k-connectivity in Random K-out graphs intersecting Erdős-Rényi Graphs *

Faruk Yavuz, Jun Zhao, Osman Yağan, and Virgil Gligor
fyavuz@andrew.cmu.edu, jun@cmu.edu, oyagan@andrew.cmu.edu, virgil@cmu.edu
Department of Electrical and Computer Engineering and CyLab
Carnegie Mellon University.

May 3, 2016

Abstract

We investigate k-connectivity in secure wireless sensor networks under the random pairwise key predistribution scheme with unreliable links. With wireless communication links modeled as independent on-off channels, this amounts to analyzing a *random graph model* formed by *intersecting* a random K-out graph and an Erdős-Rényi graph. We present conditions on how to scale the parameters of this intersection model so that the resulting graph is k-connected with probability approaching to one (resp. zero) as the number of nodes gets large. The resulting zero-one law is shown to improve and sharpen the previous result on the 1-connectivity of the same model. We also provide numerical results to support our analysis.

Keywords: Wireless Sensor Networks, Security, Key Predistribution, Random Key Graphs, k-connectivity.

1 Introduction

1.1 Motivation and Background

Random key predistribution schemes are widely regarded as the appropriate solution to ensure *secure* communications in resource-constrained wireless sensor networks (WSNs) [6, 28, 7, 11, 24, 25, 30, 31, 32, 33, 34, 36]. The idea of randomly assigning symmetric cryptographic keys to sensors before deployment has been introduced by Eschenauer and Gligor [11]. Since then many alternative methods have been proposed and random key predistribution schemes have been extensively studied in the last decade; e.g., see survey articles [22, 23] and references therein.

In this work, we consider the random *pairwise* key predistribution scheme proposed by Chan *et al.* [6], which works as follows: Prior to deployment, each of the n sensors is assigned to a set of K distinct sensors that are selected uniformly at random from among all other $(n-1)$ sensors. Then, two sensors are deemed to be *paired* if at least one of them has been assigned to the other.

*A preliminary version of this paper was presented at IEEE International Conference on Communications (ICC 2015), London (UK), June 2015.

For each such sensor pair, a unique pairwise key is then generated and loaded into the memory modules of the paired sensors together with their IDs. After deployment, any two sensors that share at least one common key can securely communicate over an existing wireless link. Precise implementation details are given in Section 2.1. Interest in the pairwise scheme arises from the fact that it has several advantages over other key predistribution techniques including the original one by Eschenauer and Gligor. For instance, the pairwise scheme enables distributed node-to-node authentication and quorum-based key revocation by construction [6]. These advantages have motivated a significant body of research devoted to analyzing various properties of the pairwise scheme including its connectivity [25, 28, 31, 32, 33, 34], resilience against node capture attacks [27], scalability [26, 30], etc.

This paper is devoted to analyzing the k-connectivity of WSNs under the pairwise scheme. Namely, given the randomness involved in the key predistribution mechanism, as well as in the availability of wireless communication links, we study how the scheme parameter K can be adjusted (possibly as a function of n) so that k-connectivity is achieved with high probability. A network (or a graph) is said to be k-connected if it remains connected despite the deletion of any $(k-1)$ nodes or links; a graph is simply deemed connected if it is 1-connected. Therefore, k-connectivity provides a guarantee of network reliability against the possible failures of $(k-1)$ sensors or links due to adversarial attacks, battery depletion, harsh environmental conditions, etc. Reliability against the failure of sensors or links is particularly important in WSN applications where sensors are deployed in hostile environments (e.g., battlefield surveillance), or, are unattended for long periods of time (e.g., environmental monitoring), or, are used in life-critical applications (e.g., patient monitoring).

Our approach is based on modeling the WSN by an appropriate random graph model and then studying its k-connectivity. The pairwise scheme naturally induces what is known in the literature as a *random K-out graph* [4, 14, 12]: each of the n vertices is assigned K arcs towards K distinct vertices that are selected uniformly at random, and then the orientation of the arcs is ignored; see Section 2.2 for details. Put differently, the edges in a random K-out graph $\mathbb{H}(n;K)$ identify pairs of sensors that share at least one key. In addition, we consider a wireless communication model comprising independent channels that are either *on* with probability p or *off* with probability $(1-p)$. The on/off channel model has been extensively used recently [24, 32, 36, 28, 33, 34] in the context of secure WSNs, and was also shown to exhibit very similar connectivity properties with the disk model [24, 32] (whereby any two sensors need to be within a certain distance of each other to have a wireless link in between). Thus, the communication connectivity in the network can be modeled by a classical Erdős-Rényi (ER) graph [4], $\mathbb{G}(n;p)$. Combining, our modeling framework consists of the *intersection* of a random K-out graph and an ER graph, denoted $\mathbb{H} \cap \mathbb{G}(n;K,p)$, or simply $\mathbb{I}(n;K,P)$, whereby pairs of nodes are connected by an edge if they share a key *and* have wireless channel in between that is on; i.e., edges in $\mathbb{I}(n;K,p)$ represent pairs of sensor that can establish a direct *secure* communication link in between[1].

[1]The main structural difference between these random graph models lies in the *regularity* of the resulting degree distributions. In particular, with a positive scalar m denoting the *mean* degree for all graphs involved, the degree of an arbitrary node in ER and K-out graph are distributed as $\text{Bin}(n-1,\frac{m}{n-1})$ and $\text{Bin}(n-1,\frac{m}{2(n-1)})+\frac{m}{2}$, respectively, meaning that K-out graphs have more regular degree distributions than ER graphs. As would be expected, the intersection of these graphs, defined as $\mathbb{I}(n;K,p)$, have more (resp. less) regular degree distributions than ER graphs (resp. K-out graphs). Put differently, we have $d_{\text{ER}} \succeq_{cx} d_{\mathbb{I}} \succeq_{cx} d_{\text{K-out}}$, where $\succeq_{cx}$ denotes the *convex ordering* and $d_{\text{ER}}, d_{\mathbb{I}}, d_{\text{K-out}}$ denote random variables standing for the degree of a node in ER graph, $\mathbb{I}(n;K,p)$, and K-out graph, respectively.

1.2 Contributions

With these in mind, we finalize in this work our analysis of the k-connectivity of $\mathbb{I}(n; K, p)$ that was initiated in [34]. There, the authors had established a zero-one law for a property that is related but *weaker* than k-connectivity, i.e., that of minimum node degree in $\mathbb{I}(n; K, p)$ being at least k. Here, we complement the analysis in [34] and establish a zero-one law for the k-connectivity of $\mathbb{I}(n; K, p)$, confirming the validity of [34, Conjecture 4.1]. Namely, we present scaling conditions on the parameters p and K with respect to the number of sensors n, such that $\mathbb{I}(n; K_n, p_n)$ is k-connected with probability approaching to zero, or one, as the number of sensors n gets large. We identify the *critical* scaling and show that it coincides with that obtained for the property of minimum node degree being at least k.

Our main result is shown to extend and improve upon the zero-one law for 1-connectivity obtained by Yağan and Makowski [28, 32]. In particular, we show that the critical scaling for 1-connectivity was incorrectly stated in [28, 32], leading to an unnecessary and unnatural additional condition that $\lim_{n\to\infty} p_n$ should exist. Here we identify the correct critical scaling and show that this additional condition is not required for the zero-one law for 1-connectivity (or, k-connectivity for $k = 2, 3, \ldots$) to hold. In fact, we show that a special case of our main result with $k = 1$ establishes a *stronger* and more fine-grained version of the zero-one law for 1-connectivity than obtained in [28, 32], while dropping the assumption on the existence of the limit of p_n; see Section 4.2 for details. Finally, we present numerical results for the finite node case under various parameter settings. This extensive simulation study suggests that although asymptotic in nature, our result can still help design WSNs (in a secure and reliable manner) in practical scenarios.

1.3 Notation and Conventions

A word on the notation: All statements involving limits 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 indicator function of an event E is denoted by $\mathbf{1}[E]$. In comparing the asymptotic behaviors of the sequences $\{a_n\}, \{b_n\}$, we use $a_n = o(b_n)$, $a_n = \omega(b_n)$, $a_n = O(b_n)$, $a_n = \Omega(b_n)$, and $a_n = \Theta(b_n)$, with their meaning in the standard Landau notation. Namely, we write $a_n = o(b_n)$ (resp. $a_n = \omega(n)$) as a shorthand for the relation $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ (resp. $\lim_{n\to\infty} \frac{a_n}{b_n} = \infty$), whereas $a_n = O(b_n)$ means that there exists $c > 0$ such that $a_n \leq c b_n$ for all n sufficiently large. Also, we have $a_n = \Omega(b_n)$ if $b_n = O(a_n)$, or equivalently, if there exists $c > 0$ such that $a_n \geq c b_n$ for all n sufficiently large. Finally, we write $a_n = \Theta(b_n)$ if we have $a_n = O(b_n)$ and $a_n = \Omega(b_n)$ at the same time.

1.4 Organization of the Paper

We organize the rest of the paper as follows. The paper is organized as follows: In Section 2, we give a formal model for the random pairwise key predistribution scheme of Chan et al. [6], and introduce the induced random K-out graph. In particular, the main model $\mathbb{I}(n; K, p)$ considered in this paper, i.e., the intersection of a random K-out graph with an ER graph, is introduced in Section 2.3. The main result of the paper concerning the k-connectivity of $\mathbb{I}(n; K, p)$ is presented in Section 3. In Section 4, we compare our results against the classical results of Erdős-Rényi and then against earlier results by Yağan and Makowski [32] on the 1-connectivity of $\mathbb{I}(n; K, p)$. Also in Section 4.3, we provide numerical results in support of our analytical results. The proof of the

main result is initiated in Section 5 where we present some basic ideas and preliminary results to be used throughout. The necessary steps for completing the proof are established in Sections 6, 7, and 8.

2 Basic Building Blocks

In what follows, we provide precise definitions of the intersection model to be studied. For convenience, these definitions are adopted by and large from [32, Sec. II and III].

2.1 The random pairwise key predistribution scheme

The random pairwise key predistribution scheme of Chan et al. [6] is motivated by the following advantages over the original EG scheme: (i) Even if some nodes are captured, the secrecy of the remaining nodes is *perfectly* preserved; (ii) Unlike earlier schemes, this pairwise scheme enables both node-to-node authentication and quorum-based revocation. See also [25] for a detailed comparison of these two classical key predistribution schemes.

We parametrize the pairwise key predistribution scheme by two positive integers n and K such that $K < n$. There are n nodes, labelled $i = 1, \ldots, n$, with unique ids $\mathrm{Id}_1, \ldots, \mathrm{Id}_n$. Write $\mathcal{V} = \{1, \ldots, n\}$ and set $\mathcal{V}_{-i} = \mathcal{V} - \{i\}$ for each $i = 1, \ldots, n$. With node i, we associate a subset $\Gamma_i(K)$ of K nodes selected uniformly at *random* from $\mathcal{V}_{-i}$. We say that each of the nodes in $\Gamma_i(K)$ is paired to node i. Thus, for any subset $A \subseteq \mathcal{V}_{-i}$, we require

$$\mathbb{P}\left[\Gamma_i(K) = A\right] = \begin{cases} \binom{n-1}{K}^{-1} & \text{if } |A| = K \\ \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Put differently, the selection of $\Gamma_i(K)$ is done *uniformly* amongst all subsets of $\mathcal{V}_{-i}$ which are of size K and we further assume that rvs $\Gamma_1(K), \ldots, \Gamma_n(K)$ are mutually independent.

Once this *offline* random pairing has been created, we construct the key rings $\Sigma_1(K), \ldots, \Sigma_n(K)$, one for each node, as follows: Assumed available is a collection of nK distinct cryptographic keys $\{\omega_{i|\ell},\ i = 1, \ldots, n;\ \ell = 1, \ldots, K\}$. Fix $i = 1, \ldots, n$ and let $\ell_i : \Gamma_i(K) \to \{1, \ldots, K\}$ denote a labeling of $\Gamma_i(K)$. For each node j in $\Gamma_i(K)$ paired to i, the cryptographic key $\omega_{i|\ell_i(j)}$ is associated with j. For instance, if the random set $\Gamma_i(K)$ is realized as $\{j_1, \ldots, j_K\}$ with $1 \le j_1 < \ldots < j_K \le n$, then an obvious labeling consists in $\ell_i(j_k) = k$ for each $k = 1, \ldots, K$ so that key $\omega_{i|k}$ is associated with node j_k. Of course other labelings are possible. Finally, with node j paired to node i, the pairwise key $\omega^{\star}_{n,ij} = [\mathrm{Id}_i|\mathrm{Id}_j|\omega_{i|\ell_i(j)}]$ is constructed and inserted in the memory modules of both nodes i and j. The key $\omega^{\star}_{n,ij}$ is assigned *exclusively* to the pair of nodes i and j, hence the terminology pairwise predistribution scheme. The key ring $\Sigma_i(K)$ of node i is the set

$$\Sigma_i(K) = \left\{\omega^{\star}_{n,ij},\ j \in \Gamma_i(K)\right\} \cup \left\{\omega^{\star}_{n,ji},\ \begin{array}{c} j = 1, \ldots, n \\ i \in \Gamma_j(K) \end{array}\right\}$$

Two nodes i and j, can secure an existing communication link if and only if $\Sigma_i(K) \cap \Sigma_j(K) \neq \emptyset$ which holds if at least one of the events $i \in \Gamma_j(K)$ or $j \in \Gamma_i(K)$ takes place. Namely, it is plain that

$$\left[\Sigma_i(K) \cap \Sigma_j(K) \neq \emptyset\right] = \left[i \in \Gamma_j(K)\right] \cup \left[j \in \Gamma_i(K)\right]$$

Both events can take place, in which case the memory modules of node i and j both contain the distinct keys $\omega^{\star}_{n,ij}$ and $\omega^{\star}_{n,ji}$. It is plain by construction that this scheme supports *distributed* node-to-node authentication.

