Exact Analysis of k-Connectivity in Secure Sensor Networks with Unreliable Links

Jun Zhao
CyLab and Dept. of ECE
Carnegie Mellon University
Email: junzhao@cmu.edu

Osman Yağan
CyLab and Dept. of ECE
Carnegie Mellon University
Email: oyagan@ece.cmu.edu

Virgil Gligor
CyLab and Dept. of ECE
Carnegie Mellon University
Email: gligor@cmu.edu

***Abstract*—The Eschenauer–Gligor (EG) random key predistribution scheme has been widely recognized as a typical approach to secure communications in wireless sensor networks (WSNs). However, there is a lack of precise probability analysis on the reliable connectivity of WSNs under the EG scheme. To address this, we rigorously derive the asymptotically exact probability of k-connectivity in WSNs employing the EG scheme with unreliable links represented by independent on/off channels, where k-connectivity ensures that the network remains connected despite the failure of any $(k-1)$ sensors or links. Our analytical results are confirmed via numerical experiments, and they provide precise guidelines for the design of secure WSNs that exhibit a desired level of reliability against node and link failures.**

***Index Terms*—Connectivity, key predistribution, minimum degree, random graphs, security, wireless sensor networks.**

I. Introduction

The Eschenauer–Gligor (EG) random key predistribution scheme [4] has been widely regarded as a typical solution to secure communications in wireless sensor networks (WSNs) [5], [6], [7], [8], [9], [10], [12], [15]. The scheme operates as follows. In a WSN with n sensors, before deployment, each sensor is independently assigned K_n distinct keys which are selected *uniformly at random* from a pool of P_n keys, where K_n and P_n are both functions of n. After deployment, any two sensors can securely communicate over an existing wireless link if and only if they share at least one key.

Wireless links between nodes may become unavailable due to the presence of physical barriers between nodes or because of harsh environmental conditions severely impairing transmission. We model unreliable links as independent channels, each being *on* with probability p_n or being *off* with probability $(1-p_n)$, where p_n is a function of n for generality. Such on/off channel model has been used in the context of secure WSNs [9], [15], [12], and is shown to well approximate the disk model [5], [6], [9], [15], [12], where any two nodes need to be within a certain distance to establish a wireless link in between.

Given the randomness involved in the EG key predistribution scheme, and the unreliability of wireless links, there arises a basic question as to how one can adjust the EG scheme parameters K_n and P_n, and the link parameter p_n, so that the resulting network is securely and reliably connected. 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). To answer the question above, this paper presents the asymptotically exact probability of k-connectivity in secure WSNs under the EG scheme with unreliable links. 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. An equivalent definition is that each node can find at least k internally node-disjoint paths to any other node. With $k=1$, k-connectivity simply means connectivity.

Our result on the asymptotically exact probability of k-connectivity complements a zero-one law established in our prior work [15], [12], and is significant to obtain a precise understanding of the connectivity behavior of secure WSNs. First, with the zero-one law, one is only provided with design choices which lead to networks that are k-connected with high probability or to that are not k-connected with high probability, where an event happens "with high probability" if its probability asymptotically converges to 1. Given the trade-offs involved between connectivity, security and memory load [4], [9], it would be more useful to have a complete picture by obtaining the asymptotically exact probability of k-connectivity. In addition, there may be situations where the network designer is interested in having a guaranteed level of k-connectivity (one-laws would provide conditions for that) but may also be interested in having some level of k-connectivity without such guarantees (one-laws would fall short in providing this). Our result fills this gap. Finally, it is not possible to determine the width of the phase transition from zero-one laws; the width of the phase transition is often calculated by the difference in parameters that it takes to increase the probability of k-connectivity from ϵ to $(1-\epsilon)$, for some $\epsilon < 0.5$. In other words, it is not clear from zero-one laws how sensitive the probability of k-connectivity is to the variations in the EG scheme parameters K_n and P_n, and the link parameter p_n. By providing the asymptotically exact probability of k-connectivity, our findings provide a clear picture of these intricate relationships.

