tau)]\right) \\ \leq \ & \frac{\lambda_n(\tau) - \mathrm{Var}[C_n(\tau)]}{\lambda_n(\tau)}. \end{aligned} \tag{61}$$

Direct substitution from (59) and (60) gives

$$\begin{aligned} & \lambda_n(\tau) - \mathrm{Var}[C_n(\tau)] \\ = \; & \lambda_n(\tau) - \left(\mathbb{E}\left[C_n(\tau)^2\right] - \mathbb{E}\left[C_n(\tau)\right]^2\right) \\ = \; & -(n-1)(n-2)\left(1-2\tau\right)_+^n \\ & \quad + (n-1)^2 \left(1-\tau\right)^{2n}, \end{aligned}$$

and the conclusion (48) follows from (61) upon noting that

$$\frac{\lambda_n(\tau) - \mathrm{Var}[C_n(\tau)]}{\lambda_n(\tau)} = B_n(\tau).$$

■

XI. Additional comments

The zero-one laws for graph connectivity and the asymptotics of the transition widths were shown to all flow from the single key convergence result (5). We presented two very different approaches to this result – One relies on results by Lévy on maximal spacings, while the other exploits a Poisson convergence result for the number of breakpoint users under the scalings (4). It is worth pointing out that a similar situation exists for geometric random graphs in higher dimensions, with appropriate modifications and at the expense of significant technical difficulties.

The discussion begins with the observation that in higher dimensions it is easier to begin with the absence of isolated nodes rather than attack directly graph connectivity. The argument is then completed by showing the asymptotic equivalence between these two graphs properties [1], [33], [34]. The absence of isolated nodes is naturally associated with the *largest nearest-neighbor link* [1], [8], a quantity which in the one-dimensional case reduces to the maximal spacing. The appropriate version of Theorem 7.1 should then be a limiting result for the largest nearest-neighbor link (suitably normalized and centered); such results are indeed available in the literature under various assumptions, e.g., see [8, Thms. 1.2-1.5, p.68] and [33, Eqn. (3), p. 341]. As the notion of breakpoint nodes is meaningless in dimension two and higher, attention shifts instead to the number of isolated nodes and Poisson convergence is shown to hold for this quantity under the appropriate scaling [27].

At this point the reader may wonder as to what becomes of the results given here when the nodes are distributed independently on the interval $[0,1]$ according to a *non*-uniform distribution F. When F admits a non-vanishing density f, the strong zero-one law was developed in [23] (through the appropriate version of (37)) while the very strong zero-one law can be found in [24] (where it is established by means of the method of first and second moments). At the time of this writing the analogs of Theorem 5.1 and Corollary 5.2 are not known. The case when the density vanishes f at isolated points is explored for a particular family of densities in [22].

References

[1] M.J.B. Appel and R.P. Russo, "The connectivity of a graph on uniform points on $[0,1]^d$," *Statistics & Probability Letters* **60** (2002), pp. 351-357.

[2] A. D. Barbour and L. Holst, "Some applications of the Stein-Chen method for proving Poisson convergence," *Advances in Applied Probability* **21** (1989), pp. 74-90.

[3] A. D. Barbour, L. Holst and S. Janson, *Poisson Approximation,* Oxford Studies in Probability **2**, Oxford University Press, Oxford (UK), 1992.

[4] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York (NY), 1968.

[5] D.A. Darling, "On a class of problems related to the random division of an interval," *Annals of Mathematical Statistics* **24** (1953), pp. 239-253.

[6] H.A. David and H.N. Nagaraja, *Order Statistics* (Third Edition), Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken (NJ), 2003.

[7] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *IEEE Communications Letters* **6** (2002), pp. 437-439.

[8] H. Dette and N. Henze, "The limit distribution of the largest nearest-neighbour link in the unit d-cube," *Journal of Applied Probability* **26** (1989), pp. 67-80.

[9] L. Devroye, "Laws of the iterated logarithm for order statistics of uniform spacings," *The Annals of Probability* **9** (1981), pp. 860-867.

[10] P. Embrechts, C. Klüppelberg and Th. Mikosch, *Modelling Extremal Events*, Applications of Mathematics: Stochastic Modelling and Applied Probability **33**, Springer- Verlag, Berlin Heidelberg (1997).

[11] C.H. Foh and B.S. Lee, "A closed form network connectivity formula for one-dimensional MANETs," 2004 IEEE International Conference on Communications (ICC 2004), Paris (France), June 2004.