2.2 Random K-out graphs

The pairwise key predistribution scheme naturally gives rise to the following class of random graphs: With $n = 2, 3, \ldots$ and positive integer $K < n$, we say that the distinct nodes i and j are K-adjacent, written $i \sim_K j$, if and only if they have at least one key in common in their key rings, namely

$$i \sim_K j \quad \text{iff} \quad \Sigma_i(K) \cap \Sigma_j(K) \neq \emptyset. \tag{2}$$

Let $\mathbb{H}(n; K)$ denote the undirected random graph on the vertex set $\{1, \ldots, n\}$ induced by the adjacency notion (2). This ensures that edges in $\mathbb{H}(n; K)$ represent pairs of sensors that have at least one cryptographic key in common, and thus that can securely communicate over an *existing* communication channel. Let $\lambda_n(K)$ define the edge assignment probability in $\mathbb{H}(n; K)$; i.e., we have

$$\mathbb{P}\left[i \sim_K j\right] = \lambda_n(K) \tag{3}$$

for any distinct $i, j \in \mathcal{V}$. It is easy to check that

$$\lambda_n(K) = 1 - \mathbb{P}\left[i \notin \Gamma_j(K) \ \cap \ j \notin \Gamma_i(K)\right] = 1 - \left(\frac{\binom{n-2}{K}}{\binom{n-1}{K}}\right)^2 = \frac{2K}{n-1} - \left(\frac{K}{n-1}\right)^2. \tag{4}$$

The random graph $\mathbb{H}(n; K)$ is known in the literature on random graphs as the random K-out graph [4, 14], or random K-orientable graph [12]. In these references, $\mathbb{H}(n; K)$ is defined in the following manner that can easily be seen to be equivalent to the adjacency condition (2): To each of the n vertices assign exactly K arcs towards K distinct vertices that are selected uniformly at random, and then ignore the orientation of the arcs. The directed version of this graph (i.e., with the orientation of the arcs preserved) has also been studied; e.g., see the work by Philips et al. [20], who showed that the *diameter* of the directed K-out graph concentrates almost surely on two values.

2.3 Intersection of $\mathbb{H}(n; K)$ with Erdős-Rényi (ER) graphs

As mentioned earlier, we assume a simple wireless communication model that consists of independent channels, each of which can be either on or off. Thus, with p in $[0, 1)$, let $\{B_{ij}(p), 1 \leq i < j \leq n\}$ denote i.i.d. $\{0, 1\}$-valued rvs with success probability p. The channel between nodes i and j is available (resp. up) with probability p and unavailable (resp. down) with the complementary probability $1 - p$. Distinct nodes i and j are said to be B-adjacent, written $i \sim_B j$, if $B_{ij}(p) = 1$. B-adjacency defines the standard ER graph $\mathbb{G}(n; p)$ on the vertex set $\{1, \ldots, n\}$ [4]. Obviously, $\mathbb{P}\left[i \sim_B j\right] = p$.

The random graph model studied here is obtained by *intersecting* the random graphs induced by the pairwise key predistribution scheme, and by the on-off communication model, respectively. Namely, we consider the intersection of $\mathbb{H}(n; K)$ with the ER graph $\mathbb{G}(n; p)$. In this case, distinct nodes i and j are said to be adjacent, written $i \sim j$, if and only they are both K-adjacent and B-adjacent, namely

$$i \sim j \quad \text{iff} \quad \Sigma_i(K) \cap \Sigma_j(K) \neq \emptyset \ \text{ and } \ B_{ij}(p) = 1. \tag{5}$$

The resulting *undirected* random graph defined on the vertex set $\{1,\dots,n\}$ through this notion of adjacency is denoted $\mathbb{I}(n;K,p)$; i.e., we have

$$\mathbb{I}(n;K,p) := \mathbb{H}(n;K) \cap \mathbb{G}(n;p)$$

The relevance of $\mathbb{I}(n;K,p)$ in the context of secure WSNs is now clear. Two nodes that are connected by an edge in $\mathbb{I}(n;K,p)$ share at least one cryptographic key *and* have a wireless link available to them, so that they can establish a *secure communication link*.

Throughout we assume the collections of rvs $\{\Gamma_1(K),\dots,\Gamma_n(K)\}$ and $\{B_{ij}(p), 1 \le i < j \le n\}$ to be independent, in which case the edge occurrence probability in $\mathbb{I}(n;K,p)$ is given by

$$\mathbb{P}[i \sim j] = \mathbb{P}[i \sim_K j]\,\mathbb{P}[i \sim_B j] = p\lambda_n(K). \tag{6}$$

2.4 k-connectivity vs. minimum node degree

Consider an undirected graph $\mathbb{G}$ defined on the vertices $1,\dots,n$. The degree of a node i, denoted d_i, is defined as the total number of links incident on it. Let δ be the *minimum* node degree in $\mathbb{G}$, i.e., $\delta = \min\{d_1,\dots,d_n\}$. We also let κ_v denote the *vertex connectivity* of $\mathbb{G}$ defined as the *minimum* number of vertices whose deletion renders $\mathbb{G}$ disconnected; if $\mathbb{G}$ is not connected we clearly have $\kappa_v = 0$. The *edge connectivity* κ_e is defined similarly in terms of edges, as the minimum number of edges that needs to be deleted to make $\mathbb{G}$ disconnected. For any graph $\mathbb{G}$, it is not difficult to see that [10]

$$\kappa_v \le \kappa_e \le \delta. \tag{7}$$

For each $k = 0, 1, 2, \dots$, we say that the graph $\mathbb{G}$ is k-vertex-connected if it holds that $\kappa_v \ge k$, whereas it is said to be k-edge-connected if $\kappa_e \ge k$. It is immediate from (7) that if a graph is k-vertex-connected, then it is also k-edge-connected, and its minimum degree is at least k. With this in mind, throughout we shall say that a graph is k-connected (without referring to vertex-connectivity) to refer to the fact that it is k-vertex-connected, and hence k-edge-connected. The terminology that has been in use thus far is now clear. If a graph is k-connected, then we have $\kappa_v, \kappa_e \ge k$, meaning that the graph can not be made disconnected even if any $k-1$ vertices or links are deleted. This is what makes the study of the property of k-connectivity appealing in seeking secure and reliable WSN designs.

2.5 Some Notation for Graph $\mathbb{I}(n;K,p)$

This section introduces some notation that will be used throughout. The complement of an event E is denoted by $\overline{E}$. For arbitrary sets S_x and S_y, the part of S_x that is not in S_y is denoted as $S_x \setminus S_y$. Sometimes, we find it convenient to use E_{ij} to denote the event that nodes i and j have an edge in between in $\mathbb{I}(n;K,p)$; i.e.,

$$E_{ij} := [i \sim j]$$

Hence, $\overline{E_{ij}}$ denotes the event that there is no edge between nodes i and j.

3 The results

To fix the terminology, we refer to any mapping $K : \mathbb{N}_0 \to \mathbb{N}_0$ as a *scaling* (for random K-out graphs) provided it satisfies the natural conditions

$$K_n < n, \quad n = 1, 2, \dots. \tag{8}$$

Similarly, we let any mapping $p : \mathbb{N}_0 \to [0,1]$ define a scaling for ER graphs. To lighten the notation we often group the parameters K and p into the ordered pair $\theta \equiv (K,p)$. Hence, a mapping $\theta : \mathbb{N}_0 \to \mathbb{N}_0 \times [0,1]$ defines a scaling for the intersection graph $\mathbb{I}(n;\theta)$ provided that the condition (8) holds. Oftentimes, we denote the scalings simply as $\{K_n\}_{n\geq 1}$, $\{p_n\}_{n\geq 1}$, and $\{\theta_n\}_{n\geq 1}$ without explicitly stating the associated mappings.

Theorem 3.1 *Consider scalings* $\{K_n\}_{n\geq 1}$ *and* $\{p_n\}_{n\geq 1}$ *such that* $\lim_{n\to\infty}(n-2K_n) = \infty$ *and* $\limsup_{n\to\infty} p_n < 1$. *With the sequence* $\gamma : \mathbb{N}_0 \to \mathbb{R}$ *defined through*

$$p_n K_n \left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) = \log n + (k-1)\log\log n + \gamma_n \tag{9}$$

we have

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ is } k\text{-connected }\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \gamma_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \gamma_n = +\infty. \end{cases} \tag{10}$$

Theorem 3.1 establishes a zero-one law for k-connectivity for random K-out graphs intersecting ER graphs, i.e., for $\mathbb{I}(n;\theta_n)$. This result was conjectured by the authors in [33, 34, Conjecture 4.1], where an analog of Theorem 3.1 was obtained for the property that minimum node degree in $\mathbb{I}(n;\theta_n)$ is at least k; see Theorem 5.1 in Section 5.1. Thus, we conclude that $\mathbb{I}(n;\theta_n)$ provides one more instance where the zero-one laws for k-connectivity and minimum node degree being at least k coincide; other examples include ER graphs [10], random key graphs [36, 37], certain classes of random geometric graphs [18], etc.

The proof of Theorem 3.1 is given in Sections 5-8. The main ideas leading to this proof are reminiscent of those used for establishing a zero-one law for k-connectivity in ER graphs [10]. However, the proof given here is much more involved, mainly because of the intricate dependencies between the edge events $\{E_{ij}\}_{1\leq i\leq j\leq n}$; see Section 7.1 for details.

We now argue that the extra conditions enforced by Theorem 3.1 are mild and do not preclude their application in realistic WSN scenarios. First, the condition $\limsup_{n\to\infty} p_n < 1$ enforces that wireless communication channels between nodes do not become available with probability one as n gets large. The situation $\limsup_{n\to\infty} p_n = 1$ that is not covered by Theorem 3.1 is reminiscent of the *full visibility* case considered in [31], and is not likely to hold in practice. In fact, it may be expected that p_n goes to zero as the number of nodes increases due to interference associated with a large number of nodes communicating simultaneously. Second, the condition $\lim_{n\to\infty}(n-2K_n) = \infty$ can be satisfied $2K_n \leq cn$ for some $c < 1$. Given that $2K_n$ is equal to the mean number of keys stored per sensor in the pairwise scheme [30], this condition needs to hold in any practical WSN scenario due to limited memory and computational capability of the sensors. In fact, Di Pietro et al. [7] noted that key ring sizes on the order of $\log n$ are feasible for WSNs.

We now present a simple corollary of Theorem 3.1, that will help compare our main result with the classical results of ER [10].

Corollary 3.2 *Consider scalings* $\{K_n\}_{n\geq 1}$ *and* $\{p_n\}_{n\geq 1}$ *such that* $\lim_{n\to\infty}(n-2K_n) = \infty$ *and* $\limsup_{n\to\infty} p_n < 1$. *With the sequence* $\gamma : \mathbb{N}_0 \to \mathbb{R}$ *defined through*

$$\frac{p_n K_n}{n-1}\left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) = \frac{\log n + (k-1)\log\log n + \gamma_n}{n} \tag{11}$$

we have

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ is } k\text{-connected }\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} \gamma_n = -\infty \\ \\ 1 & \text{if } \lim_{n\to\infty} \gamma_n = +\infty. \end{cases} \tag{12}$$

The proof of Corollary 3.2 is trivial and omitted here for brevity.

The main benefit of having Corollary 3.2 is to obtain the zero-one law for k-connectivity under the scaling (11), where the left-hand side is easily comparable with the link probability $p_n\lambda_n(K_n)$ in $\mathbb{I}(n;\theta_n)$; with the help of (4) we easily see that

$$\mathbb{P}\left[i \sim j\right] = p_n\lambda_n(K_n) = \frac{p_n K_n}{n-1}\left(2 - \frac{K_n}{n-1}\right). \tag{13}$$

4 Comments and Discussion

4.1 Comparison with Erdős-Rényi (ER) Graphs

For each p in $[0,1]$ and $n = 2, 3, \ldots$, let $\mathbb{G}(n;p)$ denote the ER graph on the vertex set $\{1,\ldots,n\}$ with edge probability p. It is known that edge assignments are mutually independent in $\mathbb{G}(n;p)$, whereas they are shown to be *negatively associated* in $\mathbb{H}(n;K)$ in sense of Joag-Dev and Proschan [15]; see Section 7.1 for details. Therefore neither the random K-out graph $\mathbb{H}(n;K)$ nor the intersection model $\mathbb{I}(n;\theta)$ can be equated with $\mathbb{G}(n;p)$. This is so even when the parameters p and K are selected so that the edge assignment probabilities in these graphs coincide, say $\lambda(n;K) = p$.