The rest of the paper is organized as follows. We describe the system model in Section II. Section III presents the main results as Theorem 1, which is established in Section IV. In Section VI, we present numerical experiments that confirm our analytical findings. Afterwards, Section VII surveys related work, and Section VIII concludes the paper. The Appendix presents a few useful lemmas and their proofs.

II. System Model

We now explain the system model. Consider a WSN with n sensors operating under the EG scheme and with wireless

978-3-9018-8273-9/15/ ©2015 IFIP

links modeled by independent on/off channels. Let a node set $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ represent the n sensors. According to the EG scheme, each node $v_i \in \mathcal{V}$ is independently assigned a set (denoted by S_i) of K_n distinct cryptographic keys, which are selected *uniformly at random* from a key pool of P_n keys. Any pair of nodes can then secure an existing communication link as long as they have at least one key in common.

The EG scheme results in a *random key graph* [1], [7], [10], also known as a *uniform random intersection graph*. This graph denoted by $G(n, K_n, P_n)$ is defined on the node set $\mathcal{V}$ such that any two distinct nodes v_i and v_j have an edge in between, an event denoted by Γ_{ij}, if and only if they share at least one key. Thus, the event Γ_{ij} means $\big(S_i \cap S_j \neq \emptyset\big)$.

Under the on/off channel model for unreliable links, each wireless link is independently being *on* with probability p_n or being *off* with probability $(1 - p_n)$. Defining C_{ij} as the event that the channel between v_i and v_j is *on*, we have $\mathbb{P}[C_{ij}] = p_n$, with $\mathbb{P}[A]$ throughout the paper meaning the probability that event A happens. The on/off channel model induces an *Erdős-Rényi graph* $G(n, p_n)$ [2] defined on the node set $\mathcal{V}$ such that v_i and v_j have an edge in between if C_{ij} takes place.

Finally, we denote by $\mathbb{G}(n, K_n, P_n, p_n)$ the underlying graph of the n-node WSN under the EG scheme with unreliable links. We often write $\mathbb{G}$ rather than $\mathbb{G}(n, K_n, P_n, p_n)$ for brevity. Graph $\mathbb{G}$ is defined on the node set $\mathcal{V}$ such that there exists an edge between nodes v_i and v_j if events Γ_{ij} and C_{ij} happen at the same time. We set event $E_{ij} := \Gamma_{ij} \cap C_{ij}$ and also write E_{ij} as $E_{v_i v_j}$ when necessary. It is clear that $\mathbb{G}$ is the intersection of $G(n, K_n, P_n)$ and $G(n, p_n)$; i.e.,

$$\mathbb{G} = G(n, K_n, P_n) \cap G(n, p_n). \tag{1}$$

We define s_n as the probability that two distinct nodes share at least one key and q_n as the probability that two distinct nodes have an edge in between in graph $\mathbb{G}$. Clearly, s_n and q_n both depend on K_n and P_n, while q_n depends also on p_n. As shown in previous work [1], [7], [10], s_n is determined through

$$s_n = \mathbb{P}[\Gamma_{ij}] = \begin{cases} 1 - \binom{P_n - K_n}{K_n} \big/ \binom{P_n}{K_n}, & \text{if } P_n > 2K_n, \\ 1, & \text{if } P_n \leq 2K_n. \end{cases}$$

Then by the independence of C_{ij} and Γ_{ij}, we have

$$q_n = \mathbb{P}[E_{ij}] = \mathbb{P}[C_{ij}] \cdot \mathbb{P}[\Gamma_{ij}] = p_n \cdot s_n \tag{2}$$

$$= \begin{cases} p_n \cdot \big[1 - \binom{P_n - K_n}{K_n} \big/ \binom{P_n}{K_n}\big], & \text{if } P_n > 2K_n, \\ p_n, & \text{if } P_n \leq 2K_n. \end{cases} \tag{3}$$

III. The Main Results

We present the main results below. Throughout the paper, k is a positive integer and does not scale with n, and e is the base of the natural logarithm function, ln. We use the standard asymptotic notation $o(\cdot), O(\cdot), \omega(\cdot), \Omega(\cdot), \Theta(\cdot)$ and $\sim$; in particular, for two positive sequences a_n and b_n, the relation $a_n \sim b_n$ means $\lim_{n\to\infty} a_n/b_n = 1$.