[12] C.H. Foh, G. Liu, B.S. Lee, B.-C. Seet, K.-J. Wong and C.P. Fu, "Network connectivity of one-dimensional MANETs with random waypoint movement," *IEEE Communications Letters* **9** (2005), pp. 31-33.

[13] A. Ghasemi and S. Nader-Esfahani, "Exact probability of connectivity in one-dimensional ad hoc wireless networks", *IEEE Communications Letters* **10** (2006), pp. 251-253.

[14] E. Godehardt, *Graphs as Structural Models: The Application of Graphs and Multigraphs in Cluster Analysis*, Vieweg, Braunschweig and Wiesbaden, 1990.

[15] E. Godehardt and J. Jaworski, "On the connectivity of a random interval graph," *Random Structures and Algorithms* **9** (1996), pp. 137-161.

[16] A.D. Gore, "Comments on "On the connectivity in finite ad hoc networks"," *IEEE Communications Letters* **10** (2006), pp. 88-90.

[17] A. Goel, S. Rai, and B. Krishnamachari, "Sharp thresholds for monotone properties in random geometric graphs," *Annals of Applied Probability* **15** (2005), pp. 2535-2552.

[18] P. Gupta and P.R. Kumar, "Critical power for asymptotic connectivity in wireless networks," Chapter in *Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, Edited by W.M. McEneany, G. Yin and Q. Zhang, Birkhäuser, Boston (MA), 1998.

[19] G. Han and A.M. Makowski, "Poisson convergence can yield very sharp transitions in geometric random graphs," Invited paper, in Proceedings of the Inaugural Workshop, Information Theory and Applications, University of California, San Diego (CA), February 2006.

[20] G. Han and A. M. Makowski, "Very sharp transitions in one-dimensional MANETs," in the Proceedings of the IEEE International Conference on Communications (ICC 2006), Istanbul (Turkey), June 2006.

[21] G. Han and A.M. Makowski, "A very strong zero-one law for connectivity in one-dimensional geometric random graphs," *IEEE Communications Letters* **11** (2007), pp. 55-57.

[22] G. Han and A. M. Makowski, "On the critical communication range under node placement with vanishing densities," in the Proceedings of the IEEE International Symposium on Information Theory (ISIT 2007), Nice (France), June 2007.

[23] G. Han and A.M. Makowski, "One-dimensional geometric random graphs with non-vanishing densities I: A strong zero-one law for connectivity," *IEEE Transactions on Information Theory* (2007), under second review.

[24] G. Han and A.M. Makowski, "One-dimensional geometric random graphs with non-vanishing densities II: A very strong zero-one law for connectivity," to be submitted to *IEEE Transactions on Information Theory* (2008).

[25] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, 2000.

[26] B. Krishnamachari, S. Wicker, S. Bejar and M. Pearlman, "Critical density thresholds in distributed wireless networks," Chapter in *Communications, Information and Network Security*, Eds. H. Bhargava, H.V. Poor, V. Tarokh and S. Yoon, Kluwer Publishers, December 2002.

[27] S.S. Kunniyur and S.S. Venkatesh, "Threshold functions, node isolation, and emergent lacunae in sensor networks," *IEEE Transactions on Information Theory* **IT-52** (2006), pp. 5352-5372.

[28] P. Lévy, "Sur la division d'un segment par des points choisis au hasard," *Comptes Rendus de l' Académie des Sciences de Paris* **208** (1939), pp. 147-149.

[29] T. Lindvall, *Lectures on the Coupling Method*, John Wiley & Sons, New York (NY), 1992.

[30] H. Maehara, "On the intersection graph of random arcles on a circle," in *Random Graphs'87*, pp. 159-173.

[31] G.L. McColm, "Threshold functions for random graphs on a line segment," *Combinatorics, Probability and Computing* **13** (2004), pp. 373-387.

[32] S. Muthukrishnan and G. Pandurangan, "The bin-covering technique for thresholding random geometric graph properties," in Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA 2005), Vancouver (BC), January 2005.

[33] M.D. Penrose, "The longest edge of the random minimal spanning tree," *Annals of Applied Probability* **7** (1997), pp. 340-361.

[34] M.D. Penrose, *Random Geometric Graphs*, Oxford Studies in Probability **5**, Oxford University Press, New York (NY), 2003.

[35] R. Pyke, "Spacings," *Journal of the Royal Statistical Society, Series B (Methodological)* **27** (1965), pp. 395-449.