Although random K-out graphs are not statistically equivalent to ER graphs, they still exhibit certain similarities. For instance, the following k-connectivity result for ER graphs is well-known [10]: For any scaling $\{p_n\}_{n\geq 1}$, define $\gamma : \mathbb{N}_0 \to \mathbb{R}$ through

$$p_n = \frac{\log n + (k-1)\log\log n + \gamma_n}{n}. \tag{14}$$

It holds that

$$\lim_{n\to\infty} \mathbb{P}\left[\ \mathbb{G}(n;p_n) \text{ is } k\text{-connected }\right] = \begin{cases} 0 & \text{if } \ \gamma_n \to -\infty \\ \\ 1 & \text{if } \ \gamma_n \to +\infty. \end{cases} \tag{15}$$

We now compare this with our main result by means of Corollary 3.2. Notice that the right-hand sides of the scalings (11) and (14) are exactly the same, and so are the corresponding zero-one laws (12) and (15), respectively. In the case of the ER graph $\mathbb{G}(n;p_n)$, the left-hand side of (14) coincides with the edge probability p_n. However, for $\mathbb{I}(n;\theta_n)$ the left-hand side of (11) is not equal to the corresponding edge probability $p_n\lambda_n(K_n)$ (viz. (13)). To see this, we recall (4) and use the fact that $\log(1-p_n) \leq -p_n$ to get

$$\frac{p_n K_n}{n-1}\left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) \geq p_n\lambda_n(K_n).$$

Hence, in ER graphs the threshold of k-connectivity appears when the link probability p_n is compared against $(\log n + (k-1)\log\log n)/n$. In $\mathbb{I}(n;\theta_n)$, our result shows that the threshold appears

when a quantity that is always larger than the link probability $p_n\lambda_n(K_n)$ is compared against $(\log n + (k-1)\log\log n)/n$. Thus, we see that $\mathbb{I}(n;\theta_n)$ tends to exhibit the k-connectivity property *easier* than ER graphs; i.e., k-connectivity can be ensured by a smaller link probability between nodes (which leads to a smaller average node degree). This is due to the fact that $\mathbb{I}(n;\theta_n)$ has a more *regular* degree distribution than ER graphs as discussed in Section 1.1.

The situation is more intricate when it holds that $\lim_{n\to\infty} p_n = 0$, whereby

$$\log(1-p_n) = -p_n - \frac{p_n^2}{2}(1+o(1)).$$

This gives

$$\begin{aligned}
\frac{p_nK_n}{n-1}\left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) &= \frac{p_nK_n}{n-1}\left(2 - \frac{K_n}{n-1} + \frac{p_n}{2}(1+o(1))\right) \\
&= \frac{p_nK_n}{n-1}\left(2-\frac{K_n}{n-1}\right)\left(1+\frac{p_n}{2}\cdot\frac{1+o(1)}{2-\frac{K_n}{n-1}}\right) \\
&= p_n\lambda_n(K_n)(1+\Theta(p_n)) \qquad (16)\\
&= p_n\lambda_n(K_n)(1+o(1)), \qquad (17)
\end{aligned}$$

where in (16), we used the fact that $1 \le 2 - \frac{K_n}{n-1} \le 2$ since $K_n \le n-1$. This interestingly shows that in the practically relevant case where wireless channels become weaker as n gets large, the threshold for the k-connectivity of $\mathbb{I}(n;\theta_n)$ appears when a quantity that is asymptotically equivalent to the *link probability* is compared against $(\log n + (k-1)\log\log n)/n$. This suggests that with $\lim_{n\to\infty} p_n = 0$, $\mathbb{I}(n;\theta_n)$ becomes reminiscent of the ER graph as far as the k-connectivity property is concerned; a similar observation was made in [32] for the threshold of 1-connectivity and absence of isolated nodes.

Nevertheless, a detailed analysis shows that even under $\lim_{n\to\infty} p_n = 0$, the zero-one laws for k-connectivity in ER graphs and $\mathbb{I}(n;\theta_n)$ are *not* exactly analogous. This is due to the fact that $(1+o(1))$ term appearing in (17) may change the limit behavior of the sequence γ_n appearing in (11). In particular, comparing (16) and (11), we have

$$\begin{aligned}
\gamma_n &= np_n\lambda_n(K_n)(1+\Theta(p_n)) - \log n - (k-1)\log\log n \\
&= np_n\lambda_n(K_n) - \log n - (k-1)\log\log n + \Theta(np_n^2\lambda_n(K_n)) \\
&= np_n\lambda_n(K_n) - \log n - (k-1)\log\log n + \Theta(K_np_n^2)
\end{aligned}$$

as we note that $\lambda_n(K_n) = \Theta(K_n/n)$. We now see that, even under $\lim_{n\to\infty} p_n = 0$ the two results, (15) under (14) and (12) under (11), may be deemed analogous if and only if the last term $K_np_n^2$ is bounded, i.e., $K_np_n^2 = O(1)$; note that only then this last term is guaranteed to *not* affect whether $\lim_{n\to\infty}\gamma_n = \pm\infty$. In conclusion, we see that for the two graphs $\mathbb{G}(n;p_n)$ and $\mathbb{I}(n;K_n,p_n)$ to exhibit asymptotically the same behavior for the property of k-connectivity, the parameter scalings should satisfy

$$p_n = o(1) \qquad \textit{and} \qquad K_np_n^2 = O(1).$$

4.2 Comparison with results by Yağan and Makowski for $k = 1$

We now compare our results with those by Yağan and Makowski [32] who established zero-one laws for 1-connectivity and for absence of isolated nodes (i.e., absence of nodes with degree zero) in

$\mathbb{I}(n;\theta)$. Here, we present their result in a slightly different form: Consider scalings $\{K_n\}_{n\geq 1}$ and $\{p_n\}_{n\geq 1}$ such that

$$p_n K_n \left(2 - \frac{K_n}{n-1}\right) \left(\frac{1 - \frac{\log(1-p_n)}{p_n}}{2}\right) \sim c \log n, \tag{18}$$

for some $c > 0$. Assume also that $\lim_{n\to\infty} p_n = p^\star$ exists. Then, we have

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H} \cap \mathbb{G}(n;\theta_n) \text{ is connected }\right] = \begin{cases} 0 & \text{if } \ c < 1 \\ \\ 1 & \text{if } \ c > 1. \end{cases} \tag{19}$$

To better compare this result with ours, we set $k = 1$ and rewrite our scaling condition (9) as

$$p_n K_n \left(2 - \frac{K_n}{n-1}\right) \left(\frac{1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}}{2 - \frac{K_n}{n-1}}\right) = \log n + \gamma_n \tag{20}$$

under which Theorem 3.1 gives

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \ \text{ is connected }\right] = \begin{cases} 0 & \text{if } \ \gamma_n \to -\infty \\ \\ 1 & \text{if } \ \gamma_n \to +\infty. \end{cases}$$

We now argue how our result on 1-connectivity constitutes an improvement on the result of [32]. This discussion will also reveal why the assumption that limit $\lim_{n\to\infty} p_n = p^\star$ exists was crucial in establishing (19) under (18) in the earlier work [32]. First of all, it is clear that if $p^\star = 0$, then

$$\lim_{n\to\infty} \left(\frac{1 - \frac{\log(1-p_n)}{p_n}}{2}\right) = 1 = \lim_{n\to\infty} \left(\frac{1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}}{2 - \frac{K_n}{n-1}}\right)$$

so that the left hand sides of (20) and (18) are asymptotically equivalent. Next, if $p^\star > 0$, then it follows that $K_n = O(\log n)$ (see [32]) under (18). This again yields the asymptotical equivalence of the left hand sides of (20) and (18). Therefore, under the assumption that p_n has a limit, a scaling condition that is *equivalent* to (18) is given by

$$p_n K_n \left(2 - \frac{K_n}{n-1}\right) \left(\frac{1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}}{2 - \frac{K_n}{n-1}}\right) \sim c \log n, \tag{21}$$

with the results (19) unchanged.

Comparing (20) with (21), we see that our 1-connectivity result is more fine-grained than the one given in [32]. In a nutshell, the scaling condition (21) enforced in [32] requires a deviation of $\gamma_n = \pm\Omega(\log n)$ (from the threshold $\log n$) to get the zero-one law, whereas in our formulation (20), it suffices to have an unbounded deviation; i.e., any $\gamma_n = \omega(1)$ will do. Put differently, we cover the case $c = 1$ in (21) that is not covered in the result (19) given in [32]. In fact, we show that if $c = 1$ in (21), then $\mathbb{I}(n;\theta_n)$ could be almost surely connected or almost surely not connected depending on the limit of γ_n defined in (20). Furthermore, our main result Theorem 3.1 indicates that if (21) holds with $c > 1$, then $\mathbb{I}(n;\theta_n)$ will be not only 1-connected, but also k-connected (almost surely) for all $k = 1, 2, \ldots$.

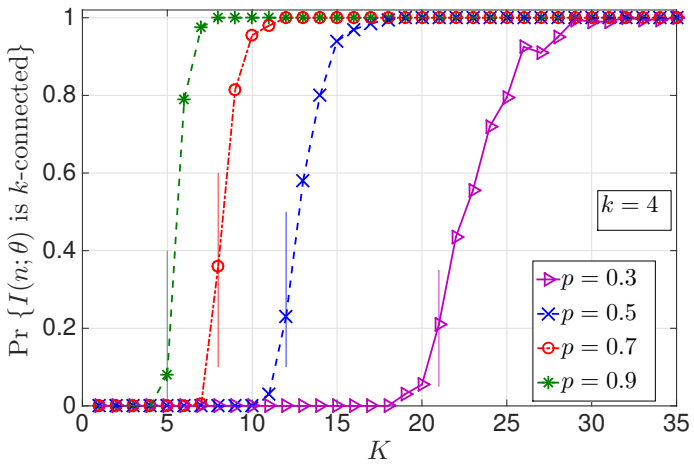

Figure 1: *Probability that* $\mathbb{I}(n; K, p)$ *is* 4*-connected as a function of* K *with* $p = 0.3, 0.5, 0.7, 0.9$ *and* $n = 2000$.

Collecting, our contributions in this paper improve the results in [32] several directions. First, we extend the results on the 1-connectivity of the model $\mathbb{I}(n; \theta_n)$ to k-connectivity with arbitrary $k = 1, 2, \ldots$. As discussed before, the k-connectivity property quantifies the reliability of the network against node or edge removals, and is desirable in a number of applications including wireless sensor networks. Second, with $k = 1$, our main result sharpens and improves the zero-one law for 1-connectivity of $\mathbb{I}(n; \theta_n)$. In particular, our result does not require the unnatural condition that $\lim_{n\to\infty} p_n$ should exist, and does establish a sharper *phase transition* result with deviation functions of the form $\gamma_n = \pm(\log n)$ being able to change the phase of the graph from being disconnected to connected. Finally, our discussion also reveals that the aforementioned unnatural condition was needed in [32] since the *critical* scaling for the connectivity of $\mathbb{I}(n; \theta_n)$ was not identified correctly; i.e., (18) was used instead of (21).

4.3 Numerical results

In this Section, we present numerical results to check the validity of Theorem 3.1 in the non-asymptotic regime, i.e., when parameter values are set in accordance with real-world wireless sensor network scenarios. In all experiments, we fix the number of nodes at $n = 2000$. Then for a given parameter pair (K, p), we generate 200 independent samples of the graph $\mathbb{I}(n; K, p)$ and count the number of times (out of a possible 200) that the obtained graph is k-connected for $k = 1, 2, \ldots$. Dividing the counts by 200, we obtain the (empirical) probabilities for k-connectivity.

For brevity, we display only three figures, namely Figures 1, 2, and 3. Each time, one of the parameters (k, p, K) is fixed at a typical value, another is varied through a wide range, while the third one is set to four different values of interest. In doing so, our goal is to understand the sensitivity of the reliability of the network (as quantified by the probability of k-connectivity) to the variations in the network parameters p and K. In particular, we want to check whether the observed sensitivity is in parallel with our analytical results given in Theorem 3.1. To that end, in each curve, we include a vertical dashed line that stands for the *critical* value of the varying parameter (i.e., of K in Figures 1 and 2, and of p in Figure 3) that results in a change of sign in the

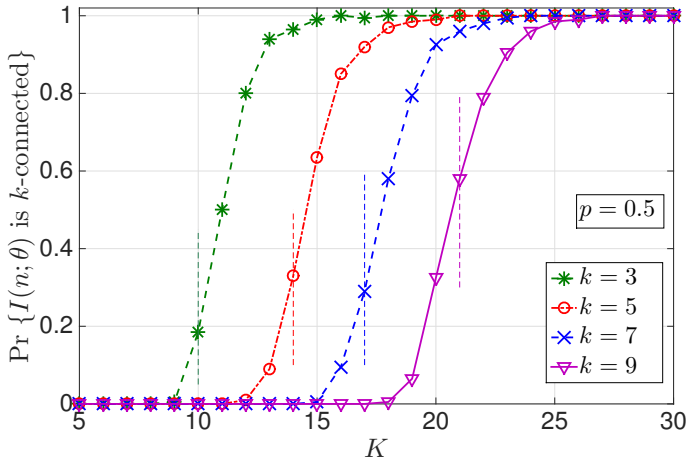

Figure 2: *Probability that* $\mathbb{I}(n; K, p)$ *is* k*-connected as a function of* K *with* $p = 0.5$ *and* $n = 2000$.