Theorem 1. *For graph $\mathbb{G}(n, K_n, P_n, p_n)$ under $P_n = \Omega(n)$ and $\frac{K_n}{P_n} = o(1)$, with q_n denoting the edge probability and a sequence α_n defined through*

$$q_n = \frac{\ln n + (k-1)\ln\ln n + \alpha_n}{n}, \tag{4}$$

if $\lim_{n\to\infty} \alpha_n = \alpha^* \in (-\infty, \infty)$, then as $n \to \infty$,

$$\mathbb{P}\left[\,\textit{Graph } \mathbb{G}(n, K_n, P_n, p_n) \textit{ is } k\textit{-connected.}\,\right] \to e^{-\frac{e^{-\alpha^*}}{(k-1)!}}.$$

Theorem 1 provides the asymptotically exact probability of k-connectivity in graph $\mathbb{G}$. Its proof is given in the next section. From (3), for all n sufficiently large, under $P_n > 2K_n$ which is clearly implied by the condition $\frac{K_n}{P_n} = o(1)$, the edge probability q_n in graph $\mathbb{G}$ is given by the expression $p_n \cdot \big[1 - \binom{P_n - K_n}{K_n} \big/ \binom{P_n}{K_n}\big]$. With a much simpler approximation $p_n \cdot \frac{{K_n}^2}{P_n}$ for q_n, we present below a corollary of Theorem 1.

Corollary 1. *For graph $\mathbb{G}(n, K_n, P_n, p_n)$ under $P_n = \Omega(n)$ and $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big)$, with a sequence β_n defined through*

$$p_n \cdot \frac{{K_n}^2}{P_n} = \frac{\ln n + (k-1)\ln\ln n + \beta_n}{n}, \tag{5}$$

if $\lim_{n\to\infty} \beta_n = \beta^* \in (-\infty, \infty)$, then as $n \to \infty$,

$$\mathbb{P}\left[\,\textit{Graph } \mathbb{G}(n, K_n, P_n, p_n) \textit{ is } k\textit{-connected.}\,\right] \to e^{-\frac{e^{-\beta^*}}{(k-1)!}}.$$

Setting $p_n = 1$ in Theorem 1 and Corollary 1, we obtain the corresponding results for random key graph $G(n, K_n, P_n)$ in view of (1). Furthermore, we can use monotonicity arguments [15] to derive the zero-one laws for k-connectivity in graph $\mathbb{G}$. Specifically, under the conditions of Theorem 1 (resp., Corollary 1), graph $\mathbb{G}$ is k-connected with high probability if $\lim_{n\to\infty}\alpha_n = \infty$ (resp., $\lim_{n\to\infty}\beta_n = \infty$), and is not k-connected with high probability if $\lim_{n\to\infty}\alpha_n = -\infty$ (resp., $\lim_{n\to\infty}\beta_n = -\infty$). The arguments are straightforward from our work [15] and are omitted here due to space limitation.

Before establishing Corollary 1 using Theorem 1, we explain the practicality of the conditions in Theorem 1 and Corollary 1: $P_n = \Omega(n)$, $\frac{K_n}{P_n} = o(1)$ and $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big)$. First, the condition $P_n = \Omega(n)$ indicates that the key pool size P_n should grow at least linearly with n, which holds in practice [4], [10], [9]. Second, the condtions $\frac{K_n}{P_n} = o(1)$ and $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big)$ (note that the latter implies the former) are also practical in secure sensor network applications since P_n is expected to be several orders of magnitude larger than K_n [4], [10], [9].

We now prove Corollary 1 using Theorem 1. We have the conditions of Corollary 1: $P_n = \Omega(n)$, $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big)$, and (5) with $\lim_{n\to\infty} \beta_n = \beta^* \in (-\infty, \infty)$. First, it is clear that $\beta_n = \beta^* \pm o(1)$. Under $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big) = o(1)$, from [15, Lemma 8], it holds that $s_n = \frac{{K_n}^2}{P_n} \cdot \big[1 \pm O\big(\frac{{K_n}^2}{P_n}\big)\big]$. In view of the above, we obtain from (2) and (5) that

$$\begin{aligned} q_n &= p_n \cdot s_n = p_n \cdot \tfrac{{K_n}^2}{P_n} \cdot \big[1 \pm O\big(\tfrac{{K_n}^2}{P_n}\big)\big] \\ &= \tfrac{\ln n + (k-1)\ln\ln n + \beta_n}{n} \cdot \big[1 \pm o\big(\tfrac{1}{\ln n}\big)\big] \\ &= \tfrac{\ln n + (k-1)\ln\ln n + \beta^* \pm o(1)}{n}. \end{aligned} \tag{6}$$

With α_n defined by (4), we use (6) to derive $\alpha_n = \beta^* \pm o(1)$, which yields that α^* denoting $\lim_{n\to\infty} \alpha_n$ equals β^*. Then in view of $\alpha^* = \beta^*$ and that the conditions of Theorem 1 all hold given the conditions of Corollary 1 (note that $\frac{{K_n}^2}{P_n} = o\big(\frac{1}{\ln n}\big)$ implies $\frac{K_n}{P_n} = o(1)$), Corollary 1 follows from Theorem 1.

IV. Establishing Theorem 1

For any graph, k-connectivity implies that its minimum degree is at least k, while the other way does not hold since a graph may have isolated components, each of which is k-connected within itself. However, for random graph $\mathbb{G}(n, K_n, P_n, p_n)$, as given by Lemma 1 below, we have shown it is unlikely under certain conditions that $\mathbb{G}(n, K_n, P_n, p_n)$ is not k-connected but has a minimum degree at least k.

Lemma 1 ([15, Section IX])**.** *For graph $\mathbb{G}(n, K_n, P_n, p_n)$ under $P_n = \Omega(n)$, $\frac{K_n}{P_n} = o(1)$ and $q_n = o(1)$, it holds that*

$$\mathbb{P}\left[\begin{array}{c}\textit{Graph } \mathbb{G} \textit{ is not } k\textit{-connected,}\\ \textit{but has a minimum degree at least } k.\end{array}\right] = o(1).$$

We show that the conditions in Lemma 1 all hold given the conditions of Theorem 1: $P_n = \Omega(n)$, $\frac{K_n}{P_n} = o(1)$ and $q_n = \frac{\ln n + (k-1)\ln\ln n + \alpha_n}{n}$ with $\lim_{n\to\infty} \alpha_n = \alpha^* \in (-\infty, \infty)$. To see this, we only need to prove $q_n = o(1)$ needed in Lemma 1 follows from the conditions of Theorem 1. Clearly, it holds that $|\alpha_n| = O(1)$ from $\lim_{n\to\infty} \alpha_n = \alpha^* \in (-\infty, \infty)$. Then in view of $|\alpha_n| = O(1)$ and the fact that k does not scale with n, we obtain from (4) that

$$q_n \sim \frac{\ln n}{n}, \tag{7}$$

which clearly implies $q_n = o(1)$.

From Lemma 1 and

$$\begin{aligned}&\mathbb{P}\left[\,\text{Graph } \mathbb{G} \text{ is } k\text{-connected.}\,\right]\\ &= \mathbb{P}\left[\,\text{Graph } \mathbb{G} \text{ has a minimum degree at least } k.\,\right]\\ &\quad - \mathbb{P}\left[\begin{array}{c}\text{Graph } \mathbb{G} \text{ is not } k\text{-connected,}\\ \text{but has a minimum degree at least } k.\end{array}\right],\end{aligned}$$

Theorem 1 on k-connectivity of $\mathbb{G}$ will be proved once we demonstrate Lemma 2 below on the minimum degree of $\mathbb{G}$.