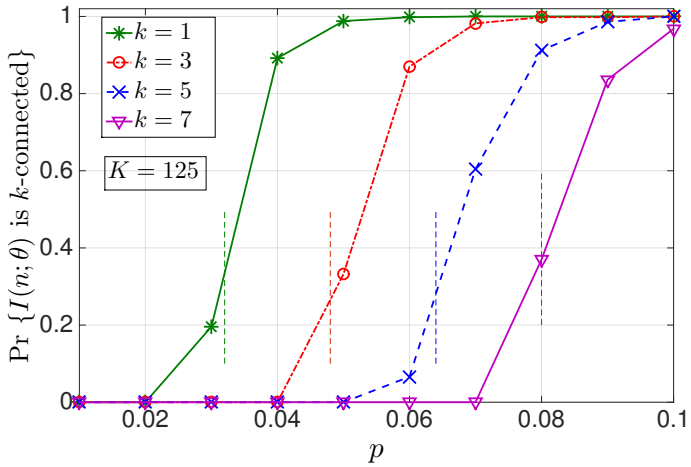

Figure 3: *Probability that* $\mathbb{I}(n; K, p)$ *is* k*-connected as a function of* p *with* $K = 125$ *and* $n = 2000$.

sequence γ_n given via (9). As an example, in Figure 1, vertical dashed lines stand for the minimum integer value of K that satisfies

$$pK\left(1 - \frac{\log(1-p)}{p} - \frac{K}{n-1}\right) > \log n + 3\log\log n. \tag{22}$$

Our main conclusions from the numerical results are twofold. First, we see that the the sharp phase transition behavior suggested by Theorem 3.1 is already observable with $n = 2000$ nodes. Namely, the probability that $\mathbb{I}(n; K, p)$ is k-connected transitions from *zero* to *one* as the parameter K (or, p) varies very slightly from a certain value. Second, we see that those critical values match well the vertical dashed lines obtained from Theorem 3.1. This prompts us to conclude that simulation outcomes are in good agreement with the analytical results.

5 Basic Ideas for Proving Theorem 3.1

5.1 From minimum node degree to k-connectivity

The proof of Theorem 3.1 will take advantage of the relationship between minimum node degree and k-connectivity in the following way. First, observe from (7) that

$$\mathbb{P}\left[\kappa_v \geq k\right] \leq \mathbb{P}\left[\delta \geq k\right] \tag{23}$$

and that

$$\begin{aligned}
\mathbb{P}[\kappa_v \geq k] &= \mathbb{P}[\delta \geq k] - \mathbb{P}[\,(\kappa_v < k) \cap (\delta \geq k)\,] \\
&= \mathbb{P}[\,\delta \geq k\,] - \bigcup_{\ell=0}^{k-1} \mathbb{P}[\,(\kappa_v = \ell) \cap (\delta \geq k)\,]. \\
&\geq \mathbb{P}[\,\delta \geq k\,] - \sum_{\ell=0}^{k-1} \mathbb{P}[\,(\kappa_v = \ell) \cap (\delta > \ell)\,]. \qquad (24)
\end{aligned}$$

The bounds (23)-(24) will pave the way to establishing the zero-one law for k-connectivity (i.e., for the property that $\kappa_v \geq k$) in $\mathbb{I}(n;\theta_n)$. Of particular importance will be the following zero-one law for the minimum node degree of $\mathbb{I}(n;\theta_n)$ (i.e., for the property that $\delta \geq k$) established by us in [33].

Theorem 5.1 ([33, 34]) *Consider scalings $\{K_n\}_{n\geq 1}$ and $\{p_n\}_{n\geq 1}$ such that $\lim_{n\to\infty}(n-2K_n) = \infty$ and $\limsup_{n\to\infty} p_n < 1$. With the sequence $\gamma : \mathbb{N}_0 \to \mathbb{R}$ defined through (9), we have*

$$\lim_{n\to\infty} \mathbb{P}\left[\begin{matrix}\textit{Min node degree of} \\ \mathbb{I}(n;\theta_n) \textit{ is at least } k\end{matrix}\right] = \begin{cases} 0 & \textit{if } \lim\limits_{n\to\infty} \gamma_n = -\infty \\ \\ 1 & \textit{if } \lim\limits_{n\to\infty} \gamma_n = +\infty. \end{cases} \tag{25}$$

It is now clear how to proceed. Pick a scaling $\theta : \mathbb{N}_0 \to \mathbb{N}_0 \times [0,1]$ as in the statements of Theorem 3.1 and Theorem 5.1. If it holds that $\lim\limits_{n\to\infty} \gamma_n = -\infty$, then we get from Theorem 5.1 that

$$\lim_{n\to\infty} \mathbb{P}\left[\delta(n;\theta_n) \geq k\right] = 0,$$

with $\delta(n;\theta_n)$ denoting the minimum node degree in $\mathbb{I}(n;\theta_n)$. From (23), this already establishes the zero-law for k-connectivity in (10), namely that

$$\lim_{n\to\infty} \mathbb{P}\left[\kappa_v(n;\theta_n) \geq k\right] = 0 \quad \text{if } \lim_{n\to\infty} \gamma_n = -\infty,$$

with $\kappa_v(n;\theta_n)$ denoting the vertex connectivity of $\mathbb{I}(n;\theta_n)$. Hence, we only need to establish the one-law of Theorem 3.1. Now, assume that $\lim\limits_{n\to\infty} \gamma_n = +\infty$. From Theorem 5.1 we have that

$$\lim_{n\to\infty} \mathbb{P}\left[\delta(n;\theta_n) \geq k\right] = 1. \tag{26}$$

Using this in (24), we see that the one-law of (10),

$$\lim_{n\to\infty} \mathbb{P}\left[\kappa_v(n;\theta_n) \geq k\right] = 1 \quad \text{if } \lim_{n\to\infty} \gamma_n = +\infty,$$

will follow if we show under $\lim_{n\to\infty} \gamma_n = +\infty$ that

$$\lim_{n\to\infty} \mathbb{P}\left[(\kappa_v(n;\theta_n) = \ell) \ \cap \ (\delta(n;\theta_n) > \ell)\right] = 0 \ \text{ for each } \ell = 0, 1, \ldots, k-1. \tag{27}$$

The rest of the paper is devoted to establishing (27), i.e., the fact that almost surely $\mathbb{I}(n;\theta_n)$ can not be made disconnected by deleting ℓ vertices when all of its nodes have degree larger than ℓ. This will be done by finding a sufficiently tight upper bound on the probability $\mathbb{P}[\kappa_v(n;\theta_n) = \ell \ \cap \ \delta(n;\theta_n) > \ell]$ and then showing that it goes to zero as $n \to \infty$. The approach is similar to the one used for proving the one-law for k-connectivity in ER graphs [4, p. 164].

5.2 Confining γ_n

In this section, we show that in establishing (27), we can restrict our attention to a subclass of structured scalings, referred throughout as *admissible* scalings. Recall that any mapping $(K, p) : \mathbb{N}_0 \to \mathbb{N}_0 \times [0, 1)$ defines a scaling provided that the condition (8) is satisfied. We now introduce the notion of an admissible scaling.

Definition 5.2 *A mapping $(K, p) : \mathbb{N}_0 \to \mathbb{N}_0 \times [0, 1)$ is said to be an admissible scaling if (8) holds, and the sequence $\gamma : \mathbb{N}_0 \to \mathbb{R}$ defined through (9) satisfies*

$$\lim_{n\to\infty} \frac{\gamma_n}{\log n} = 0. \tag{28}$$

The notion of the admissible scaling for the graph $\mathbb{I}(n;\theta_n)$ has been introduced by the authors in [34]. There it was shown via a coupling argument that (see [34, Propositions 5.3, 5.4, and 5.5] for any *monotone increasing*[2] property of the graph $\mathbb{I}(n;\theta_n)$, a zero-one law established under admissible scalings implies a zero-one law for general scalings. We present this result in more precise form below:

Lemma 5.3 ([34]) *Consider any scaling $(K, p) : \mathbb{N}_0 \to \mathbb{N}_0 \times [0, 1)$ such that $\lim_{n\to\infty}(n - 2K_n) = \infty$ and $\limsup_{n\to\infty} p_n < 1$, and let the sequence $\gamma : \mathbb{N}_0 \to \mathbb{R}$ be defined through (9). Then, for any monotone increasing graph property $\mathcal{P}$, we have the following:*

a) If

$$\lim_{n\to\infty} \gamma_n = +\infty \ \text{ and } \ (28) \implies \lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ has property } \mathcal{P}\right] = 1$$

then

$$\lim_{n\to\infty} \gamma_n = +\infty \implies \lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ has property } \mathcal{P}\right] = 1$$

b) If

$$\lim_{n\to\infty} \gamma_n = -\infty \ \text{ and } \ (28) \implies \lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ has property } \mathcal{P}\right] = 0$$

then

$$\lim_{n\to\infty} \gamma_n = -\infty \implies \lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ has property } \mathcal{P}\right] = 0.$$

Lemma 5.3 shows that for the purposes of our discussion, there is no loss of generality in considering only the admissible scalings. In view of this, the desired result (27) and hence the main result of our paper will be established upon proving the following result.

[2] A graph property is called monotone increasing if it holds under the addition of edges in a graph.

Proposition 5.4 *Consider scalings* $\{K_n\}_{n\geq 1}$ *and* $\{p_n\}_{n\geq 1}$ *such that* $\limsup_{n\to\infty} p_n < 1$ *and*

$$p_n K_n \left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) \sim \log n. \tag{29}$$

Then it holds for each $\ell = 0, 1, \ldots$ *that*

$$\lim_{n\to\infty} \mathbb{P}\left[(\kappa_v(n;\theta_n) = \ell) \ \cap \ (\delta(n;\theta_n) > \ell)\right] = 0. \tag{30}$$

In the special case with $\ell = 0$, Proposition 5.4 becomes analogous to [32, Proposition 11.1]. In order to see why (27) is implied by Proposition 5.4 under admissible scalings, just note that with $\gamma : \mathbb{N}_0 \to \mathbb{R}$ defined through (9), the admissibility condition (28) is equivalent to (29). Put differently, Proposition 5.4 establishes (27) under any admissible scaling $(K, p) : \mathbb{N}_0 \to \mathbb{N}_0 \times [0, 1)$, even those that do not satisfy $\lim_{n\to\infty} \gamma_n = +\infty$. This last condition is nevertheless required for the one-law $\lim_{n\to\infty} \mathbb{P}\left[\mathbb{I}(n;\theta_n) \text{ is } k\text{-connected }\right] = 1$ to hold as it is crucial in obtaining (26).

5.3 Useful consequences of the scaling condition (29)

We collect in this section some useful consequences of the scaling condition (9) that follow under the assumptions of admissibility and $\limsup_{n\to\infty} p_n < 1$.

Lemma 5.5 *Consider scalings* $\{K_n\}_{n\geq 1}$ *and* $\{p_n\}_{n\geq 1}$ *such that* $\limsup_{n\to\infty} p_n < 1$ *and (29) is satisfied. Namely, the sequence* $\gamma : \mathbb{N}_0 \to \mathbb{R}$ *defined through (9) satisfies* $\gamma_n = \pm o(\log n)$*. Then, we have*

$$p_n K_n = \Theta(\log n) \tag{31}$$

and

$$p_n \lambda_n(K_n) \leq (1 + o(1)) \frac{\log n}{n}. \tag{32}$$

Proof. It is easy to see that (31) follows directly from (29) under $\limsup_{n\to\infty} p_n < 1$. A complete proof was given by the authors in [34, Lemma 5.7].

In order to see (32), note that

$$1 - \frac{\log(1-p_n)}{p_n} \geq 2$$

so that

$$p_n \lambda_n(K) = \frac{p_n K_n}{n-1}\left(2 - \frac{K_n}{n-1}\right) \ \leq \ \frac{p_n K_n}{n-1}\left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right). \tag{33}$$

Reporting (29) into (33), we get (32) as we note that

$$\frac{p_n K_n}{n-1}\left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) = \frac{(1+o(1))\log n}{n-1} = (1+o(1))\frac{\log n}{n}.$$

■

Finally, we will also make use of the following consequence of (29). For any $\epsilon > 0$, it holds that

$$p_n K_n \left(1 - \frac{\log(1-p_n)}{p_n} - \frac{K_n}{n-1}\right) \geq (1-\epsilon)\log n, \text{ for all } n \text{ sufficiently large.} \tag{34}$$

6 A Proof of Proposition 5.4

The arguments given next are analogous to those presented in [32, Sec. XI] for the case $\ell = 0$. We start by introducing some notation. Let $\mathcal{N}$ denote the collection of all non-empty subsets of $\{1, \ldots, n\}$. For any $U \in \mathcal{N}$, we define the graph $\mathbb{I}(n; K, p)(U)$ (with vertex set U) as the subgraph of $\mathbb{I}(n; K, p)$ restricted to the nodes in U. If all nodes in U are deleted from $\mathbb{I}(n; K, p)$, the remaining graph is given by $\mathbb{I}(n; K, p)(U^c)$ on the vertices $U^c = \{1, \ldots, n\} \setminus U$. Let $\mathcal{N}_{U^c}$ denote the collection of all non-empty subsets of U^c. We say that $T \in \mathcal{N}_{U^c}$ is *isolated* in $\mathbb{I}(n; K, p)(U^c)$ if there are no edges (in $\mathbb{I}(n; K, p)$) between the nodes in T and the nodes in $U^c \setminus T$. This is characterized by the events $\{\overline{E_{ij}} : \ i \in T, \ j \in U^c \setminus T\}$, and defines the notion of $\mathbb{I}(n; K, p)$ being $(U, T)-$separable; i.e., deleting all nodes in U renders $\mathbb{I}(n; K, p)$ *disconnected* into a subgraph on T and a subgraph on U^c/T with no edges in between the two subgraphs.