Lemma 2. *Under the conditions of Theorem 1, it holds that $\lim_{n\to\infty} \mathbb{P}[\mathbb{G}$ has a minimum degree at least $k.] = e^{-\frac{e^{-\alpha^*}}{(k-1)!}}$.*

To prove Lemma 2, we first show that the number of nodes in $\mathbb{G}$ with a certain degree converges in distribution to a Poisson random variable. With ϕ_h denoting the number of nodes with degree h in $\mathbb{G}$, $h = 0, 1, \ldots$, we use the method of moments to prove that ϕ_h asymptotically follows a Poisson distribution with mean λ_h. Specifically, from [11, Theorem 7], it follows for any integers $h \geq 0$ and $\ell \geq 0$ that

$$\mathbb{P}[\phi_h = \ell] \sim (\ell!)^{-1} {\lambda_h}^{\ell} e^{-\lambda_h}, \tag{8}$$

since $\mathbb{P}[\text{Nodes } v_1, v_2, \ldots, v_m \text{ all have degree } h] \sim {\lambda_h}^m / n^m$, which is shown by Lemma 3 below with

$$\lambda_h = n(h!)^{-1}(nq_n)^h e^{-nq_n}. \tag{9}$$

Lemma 3. *For graph $\mathbb{G}$ under the conditions of Theorem 1, $\mathbb{P}[v_1, v_2, \ldots, v_m$ all have degree $h] \sim (h!)^{-m}(nq_n)^{hm} e^{-mnq_n}$ holds for any integers $m \geq 1$ and $h \geq 0$.*

As explained above, Lemma 3 shows (8) with λ_h given by (9). Then the proof of Lemma 2 will be completed once we establish Lemma 3 and the result that (8) implies Lemma 2. Below we will demonstrate that (8) implies Lemma 2, and then detail the proof of Lemma 3.

A. Proving that (8) implies Lemma 2

Recall that ϕ_h denotes the number of nodes with degree h in graph $\mathbb{G}$. With δ defined as the minimum degree of graph $\mathbb{G}$, then the event $(\delta \geq k)$ is the same as $\bigcap_{h=0}^{k-1}(\phi_h = 0)$ (i.e., the event that no node has a degree falling in $\{0, 1, \ldots, k-1\}$). Hence, we obtain

$$\mathbb{P}[\delta \geq k] = \mathbb{P}\left[\bigcap_{h=0}^{k-1}(\phi_h = 0)\right] \leq \mathbb{P}[\phi_{k-1} = 0]; \tag{10}$$

and by the union bound, it holds that

$$\begin{aligned}\mathbb{P}[\delta \geq k] &= \mathbb{P}\left[\,(\phi_{k-1} = 0) \cap \overline{\left(\bigcup_{h=0}^{k-2}(\phi_h \neq 0)\right)}\,\right]\\ &\geq \mathbb{P}[\phi_{k-1} = 0] - \sum_{h=0}^{k-2}\mathbb{P}[\phi_h \neq 0].\end{aligned} \tag{11}$$

To use (10) and (11), we compute $\mathbb{P}[\phi_h \neq 0]$ given (8) and thus evaluate λ_h specified in (9). Applying (4) and (7) to (9), and considering $\lim_{n\to\infty} \alpha_n = \alpha^*$ with $|\alpha^\star| < \infty$, we establish

$$\begin{aligned}\lambda_h &= n(h!)^{-1}(nq_n)^h e^{-nq_n}\\ &\sim n(h!)^{-1}(\ln n)^h \cdot e^{-\ln n - (k-1)\ln\ln n - \alpha_n}\\ &= (h!)^{-1}(\ln n)^{h+1-k} e^{-\alpha_n}\\ &\to \begin{cases} 0, & \text{for } h = 0, 1, \ldots, k-2,\\ \frac{e^{-\alpha^*}}{(k-1)!}, & \text{for } h = k-1,\\ \infty, & \text{for } h = k, k+1, \ldots\end{cases}\end{aligned} \tag{12}$$

By (8) and (12), we derive that as $n \to \infty$,

$$\mathbb{P}[\phi_h = 0] \to \begin{cases} 1, & \text{for } h = 0, 1, \ldots, k-2,\\ e^{-\frac{e^{-\alpha^*}}{(k-1)!}}, & \text{for } h = k-1,\\ 0, & \text{for } h = k, k+1, \ldots\end{cases} \tag{13}$$

Using (13) in (10) and (11), we obtain $\mathbb{P}[\delta \geq k] \to e^{-\frac{e^{-\alpha^*}}{(k-1)!}}$; i.e., Lemma 2 is proved.