The proof starts with the following observations:

1. If the vertex-connectivity of $\mathbb{I}(n; K, p)$ is ℓ (i.e., if $\kappa_v(n; \theta) = \ell$) and yet each node has degree at least $\ell+1$ (i.e., $\delta(n; \theta) > \ell$), then there must exist subsets U, T of nodes with $U \in \mathcal{N}$, $|U| = \ell$, and $T \in \mathcal{N}_{U^c}$, $|T| \geq 2$ such that $\mathbb{I}(n; K, p)(T)$ is connected while T is isolated in $\mathbb{I}(n; K, p)(U^c)$. This ensures that $\mathbb{I}(n; K, p)$ is $(U, T)-$separable, and hence can be disconnected by deleting an appropriately selected set (i.e., U) of ℓ nodes. Notice that, this would *not* be possible for sets T in $\mathcal{N}_{U^c}$ with $|T| = 1$, since the degree of a node in T is at least $\ell + 1$ by virtue of the event $\delta(n; \theta) > \ell$; this ensures that a single node in T is connected to at least one node in $U^c \setminus T$.

2. Secondly, the event $\kappa_v(n; \theta) = \ell$ enforces $\mathbb{I}(n; K, p)$ to remain connected after the deletion of any $\ell - 1$ nodes. Therefore, if there exists a subset U (with $|U| = \ell$) such that some T in $\mathcal{N}_{U^c}$ is isolated in $\mathbb{I}(n; K, p)(U^c)$, then each of the ℓ nodes in U should be connected to at least one node in T *and* to at least one node in $U^c \setminus T$. This can easily be seen by contradiction: Consider subsets $U \in \mathcal{N}$ with $|U| = \ell$, and $T \in \mathcal{N}_{U^c}$ with $|T| \geq 2$, such that there exists no edge between the nodes in T and the nodes in $U^c \setminus T$. Suppose there exists $i \in U$ that is *not* connected to any node in T. Then, $\mathbb{I}(n; K, p)$ can be disconnected by deleting the nodes in $U \setminus \{i\}$ since there will be no edge between the nodes in T and the nodes in $\{i\} \cup U^c \setminus T$. But, $|U \setminus \{i\}| = \ell - 1$, and this contradicts the fact that $\kappa_v(n; \theta) = \ell$. Similar arguments can be used to show that it is *not* possible for $i \in U$ to be disconnected from all nodes in $U^c \setminus T$.

We now use the above arguments to characterize the event $[(\kappa_v(n; \theta) = \ell) \cap (\delta(n; \theta) > \ell)]$. With each non-empty subset $T \subseteq U^c$ of nodes, we associate several events of interest: Let $\mathcal{C}_T$ denote the event that the subgraph $\mathbb{I}(n; K, p)(T)$ is itself connected. The event $\mathcal{C}_T$ is completely determined by the rvs $\{\Gamma_i(K), \ i \in T\}$ and $\{B_{ij}, \ i, j \in T\}$. We also introduce the event $\mathcal{D}_{U,T}$ to capture the fact that T is isolated in $\mathbb{I}(n; K, p)(U^c)$, i.e.,

$$\mathcal{D}_{U,T} \ := \bigcap_{\substack{i \in T \\ j \in U^c \setminus T}} \overline{E_{ij}}.$$

Finally, we let $\mathcal{B}_{U,T}$ denote the event that each node in U has an edge with at least one node in T, i.e.,

$$\mathcal{B}_{U,T} := \bigcap_{i \in U} \bigcup_{j \in T} E_{ij}. \tag{35}$$

We also set[3]

$$\mathcal{A}_{U,T} := \mathcal{B}_{U,T} \cap \mathcal{C}_T \cap \mathcal{D}_{U,T}.$$

The inclusion

$$[(\kappa_v(n;\theta) = \ell) \ \cap \ (\delta(n;\theta) > \ell)] \ \subseteq \bigcup_{U \in \mathcal{N}_{n,\ell},\, T \in \mathcal{N}_{U^c}:\ |T| \geq 2} \mathcal{A}_{U,T}$$

is now immediate with $\mathcal{N}_{n,r}$ denoting the collection of all subsets of $\{1,\ldots,n\}$ with exactly r elements. It is also easy to check that this union need only be taken over all subsets T of $\{1,\ldots,n\}$ with $2 \leq |T| \leq \lfloor \frac{n-\ell}{2} \rfloor$.

We now use a standard union bound argument to obtain

$$\begin{aligned}\mathbb{P}\left[(\kappa_v(n;\theta) = \ell) \ \cap \ (\delta(n;\theta) > \ell)\right] \ &\leq \sum_{U \in \mathcal{N}_{n,\ell}, T \in \mathcal{N}_{U^c}:\ 2 \leq |T| \leq \lfloor \frac{n-\ell}{2} \rfloor} \mathbb{P}\left[\mathcal{A}_{U,T}\right] \\ &= \sum_{r=2}^{\lfloor \frac{n-\ell}{2} \rfloor} \sum_{U \in \mathcal{N}_{n,\ell}, T \in \mathcal{N}_{U^c,r}} \mathbb{P}\left[\mathcal{A}_{U,T}\right]\end{aligned}$$

with $\mathcal{N}_{U^c,r}$ denoting the collection of all subsets of U^c with exactly r elements.

For each $r = 1,\ldots,n-\ell-1$, we simplify the notation by writing $\mathcal{A}_{\ell,r} := \mathcal{A}_{\{1,\ldots,\ell\},\{\ell+1,\ldots,\ell+r\}}$, $\mathcal{D}_{\ell,r} := \mathcal{D}_{\{1,\ldots,\ell\},\{\ell+1,\ldots,\ell+r\}}$, $\mathcal{B}_{\ell,r} := \mathcal{B}_{\{1,\ldots,\ell\},\{\ell+1,\ldots,\ell+r\}}$ and $\mathcal{C}_r := \mathcal{C}_{\{\ell+1,\ldots,\ell+r\}}$. Under the enforced assumptions on the system model, exchangeability yields

$$\mathbb{P}\left[\mathcal{A}_{U,T}\right] = \mathbb{P}\left[\mathcal{A}_{\ell,r}\right], \quad U \in \mathcal{N}_{n,\ell}, \ T \in \mathcal{N}_{U^c,r}$$

and the expression

$$\sum_{U \in \mathcal{N}_{n,\ell}, T \in \mathcal{N}_{U^c,r}} \mathbb{P}\left[\mathcal{A}_{U,T}\right] = \binom{n}{\ell}\binom{n-\ell}{r}\, \mathbb{P}\left[\mathcal{A}_{\ell,r}\right]$$

follows since $|\mathcal{N}_{n,\ell}| = \binom{n}{\ell}$ and $|\mathcal{N}_{U^c,r}| = \binom{n-\ell}{r}$. Substituting into (36) we obtain the key bound

$$\mathbb{P}\left[(\kappa_v(n;\theta) = \ell) \ \cap \ (\delta(n;\theta) > \ell)\right] \leq \sum_{r=2}^{\lfloor \frac{n-\ell}{2} \rfloor} \binom{n}{\ell}\binom{n-\ell}{r}\, \mathbb{P}\left[\mathcal{A}_{\ell,r}\right]. \tag{36}$$

The proof of Proposition 5.4 will be completed once we show the following result.

Lemma 6.1 *Consider scalings $\{K_n\}_{n\geq 1}$ and $\{p_n\}_{n\geq 1}$ such that $\limsup_{n\to\infty} p_n < 1$ and (29) holds. Then, for each $\ell = 0, 1, \ldots$, we have*

$$\lim_{n\to\infty} \sum_{r=2}^{\lfloor \frac{n-\ell}{2} \rfloor} \binom{n}{\ell}\binom{n-\ell}{r}\, \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0. \tag{37}$$

[3]For brevity of notation, the dependence of the events $\mathcal{A}_{U,T}, \mathcal{B}_{U,T}, \mathcal{C}_T$ and $\mathcal{D}_{U,T}$ on the parameters n, K, p are suppressed.

In view of (36), Lemma 6.1 implies (30) and hence the proof of Proposition 5.4 is completed. As explained before, Proposition 5.4 establishes (27), which in turn leads to the proof of Theorem 3.1 by means of the arguments given in Section 5.1. ■

We will establish (37) by first obtaining sufficiently tight upper bounds for the term $\mathbb{P}[\mathcal{A}_{\ell,r}]$. The means to do so are provided in the following section.

7 Bounding the probabilities $\mathbb{P}[\mathcal{A}_{\ell,r}]$

7.1 Negative Association of link assignment random variables

Fix positive integers $n = 2, 3, \ldots$ and K with $K < n$. Recall that $[i \sim j]$ denotes the event that nodes i and j are adjacent in $\mathbb{I}(n; K, p)$. Several properties of the $\{0,1\}$-valued rvs

$$\{\mathbf{1}[i \sim j], \quad i \neq j, \quad i, j = 1, \ldots, n\} \tag{38}$$

will play a key role in some of the forthcoming arguments (which are mainly adopted from [32, Sec. IX]).

The properties of interest can be couched in terms of *negative association*, a form of negative correlation introduced by Joag-Dev and Proschan [15]. We first develop the needed definitions and properties: Let $\{X_\lambda,\ \lambda \in \Lambda\}$ be a collection of $\mathbb{R}$-valued rvs indexed by the finite set Λ. For any non-empty subset A of Λ, we write X_A to denote the $\mathbb{R}^{|A|}$-valued $X_A = (X_\lambda,\ \lambda \in A)$. The rvs $\{X_\lambda,\ \lambda \in \Lambda\}$ are then said to be *negatively associated* if for any non-overlapping subsets A and B of Λ and for any monotone increasing mappings $\varphi : \mathbb{R}^{|A|} \to \mathbb{R}$ and $\psi : \mathbb{R}^{|B|} \to \mathbb{R}$, the covariance inequality

$$\mathbb{E}[\varphi(X_A)\psi(X_B)] \leq \mathbb{E}[\varphi(X_A)]\,\mathbb{E}[\psi(X_B)] \tag{39}$$

holds whenever the expectations in (39) are well defined and finite.

This definition has some useful consequences to be utilized repeatedly: First of all, it is also well known [15, P2, p. 288] that the negative association of the rvs $\{X_\lambda,\ \lambda \in \Lambda\}$ implies the inequality

$$\mathbb{E}\left[\prod_{\lambda \in A} f_\lambda(X_\lambda)\right] \leq \prod_{\lambda \in A} \mathbb{E}[f_\lambda(X_\lambda)] \tag{40}$$

where A is a subset of Λ and the collection $\{f_\lambda,\ \lambda \in A\}$ of mappings $\mathbb{R} \to \mathbb{R}_+$ are all monotone increasing; by non-negativity all the expectations exist and finiteness is moot. Also, the negative association of $\{X_\lambda,\ \lambda \in \Lambda\}$ implies the negative association of the collection $\{X_\lambda,\ \lambda \in \Lambda'\}$ where Λ' is any subset of Λ.

We can apply these ideas to collections of indicator rvs, namely for each λ in Λ, $X_\lambda = \mathbf{1}[E_\lambda]$ for some event E_λ. From the definitions, it is easy to see that if the rvs $\{\mathbf{1}[E_\lambda],\ \lambda \in \Lambda\}$ are negatively associated, so are the rvs $\{\mathbf{1}[E_\lambda^c],\ \lambda \in \Lambda\}$. Moreover, for any subset A of Λ, we have

$$\mathbb{P}[E_\lambda,\ \lambda \in A] \leq \prod_{\lambda \in A} \mathbb{P}[E_\lambda]. \tag{41}$$

This follows from (40) by taking $f_\lambda(x) = x^+$ on $\mathbb{R}$ for each λ in Λ.

A key observation for our purpose is as follows: For each $i = 1, \ldots, n$, the rvs

$$\{\mathbf{1}\left[j \in \Gamma_i(K)\right],\ j \in \mathcal{N}_{-i}\} \tag{42}$$

form a collection of negatively associated rvs. This is a consequence of the fact that the set $\Gamma_i(K)$ represents a random sample (without replacement) of size K from $\mathcal{N}_{-i}$; see [15, Example 3.2(c)] for details. Using this fact, it is not difficult to show that the rvs

$$\left\{\mathbf{1}\left[j \in \Gamma_i(K)\ \vee\ i \in \Gamma_j(K)\right], \quad \begin{matrix} i \neq j \\ i,j = 1,\ldots,n \end{matrix}\right\}, \tag{43}$$

or equivalently the rvs

$$\{\mathbf{1}\left[i \sim_K j\right], \quad i \neq j, \quad i,j = 1,\ldots,n\} \tag{44}$$

are also negatively associated. This is already established by Yağan and Makowski in [32, Sec. IX] and indicates that link assignment events in $\mathbb{H}(n;K)$ are negatively associated.

We now extend this argument to show that link assignment events in the intersection graph $\mathbb{H} \cap \mathbb{G}(n;K,p)$ are also negatively associated. We start by noting that the collection of rvs

$$\{\mathbf{1}\left[i \sim_B j\right], \quad i \neq j, \quad i,j = 1,\ldots,n\}$$

are mutually independent, and hence they are negatively associated since (39) holds with equality. It is clear that the collection of rvs $\{\mathbf{1}\left[i \sim_B j\right],\ \ i \neq j,\ \ i,j = 1,\ldots,n\}$ and $\{\mathbf{1}\left[i \sim_K j\right],\ \ i \neq j,\ \ i,j = 1,\ldots,n\}$ are mutually independent. Thus, by the "closure under products" property of negative association [15, P7, p. 288] [9, p. 35], the rvs

$$\{\mathbf{1}\left[i \sim_B j\right], \mathbf{1}\left[i \sim_K j\right], \quad i \neq j, \quad i,j = 1,\ldots,n\}$$

also form a collection of negatively associated rvs. With distinct $i,j = 1,\ldots,n$, we note that

$$\mathbf{1}\left[i \sim j\right] := \mathbf{1}\left[(i \sim_B j)\ \cap\ (i \sim_K j)\right] = g\left(\mathbf{1}\left[i \sim_B j\right], \mathbf{1}\left[i \sim_K j\right]\right)$$

with mapping $g : \mathbb{R}^2 \to \mathbb{R}$ given by $g(x,y) = x^+y^+$ for all x,y in $\mathbb{R}$. This mapping being non-decreasing on $\mathbb{R}^2$, it follows [15, P6, p. 288] that the rvs

$$\left\{\mathbf{1}\left[i \sim j\right], \quad \begin{matrix} i \neq j \\ i,j = 1,\ldots,n \end{matrix}\right\} \tag{45}$$

are negatively associated. This establishes the fact that edge assignment events are negatively associated in $\mathbb{H} \cap \mathbb{G}(n;K,p)$.

7.2 Bounding the probabilities $\mathbb{P}\left[\mathcal{B}_{\ell,r}\right]$

Recall from (35) that $\mathcal{B}_{\ell,r}$ denotes the event that each of the nodes in $\{1,\ldots,\ell\}$ has at least one edge with the nodes in $\{\ell+1,\ldots,\ell+r\}$; i.e., $\mathcal{B}_{\ell,r} := \bigcap_{i=1}^{\ell} \bigcup_{j=\ell+1}^{\ell+r} E_{ij}$. Also, recall from (6) that $\mathbb{P}\left[E_{ij}\right] = p\lambda_n(K)$. The following bound is a direct consequence of the negative association of link assignment events $\{E_{ij}\}$ (and a standard application of the union bound) and will prove useful in various occasions.

Lemma 7.1 *For each $r = 2, \ldots, n$, we have*

$$\mathbb{P}\left[\mathcal{B}_{\ell,r}\right] \leq (rp\lambda_n(K))^{\ell}. \tag{46}$$

Proof. We start by recalling the definition $\mathcal{B}_{\ell,r} := \bigcap_{i=1}^{\ell} \bigcup_{j=\ell+1}^{\ell+r} E_{ij}$, and noting that the collection of rvs $\{\mathbf{1}\left[E_{ij}\right], \ \ i \neq j, \ \ i,j = 1, \ldots, n\}$ are negatively associated as shown in Section 7.1. Next, we argue that the collection of rvs

$$\left\{\mathbf{1}\left[\cup_{j=\ell+1}^{\ell+r} E_{ij}\right], \quad i = 1, \ldots, \ell\right\}$$

are also negatively associated. This follows from the "disjoint monotone aggregation" property of negative association [9, p. 35] upon noting that the collections $\{\mathbf{1}\left[E_{1j}\right], \ \ j = \ell+1, \ldots, \ell+r\}$, $\{\mathbf{1}\left[E_{2j}\right], \ \ j = \ell+1, \ldots, \ell+r\}, \ldots, \{\mathbf{1}\left[E_{\ell j}\right], \ \ j = \ell+1, \ldots, \ell+r\}$ are mutually disjoint (i.e., that they have no common terms), and the mapping

$$f_i\left(\{\mathbf{1}\left[E_{ij}\right], \ \ j = \ell+1, \ldots, \ell+r\}\right) = \mathbf{1}\left[\cup_{j=\ell+1}^{\ell+r} E_{ij}\right] = 1 - \prod_{j=\ell+1}^{\ell+r} \left(1 - \mathbf{1}\left[E_{ij}\right]\right)$$

is non-decreasing for each $i = 1, 2, \ldots, \ell$. In view of (40), this leads to

$$\mathbb{P}\left[\mathcal{B}_{\ell,r}\right] = \mathbb{E}\left[\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\right] = \mathbb{E}\left[\prod_{i=1}^{\ell} \mathbf{1}\left[\cup_{j=\ell+1}^{\ell+r} E_{ij}\right]\right] \leq \prod_{i=1}^{\ell} \mathbb{E}\left[\mathbf{1}\left[\cup_{j=\ell+1}^{\ell+r} E_{ij}\right]\right] \leq \prod_{i=1}^{\ell} r\mathbb{P}\left[E_{ij}\right] = (rp\lambda_n(K))^{\ell},$$

where we also used a standard union bound. This establishes Lemma 7.1. ■

7.3 Bounding the probabilities $\mathbb{P}\left[\mathcal{D}_{\ell,r}\right]$

The following two results will be used to efficiently bound the probability $\mathbb{P}\left[D_{\ell,r}\right]$, i.e., the event that vertices $\{\ell+1, \ldots, \ell+r\}$ are isolated from the vertices $\{\ell+r+1, \ldots, n\}$ in $\mathbb{I}(n;\theta)$. Their proofs, based on the negative association of the rvs in (42), are very similar to the proofs of [32, Lemma 12.1] and [32, Lemma 12.2], respectively, with only small modifications needed. We skip these details here for brevity, but comment on the intuition behind them.

The first bound leverages the fact that $D_{\ell,r}$ takes place *only if* all the edges in K-out graph $\mathbb{H}(n;K)$ *from* $\{\ell+r+1, \ldots, n\}$ *to* $\{\ell+1, \ldots, \ell+r\}$ are *deleted* during the intersection with the ER graph $\mathbb{G}(n;p)$; more precisely, all $\mathcal{E}^{\star}_{n,r}(K)$ edges (see (48) below) between the two sets of vertices that result from the selections $\{\Gamma_{\ell+r+1}, \ldots, \Gamma_n\}$ of the first set shall be *off* in the ER graph.

Lemma 7.2 *For each $r = 2, \ldots, n-1$, the inequality*

$$\mathbb{P}\left[\mathcal{D}_{\ell,r} \ \middle| \ \begin{array}{c} B_{ij}(p), \ 1 \leq i < j \leq r+\ell \\ \Gamma_1(K), \ldots, \Gamma_n(K) \end{array}\right] \leq (1-p)^{\mathcal{E}^{\star}_{n,r}(K)} \tag{47}$$

holds where the rv $\mathcal{E}^{\star}_{n,r}(K)$ is given by

$$\mathcal{E}^{\star}_{n,r}(K) = \sum_{i=\ell+1}^{\ell+r} \sum_{j=r+\ell+1}^{n} \mathbf{1}\left[i \in \Gamma_j(K)\right]. \tag{48}$$

The bound (47) is fairly intuitive and could be inferred directly by inspection. To see this, recall that

$$\mathcal{D}_{\ell,r} = \bigcap_{i=\ell+1}^{\ell+r} \bigcap_{j=\ell+r+1}^{n} \overline{E_{ij}}.$$

Namely, $\mathcal{D}_{\ell,r}$ stands for the event that vertices $\{\ell+1,\ldots,\ell+r\}$ are isolated from the vertices $\{\ell+r+1,\ldots,n\}$ in $\mathbb{I}(n;\theta)$. But, the individual K-out graph $\mathbb{H}(n;K)$ contains at least

$$\sum_{i=\ell+1}^{\ell+r} \sum_{j=r+\ell+1}^{n} \mathbf{1}\left[i \in \Gamma_j(K)\right]$$

many distinct edges between these two sets of vertices (counting only the edges resulting from the selections $\{\Gamma_j\}_{j=\ell+r+1}^{n}$ of the second set). For the event $\mathcal{D}_{\ell,r}$ to take place, all these edges should disappear when intersecting with the ER graph $\mathbb{G}(n;p)$. We then get (47) by independence.

The arguments leading to (47) can be modified to obtain the next bound, this time with the help of negative association.

Lemma 7.3 *For each $r = 2,\ldots,n-1$, the inequality*

$$\mathbb{P}\left[\mathcal{D}_{\ell,r} \;\middle|\; \begin{array}{c} B_{ij}(p),\ 1 \le i < j \le r+\ell \\ \Gamma_1(K),\ldots,\Gamma_{\ell+r}(K) \end{array}\right] \le (1-p)^{\mathcal{E}^{\star\star}_{n,r}(K)} \cdot \left(1 - \frac{pK}{n-1}\right)^{r(n-r-\ell)-\mathcal{E}^{\star\star}_{n,r}(K)} \tag{49}$$

holds with $\mathcal{E}^{\star\star}_{n,r}(K)$ given by

$$\mathcal{E}^{\star\star}_{n,r}(K) = \sum_{i=\ell+1}^{\ell+r} \sum_{j=r+\ell+1}^{n} \mathbf{1}\left[j \in \Gamma_i(K)\right]. \tag{50}$$

In order to understand the intuition behind the bound (49), note again that for $\mathcal{D}_{\ell,r}$ to take place, there should be no edge between the vertices $\{\ell+1,\ldots,\ell+r\}$ and $\{\ell+r+1,\ldots,n\}$ in $\mathbb{I}(n;\theta)$. For this to happen, we need

i) All edges that are assigned in $\mathbb{H}(n;K)$ due to the selections $\{\Gamma_{\ell+1},\ldots,\Gamma_{\ell+r}\}$ of vertices $\{\ell+1,\ldots,\ell+r\}$ shall be deleted upon intersection with the ER graph $\mathbb{G}(n;p)$. By independence, this gives the first term of the bound given at (49).

ii) Once the edges in item (i) above are taken out of consideration, there remains $r(n-r-\ell) - \mathcal{E}^{\star\star}_{n,r}(K)$ possible edges between the two aforementioned sets of vertices, each of which does *not* exist in $\mathbb{I}(n;\theta)$ with probability

$$\mathbb{P}\left[(i \in \Gamma_j(K)\ \cap\ B_{ij}(p) = 1)^c\right] = 1 - \frac{Kp}{n-1}.$$

The second term in (49) follows from the fact that indicators of these $r(n-r-\ell) - \mathcal{E}^{\star\star}_{n,r}(K)$ events (each indicating the *non*-existence of a particular edge) are negatively associated.

The following bounds follow from the facts that $|\Gamma_i(K)| = K$ and $\sum_{j=1}^{n} \mathbf{1}\left[j \in \Gamma_i(K)\right] = K$, and will be used repeatedly.

$$\begin{aligned} \mathcal{E}^{\star\star}_{n,r}(K) &\geq r(K-r-\ell)^+ & (51)\\ r(n-r-\ell) - \mathcal{E}^{\star\star}_{n,r}(K) &\geq r(n-r-\ell-K)^+ & (52)\end{aligned}$$

The tail of the rv $\mathcal{E}^{\star}_{n,r}(K)$ is controlled through the next result, obtained from the Chernoff-Hoeffding bound applied to the negatively associated rvs [9, Thm 1.1., p.6].

Lemma 7.4 *Fix $\ell = 0, 1, \ldots, n-3$ and $r = 2, \ldots, n-\ell-1$. For any t in $(0,1)$ we have*

$$\mathbb{P}\left[\mathcal{E}^{\star}_{n,r}(K) \leq (1-t)rK \cdot \frac{n-r-\ell}{n-1}\right] \leq e^{-\frac{t^2}{2} rK \cdot \frac{n-r-\ell}{n-1}}. \tag{53}$$

Proof. It is easy to see that $\mathbb{E}\left[\mathcal{E}^{\star}_{n,r}(K)\right] = rK\frac{n-r-\ell}{n-1}$. The proof then follows from the Chernoff-Hoeffding bound, and is identical to the proof of [32, Lemma 12.3].

7.4 Bounding the probabilities $\mathbb{P}\left[\mathcal{C}_r\right]$

A key consequence of the negative association of the link assignment rvs in $\mathbb{I}(n;\theta)$ is the following bound established by Yağan and Makowski [32, Lemma 12.5].

Lemma 7.5 *[32, Lemma 12.5] For each $r = 2, \ldots, n$, we have*

$$\mathbb{P}\left[\mathcal{C}_r\right] \leq r^{r-2}(p\lambda_n(K))^{r-1}. \tag{54}$$

The proof of this bound passes through the facts that i) if a set of r vertices are to form a *connected component*, then they would have to contain – as a subgraph – at least one of the r^{r-2} possible *spanning trees* that can exist on these vertices; and ii) for any spanning tree, the $r-1$ edges it contains form a set of negatively associated edge occurrence events, meaning that probability of its existence on this particular set of vertices is less than the $r-1^{\text{th}}$ power of a single edge probability.

7.5 Establishing the negative association of $\mathbf{1}\left[\mathcal{C}_r\right]$ and $\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]$

A key ingredient in bounding the terms $\mathbb{P}\left[\mathcal{A}_{\ell,r}\right]$ and hence in our proof is the following result which states that rvs $\mathbf{1}\left[\mathcal{C}_r\right]$ and $\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]$ are negatively associated.

Lemma 7.6 *For each $\ell = 0, 1, \ldots, n-3$ and $r = 2, \ldots, n-\ell-1$, we have*

$$\mathbb{P}\left[\mathcal{B}_{\ell,r} \cap \mathcal{C}_r\right] \leq \mathbb{P}\left[\mathcal{B}_{\ell,r}\right]\mathbb{P}\left[\mathcal{C}_r\right]. \tag{55}$$

Proof. Recall that $\mathcal{C}_r$ stands for the event that nodes $\ell+1, \ldots, \ell+r$ form a connected subgraph in $\mathbb{I}(n;K,p)$. This is equivalent to stating that $\mathbb{I}(n;K,p)$ contains a spanning tree on nodes $\{\ell+1, \ldots, \ell+r\}$. Let $\mathcal{T}_{\ell+1,\ell+r}$ denote the collection of all spanning trees on the vertex set $\{\ell+1, \ldots, \ell+r\}$ – By Cayley's formula [16], we know that there are r^{r-2} trees on r vertices, i.e.,

$|\mathcal{T}_{\ell+1,\ell+r}| = r^{r-2}$. Put differently, $\mathcal{T}_{\ell+1,\ell+r}$ contains r^{r-2} distinct trees T, where each $T \in \mathcal{T}_{\ell+1,\ell+r}$ is a collection of $r-1$ node pairs (j_1, j_2) (with $j_1, j_2 \in \{\ell+1, \ldots, \ell+r\}$) that stand for the $r-1$ edges of the tree. Hence, we have

$$\mathbf{1}\left[\mathcal{C}_r\right] = \mathbf{1}\left[\bigcup_{T\in\mathcal{T}_{\ell+1,\ell+r}} \bigcap_{(j_1,j_2)\in T} E_{j_1 j_2}\right] = 1 - \prod_{T\in\mathcal{T}_{\ell+1,\ell+r}} \left(1 - \prod_{(j_1,j_2)\in T} \mathbf{1}\left[E_{j_1 j_2}\right]\right) \tag{56}$$

while, as before, $\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]$ is given by

$$\mathbf{1}\left[\mathcal{B}_{\ell,r}\right] = \mathbf{1}\left[\bigcap_{i=1}^{\ell} \bigcup_{j=\ell+1}^{\ell+r} E_{ij}\right] = \prod_{i=1}^{\ell} \left(1 - \prod_{j=\ell+1}^{\ell+r} \left(1 - \mathbf{1}\left[E_{ij}\right]\right)\right). \tag{57}$$

It is already established in Section 7.1 that the collection of rvs $\{E_{ij}, \quad i \neq j, \quad i, j = 1, \ldots, n\}$ are negatively associated. Since $j_1, j_2 \in \{\ell+1, \ldots, \ell+r\}$, it is easy to see that the $\mathbf{1}\left[E_{j_1 j_2}\right]$ terms that appear in (56) are disjoint from the $\mathbf{1}\left[E_{ij}\right]$ terms that appear in (57). It is also clear that both relations (56) and (57) constitute non-decreasing mappings of the rvs $\mathbf{1}\left[E_{j_1 j_2}\right]$ and $\mathbf{1}\left[E_{ij}\right]$, respectively. Thus, from "disjoint monotone aggregation" property of negative association [9, p. 35] we obtain that rvs $\mathbf{1}\left[\mathcal{C}_r\right]$ and $\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]$ are negatively associated and (55) follows immediately. ■

7.6 Bounds on $\mathbb{P}\left[\mathcal{A}_{\ell,r}\right]$

We now combine the findings of Sections 7.1 - 7.5 to obtain a bound on the probabilities $\mathbb{P}\left[\mathcal{A}_{\ell,r}\right]$. The corresponding result is presented next.

Lemma 7.7 *Fix $\ell = 0, 1, \ldots, n-3$ and $r = 2, \ldots, n-\ell-1$. For some $a > 0$ we have*

$$\begin{aligned} \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \quad \leq \quad & \min\left\{(rp\lambda_n(K))^{\ell}, 1\right\} \cdot r^{r-2}(p\lambda_n(K))^{r-1} \\ & \times \min\left\{(1-p)^{r(K-r-\ell)}\left(1 - \frac{pK}{n-1}\right)^{r(n-r-\ell-K)}, 2e^{-arpK\frac{n-r-\ell}{n-1}}\right\}. \end{aligned} \tag{58}$$

We will use Lemma 7.7 in Section 8 to establish (37). In doing so, the first parts of the *min* expressions appearing in the bound (58) will be used for *small* values of r (see Section 8.1), while the second parts will be used for *large* values of r (see Section 8.2), in the range $r = 2, \ldots, \lfloor \frac{n-\ell}{2} \rfloor$.

Proof. The arguments needed for the proof are similar to those used in [32, Sec. XIV]. We start by noting that the event $\mathcal{C}_r$ is determined solely by the rvs $\{\Gamma_{\ell+1}(K), \ldots, \Gamma_{\ell+r}(K)\}$ and $\{B_{ij}(p), \ \ell+1 \leq i < j \leq \ell+r\}$. Similarly, $\mathcal{B}_{\ell,r}$ depends only on the rvs $\{\Gamma_1(K), \ldots, \Gamma_{\ell+r}(K)\}$ and $\{B_{ij}(p), \quad i = 1, \ldots, \ell, \quad j = \ell+1, \ldots, \ell+r\}$. We will establish (58) by combining two bounds that are obtained next via different arguments. First, conditioning on the rvs $\{B_{ij}(p), \ 1 \leq i < j \leq r+\ell\}$ and $\{\Gamma_1(K), \ldots, \Gamma_{\ell+r}(K)\}$ we get

$$\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \quad = \quad \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbf{1}\left[\mathcal{D}_{\ell,r}\right] \;\middle|\; \begin{array}{c} B_{ij}(p), \ 1 \leq i < j \leq r+\ell \\ \Gamma_1(K), \ldots, \Gamma_{\ell+r}(K) \end{array}\right]\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbb{E}\left[\mathbf{1}\left[\mathcal{D}_{\ell,r}\right] \;\middle|\; \begin{matrix} B_{ij}(p),\ 1\le i<j\le r+\ell \\ \Gamma_1(K),\dots,\Gamma_{\ell+r}(K)\end{matrix}\right]\right] \\
&= \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbb{P}\left[\mathcal{D}_{\ell,r} \;\middle|\; \begin{matrix} B_{ij}(p),\ 1\le i<j\le r+\ell \\ \Gamma_1(K),\dots,\Gamma_{\ell+r}(K)\end{matrix}\right]\right] \\
&\le \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\cdot(1-p)^{\mathcal{E}^{\star\star}_{n,r}(K)}\left(1-\frac{pK}{n-1}\right)^{r(n-r-\ell)-\mathcal{E}^{\star\star}_{n,r}(K)}\right] \qquad (59)\\
&\le \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\cdot(1-p)^{r(K-r-\ell)}\left(1-\frac{pK}{n-1}\right)^{r(n-r-\ell-K)}\right] \qquad (60)\\
&= \mathbb{P}\left[\mathcal{C}_r\cap\mathcal{B}_{\ell,r}\right]\cdot(1-p)^{r(K-r-\ell)}\left(1-\frac{pK}{n-1}\right)^{r(n-r-\ell-K)} \qquad (61)\\
&\le \mathbb{P}\left[\mathcal{C}_r\right]\mathbb{P}\left[\mathcal{B}_{\ell,r}\right]\cdot(1-p)^{r(K-r-\ell)}\left(1-\frac{pK}{n-1}\right)^{r(n-r-\ell-K)} \qquad (62)
\end{aligned}
$$

where (59) follows from Lemma 7.3, (60) follows from the bounds (51)-(52), and (62) follows directly from Lemma 7.6.

Second, we condition on the rvs $\{B_{ij}(p),\ 1\le i<j\le r+\ell\}$ and $\{\Gamma_1(K),\dots,\Gamma_n(K)\}$ to get

$$
\begin{aligned}
\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbf{1}\left[\mathcal{D}_{\ell,r}\right] \;\middle|\; \begin{matrix} B_{ij}(p),\ 1\le i<j\le r+\ell \\ \Gamma_1(K),\dots,\Gamma_n(K)\end{matrix}\right]\right] \\
&= \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbb{E}\left[\mathbf{1}\left[\mathcal{D}_{\ell,r}\right] \;\middle|\; \begin{matrix} B_{ij}(p),\ 1\le i<j\le r+\ell \\ \Gamma_1(K),\dots,\Gamma_n(K)\end{matrix}\right]\right] \\
&= \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]\mathbb{P}\left[\mathcal{D}_{\ell,r} \;\middle|\; \begin{matrix} B_{ij}(p),\ 1\le i<j\le r+\ell \\ \Gamma_1(K),\dots,\Gamma_n(K)\end{matrix}\right]\right] \\
&\le \mathbb{E}\left[\mathbf{1}\left[\mathcal{C}_r\right]\mathbf{1}\left[\mathcal{B}_{\ell,r}\right](1-p)^{\mathcal{E}^{\star}_{n,r}(K)}\right] \qquad (63)\\
&= \mathbb{P}\left[\mathcal{C}_r\cap\mathcal{B}_{\ell,r}\right]\mathbb{E}\left[(1-p)^{\mathcal{E}^{\star}_{n,r}(K)}\right] \qquad (64)\\
&\le \mathbb{P}\left[\mathcal{C}_r\right]\mathbb{P}\left[\mathcal{B}_{\ell,r}\right]\mathbb{E}\left[(1-p)^{\mathcal{E}^{\star}_{n,r}(K)}\right] \qquad (65)
\end{aligned}
$$

where (63) follows from Lemma 7.2 and (64) follows from the fact that the rv $\mathcal{E}^{\star}_{n,r}(K)$ depends only on $\{\Gamma_{r+\ell+1}(K),\dots,\Gamma_n(K)\}$ and thus is independent from the rvs $\mathbf{1}\left[\mathcal{C}_r\right]$ and $\mathbf{1}\left[\mathcal{B}_{\ell,r}\right]$. Finally, (65) follows directly from Lemma 7.6.

Combining (62) and (65), we find

$$
\mathbb{P}\left[\mathcal{A}_{\ell,r}\right]\le\mathbb{P}\left[\mathcal{C}_r\right]\mathbb{P}\left[\mathcal{B}_{\ell,r}\right]\min\left\{(1-p)^{r(K-r-\ell)}\left(1-\frac{pK}{n-1}\right)^{r(n-r-\ell-K)},\mathbb{E}\left[(1-p)^{\mathcal{E}^{\star}_{n,r}(K)}\right]\right\}.
$$

Reporting (46) and (54) into this last bound, we see that the desired result (58) will follow if we show that

$$
\mathbb{E}\left[(1-p)^{\mathcal{E}^{\star}_{n,r}(K)}\right]\le 2e^{-arpK\frac{n-r-\ell}{n-1}} \qquad (66)
$$

for some $a>0$.

We will establish (66) by means of Lemma 7.4. Pick t arbitrary in $(0,1)$. A simple decomposition argument shows that

$$\begin{aligned}
&\mathbb{E}\left[(1-p)^{\mathcal{E}^\star_{n,r}(K)}\right] \\
&\quad\leq\ \mathbb{E}\left[(1-p)^{\mathcal{E}^\star_{n,r}(K)}\,\mathbf{1}\left[\mathcal{E}^\star_{n,r}(K) > (1-t)rK\cdot\frac{n-r-\ell}{n-1}\right]\right] + \mathbb{P}\left[\mathcal{E}^\star_{n,r}(K) \leq (1-t)rK\cdot\frac{n-r-\ell}{n-1}\right] \\
&\quad\leq\ (1-p)^{(1-t)rK\cdot\frac{n-r-\ell}{n-1}} + e^{-\frac{t^2}{2}rK\cdot\frac{n-r-\ell}{n-1}} \\
&\quad\leq\ e^{-(1-t)rpK\cdot\frac{n-r-\ell}{n-1}} + e^{-\frac{t^2}{2}rK\cdot\frac{n-r-\ell}{n-1}} \\
&\quad\leq\ e^{-(1-t)rpK\cdot\frac{n-r-\ell}{n-1}} + e^{-\frac{t^2}{2}rpK\cdot\frac{n-r-\ell}{n-1}} \\
&\quad\leq\ 2e^{-arpK\cdot\frac{n-r-\ell}{n-1}}
\end{aligned}$$

where we set $a = \min\left\{1-t, \frac{t^2}{2}\right\} > 0$. This gives (66) and Lemma 7.7 is now established. ∎

8 Establishing (37)

The rest of the proof is based on arguments similar to those used in [32, Sec. XIV and XV]: Consider a scaling $\theta : \mathbb{N}_0 \to \mathbb{N}_0 \times [0,1]$ as in the statement of Proposition 5.4. For the time being, pick an integer $R \geq 2$, and consider the decomposition

$$\sum_{r=2}^{\lfloor\frac{n-\ell}{2}\rfloor}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = \sum_{r=2}^{R}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] + \sum_{r=R+1}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right].$$

Let n go to infinity: The desired convergence (37) will be established if we show that

$$\lim_{n\to\infty}\sum_{r=2}^{R}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0, \tag{67}$$

and

$$\lim_{n\to\infty}\sum_{r=R+1}^{\lfloor\frac{n-\ell}{2}\rfloor}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0. \tag{68}$$

The next sections are devoted to proving the validity of (67) and (68). Throughout, we use the standard bound $\binom{n}{r} \leq \left(\frac{en}{r}\right)^r$ which is valid for all $r, n = 1, 2, \ldots$ with $r \leq n$. In particular, we will repeatedly use the facts that

$$\binom{n}{\ell}\binom{n-\ell}{r} \leq n^\ell \cdot n^r = n^{\ell+r}, \tag{69}$$

and

$$\binom{n}{\ell}\binom{n-\ell}{r} \leq n^\ell \cdot \binom{n}{r} \leq n^\ell\left(\frac{en}{r}\right)^r = n^{\ell+r}e^r r^{-r}. \tag{70}$$

8.1 Establishing (67)

Fix $r = 2, 3, \ldots$. From (29), (33), (69), and Lemma 7.7, it follows that

$$\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \tag{71}$$

$$\leq\; n^{\ell+r}\cdot r^{\ell+r-2}\left(p_n\lambda_n(K_n)\right)^{\ell+r-1}(1-p_n)^{r(K_n-r-\ell)}\left(1-\frac{p_nK_n}{n-1}\right)^{r(n-r-\ell-K_n)} \tag{72}$$

$$\leq\; n^{\ell+r}\cdot r^{\ell+r-2}\cdot\left\{\frac{\log n}{n}\cdot[1+o(1)]\right\}^{\ell+r-1}(1-p_n)^{r(K_n-r-\ell)}\left(1-\frac{p_nK_n}{n-1}\right)^{r(n-r-\ell-K_n)}, \tag{73}$$

where we have

$$\begin{aligned}
&(1-p_n)^{r(K_n-r-\ell)}\left(1-\frac{p_nK_n}{n-1}\right)^{r(n-r-\ell-K_n)}\\
&\leq\exp\left\{r(K_n-r-\ell)\log(1-p_n)-\frac{p_nK_n}{n-1}\cdot r(n-r-\ell-K_n)\right\}\\
&=\exp\left\{rK_n\log(1-p_n)+r(-r-\ell)\log(1-p_n)-rp_nK_n+\frac{rp_n{K_n}^2}{n-1}+\frac{p_nK_n}{n-1}r(\ell+r-1)\right\}\\
&=\exp\left\{-rp_nK_n\left(1-\frac{\log(1-p_n)}{p_n}-\frac{K_n}{n-1}\right)+r(-r-\ell)\log(1-p_n)+\frac{p_nK_n}{n-1}(\ell+r-1)\right\}. \qquad (74)
\end{aligned}$$

Under bounded r, ℓ and $\limsup_{n\to\infty} p_n < 1$, it holds that $e^{r(-r-\ell)\log(1-p_n)} = O(1)$. Under bounded r, ℓ and (31), it follows that

$$\lim_{n\to\infty} e^{\frac{p_nK_n}{n-1}(\ell+r-1)} = 1.$$

Applying the above and (34) to (74), we obtain

$$(1-p_n)^{r(K_n-r-\ell)}\left(1-\frac{p_nK_n}{n-1}\right)^{r(n-r-\ell-K_n)} \leq O(1)\cdot e^{-r(1-\epsilon)\log n} = O(1)\cdot n^{-r(1-\epsilon)}. \tag{75}$$

for arbitrary $\epsilon > 0$. Substituting (75) into (73) and using the fact that $r \geq 2$, we derive

$$\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \leq O(1)\cdot n^{1-r(1-\epsilon)}\cdot(\log n)^{\ell+r-1}\cdot r^{\ell+r-2} \leq O(1)\cdot n^{2\epsilon-1}\cdot(\log n)^{\ell+r-1}.$$

Since ϵ was arbitrary, we can select ϵ with $0 < \epsilon < \frac{1}{2}$ so that $2\epsilon - 1 < 0$. This yields

$$\lim_{n\to\infty}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0$$

for each $r = 2, 3, \ldots$. Hence, for any bounded R, we immediately obtain the desired result

$$\lim_{n\to\infty}\sum_{r=2}^{R}\binom{n}{\ell}\binom{n-\ell}{r}\,\mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0.$$

■

8.2 Establishing (68)

Consider now the range $r = R+1, \ldots, \lfloor \frac{n-\ell}{2} \rfloor$, where we have

$$\frac{n-r-\ell}{n-1} \geq \frac{n-\ell}{2(n-1)}.$$

From (33), (70), and Lemma 7.7, it follows for all n sufficiently large that

$$\begin{aligned}
\binom{n}{\ell}\binom{n-\ell}{r} \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] &\leq n^{\ell+r} e^r r^{-r} \cdot r^{r-2} \cdot \left(p_n \lambda_n(K_n)\right)^{r-1} \cdot 2e^{-arp_n K_n \frac{n-r-\ell}{n-1}} \\
&\leq n^{\ell+r} e^r r^{-r} \cdot r^{r-2} \cdot \left\{\frac{\log n}{n} \cdot [1+o(1)]\right\}^{r-1} \cdot 2e^{-arp_n K_n \frac{n-r-\ell}{n-1}} \\
&\leq 2n^{\ell+1}(e^2 \log n))^r \cdot e^{-arp_n K_n \frac{n-\ell}{2(n-1)}} \\
&= 2n^{\ell+1}\left(\exp\left\{2 + \log\log n - \frac{ap_n K_n}{2}\left(1 - \frac{\ell-1}{n-1}\right)\right\}\right)^r
\end{aligned}$$

Since $p_n K_n = \Theta(\log n)$ from (31), it holds that

$$2 + \log\log n - \frac{ap_n K_n}{2}\left(1 - \frac{\ell-1}{n-1}\right) = -\Theta(\log n).$$

This implies that there exists a positive constant c such that for all n sufficiently large.

$$2 + \log\log n - \frac{ap_n K_n}{2}\left(1 - \frac{\ell-1}{n-1}\right) \leq -c\log n.$$

Therefore, it follows for all n sufficiently large that

$$\binom{n}{\ell}\binom{n-\ell}{r} \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \leq 2n^{\ell+1} \cdot e^{-cr\log n}, \tag{76}$$

yielding that

$$\sum_{r=R+1}^{\lfloor \frac{n-\ell}{2} \rfloor} \binom{n}{\ell}\binom{n-\ell}{r} \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] \leq 2n^{\ell+1} \sum_{r=R+1}^{\infty} n^{-cr} = \frac{2n^{\ell+1-c(R+1)}}{1-n^{-c}}.$$

Now, with R selected such that

$$R > \frac{\ell+1}{c} - 1$$

we get $\ell + 1 - c(R+1) < 0$. Then, letting $n \to \infty$ we obtain the desired result

$$\lim_{n\to\infty} \sum_{r=R+1}^{\lfloor \frac{n-\ell}{2} \rfloor} \binom{n}{\ell}\binom{n-\ell}{r} \mathbb{P}\left[\mathcal{A}_{\ell,r}\right] = 0.$$

■

Acknowledgements

This work was supported by the Department of Electrical and Computer Engineering and CyLab at Carnegie Mellon University, and also by Grant DAAD19-02- 1-0389 and MURI Grant W 911 NF 0710287 from the Army Research Office, and Grants CCF-0424422 and CNS-0831440 from the National Science Foundation to the Berkeley TRUST STC.

References

[1] I. F. Akyildiz, Y. Sankarasubramaniam, W. Su and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks* **38**, pp. 393-422.

[2] S.R. Blackburn and S. Gerke, "Connectivity of the uniform random intersection graph," *Discrete Mathematics* **309** (2009), pp. 5130-5140.

[3] M. Bloznelis, J. Jaworski and K. Rybarczyk, "Component evolution in a secure wireless sensor network," *Networks* **53** (2009), pp. 19-26.

[4] B. Bollobás, *Random Graphs*, Second Edition, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge (UK), 2001.

[5] S. A. Çamtepe and B. Yener, "Key Distribution Mechanisms for Wireless Sensor Networks: a Survey," Technical Report TR-05-07, Computer Science Department, Rensselaer Polytechnic Institute, Troy (NY), March 2005.

[6] H. Chan, A. Perrig and D. Song, "Random key predistribution schemes for sensor networks," in Proceedings of the 2003 IEEE Symposium on Research in Security and Privacy (SP 2003), Oakland (CA), May 2003, pp. 197-213.

[7] R. Di Pietro, L.V. Mancini, A. Mei, A. Panconesi and J. Radhakrishnan, "Redoubtable sensor networks," *ACM Transactions on Information Systems Security* **TISSEC 11** (2008), pp. 1-22.

[8] W. Du, J. Deng, Y.S. Han and P.K. Varshney, "A pairwise key pre-distribution scheme for wireless sensor networks," in Proceedings of the 10th ACM Conference on Computer and Communications Security (CCS 2003), Washington (DC), October 2003, pp. 42-51.

[9] D. Dubhashi and A. Panconesi, *Concentration of Measure for the Analysis of Randomized Algorithms*, Cambridge University Press, New York (NY), 2009.

[10] P. Erdős and A. Rényi, "On the Strength of Connectedness of Random Graphs," *Acta Math. Acad. Sci. Hungar* **12**, 1961, pp. 261–267.

[11] L. Eschenauer and V.D. Gligor, "A key-management scheme for distributed sensor networks," in Proceedings of the ACM Conference on Computer and Communications Security (CSS 2002), Washington (DC), November 2002, pp. 41-47.

[12] T.I. Fenner and A.M. Frieze, "On the connectivity of random m-orientable graphs and digraphs," *Combinatorica* **2** (1982), pp. 347-359.

[13] J. Hwang and Y. Kim, "Revisiting random key pre-distribution schemes for wireless sensor networks," in Proceedings of the Second ACM Workshop on Security of Ad Hoc And Sensor Networks (SASN 2004), Washington (DC), October 2004.

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

[15] K. Joag-Dev and F. Proschan, "Negative association of random variables, with applications," *The Annals of Statistics* **11** (1983), pp. 266-295

[16] G.E. Martin, *Counting: The Art of Enumerative Combinatorics*, Springer Verlag New York, 2001.

[17] K. Rybarczyk, "Diameter, connectivity and phase transition of the uniform random intersection graph," *Discrete Mathematics* **311** (2011), pp. 1998-2019.

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

[19] A. Perrig, J. Stankovic and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM* **47** (2004), pp. 53–57.

[20] T. K. Philips, D. F. Towsley, and J. Wolf, "On the diameter of a class of random graphs," *IEEE Transactions on Information Theory* **36**:2, 1990, pp. 285-288.

[21] D.-M. Sun and B. He, "Review of key management mechanisms in wireless sensor networks," *Acta Automatica Sinica* **12** (2006), pp. 900-906.

[22] Y. Wang, G. Attebury and B. Ramamurthy, "A survey of security issues in wireless sensor networks," *IEEE Communications Surveys & Tutorials* **8** (2006), pp. 2-23.

[23] Y. Xiao, V.K. Rayi, B. Sun, X. Du, F. Hu and M. Galloway, "A survey of key management schemes in wireless sensor networks," *Computer Communications* **30** (2007), pp. 2314-2341.

[24] O. Yağan, "Performance of the Eschenauer-Gligor key distribution scheme under an ON/OFF channel," *IEEE Transactions on Information Theory* **IT-58** (2012), pp. 3821-3835.

[25] O. Yağan, *Random Graph Modeling of Key Distribution Schemes in Wireless Sensor Networks*, Ph.D. Thesis, Department of Electrical and Computer Engineering, University of Maryland, College Park (MD), June 2011.

[26] O. Yağan and A. M. Makowski, "On the gradual deployment of random pairwise key distribution schemes," in Proceedings of the 9th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2011), Princeton (NJ), May 2011.

[27] O. Yağan and A. M. Makowski, "On the resiliency of sensor networks under the pairwise key distribution scheme," in Proceedings of the 22nd IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2011), Toronto (CANADA), 1218-1222, 2011.

[28] O. Yağan and A. M. Makowski, "Designing securely connected wireless sensor networks in the presence of unreliable links," in Proceedings of the 2011 IEEE International Conference on Communications (ICC 2011), Kyoto (Japan), June 2011.

[29] O. Yağan and A.M. Makowski, "Zero-one laws for connectivity in random key graphs," *IEEE Transactions on Information Theory* **IT-58** (2012), pp. 2983-2999.

[30] O. Yağan and A. M. Makowski, "On the scalability of the random pairwise key distribution scheme: Gradual deployment and key ring sizes," *Performance Evaluation* **70**(7-8):493-512, July 2013

[31] O. Yağan and A. M. Makowski, "On the connectivity of sensor networks under random pairwise key predistribution," *IEEE Transactions on Information Theory* **IT-59** (2013), pp. 5754-5762.

[32] O. Yağan and A.M. Makowski, "Modeling the pairwise key predistribution scheme in the presence of unreliable links," *IEEE Transactions on Information Theory* **IT-59** (2013), pp. 1740-1760.

[33] F. Yavuz, J. Zhao, O. Yağan and V. Gligor, "On secure and reliable communications in wireless sensor networks: Towards k-connectivity under a random pairwise key predistribution scheme," in Proceedings of the IEEE International Symposium on Information Theory (ISIT 2014), Hawaii (HI), July 2014.

[34] F. Yavuz, J. Zhao, O. Yağan and V. Gligor, "Towards k-connectivity of the random graph induced by a pairwise key predistribution scheme with unreliable links," *IEEE Transactions on Information Theory* **IT-61** (2015), pp. 6251-6271.

[35] C.W. Yi, P.J. Wan, K.W. Lin and C.H. Huang, "Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with unreliable nodes and links," in Proceedings of IEEE Globecom 2006, San Francisco (CA), November 2006.

[36] J. Zhao, O. Yağan and V. Gligor, "Secure k-Connectivity in Wireless Sensor Networks under an On/Off Channel Model," in Proceedings of the IEEE International Symposium on Information Theory (ISIT 2013), Istanbul (Turkey), July 2013.

[37] J. Zhao, O. Yağan and V. Gligor, "k-Connectivity in Random Key Graphs with Unreliable Links, *IEEE Transactions on Information Theory* **IT-61** (2015), pp. 3810–3836.