B. Proving Lemma 3

We use $\mathcal{V}_m$ to denote the node set $\{v_1, v_2, \ldots, v_m\}$. Lemma 3 evaluates the probability that each of $\mathcal{V}_m$ has degree h. To compute such probability, we look at whether at least two of $\mathcal{V}_m$ have an edge in between, and whether at least two of $\mathcal{V}_m$ have at least one common neighbor. To this end, we define $\mathcal{P}_1$ as the probability of event

$$\begin{array}{c}(\text{each of } \mathcal{V}_m \text{ has degree } h)\\ \cap\,\big[(\text{at least two of } \mathcal{V}_m \text{ have an edge in between})\\ \cup\,(\text{at least two of } \mathcal{V}_m \text{ have at least one common neighbor})\big],\end{array}$$

and define $\mathcal{P}_2$ as the probability of event

$$\begin{array}{c}(\text{each of } \mathcal{V}_m \text{ has degree } h)\\ \cap\,(\text{no two of } \mathcal{V}_m \text{ have any edge in between})\\ \cap\,(\text{no two of } \mathcal{V}_m \text{ have any common neighbor}).\end{array}$$

Then $\mathbb{P}[\text{each of } \mathcal{V}_m \text{ has degree } h] = \mathcal{P}_1 + \mathcal{P}_2$. Thus, Lemma 3 will hold once we establish the following two propositions.

Proposition 1. *Under the conditions of Theorem 1, it holds that $\mathcal{P}_1 = o\big((h!)^{-m}(nq_n)^{hm}e^{-mnq_n}\big)$.*

Proposition 2. *Under the conditions of Theorem 1, it holds that $\mathcal{P}_2 \sim (h!)^{-m}(nq_n)^{hm}e^{-mnq_n}$.*

To prove Propositions 1 and 2, we analyze below how nodes in graph $\mathbb{G}$ have edges. We first look at how edges exist between $v_1, v_2, \ldots, v_m$. Recalling C_{ij} as the event that the communication channel between distinct nodes v_i and v_j is *on*, we set $\mathbf{1}[C_{ij}]$ as the indicator variable of event C_{ij} by

$$\mathbf{1}[C_{ij}] := \begin{cases} 1, & \text{if the channel between } v_i \text{ and } v_j \text{ is } \textit{on},\\ 0, & \text{if the channel between } v_i \text{ and } v_j \text{ is } \textit{off}.\end{cases}$$

We denote by $\mathcal{C}_m$ a $\binom{m}{2}$-tuple consisting of all possible $\mathbf{1}[C_{ij}]$ with $1 \leq i < j \leq m$ as follows:

$$\mathcal{C}_m := (\mathbf{1}[C_{12}], \dots, \mathbf{1}[C_{1m}], \quad \mathbf{1}[C_{23}], \dots, \mathbf{1}[C_{2m}], \\ \mathbf{1}[C_{34}], \dots, \mathbf{1}[C_{3m}], \quad \dots, \quad \mathbf{1}[C_{(m-1),m}]).$$

Recalling S_i as the key set on node v_i, we define a m-tuple $\mathcal{T}_m$ through $\mathcal{T}_m := (S_1, S_2, \dots, S_m)$. Then we define $\mathcal{L}_m$ as $\mathcal{L}_m := (\mathcal{C}_m, \mathcal{T}_m)$. With $\mathcal{L}_m$, we have the *on/off* states of all channels between nodes $v_1, v_2, \dots, v_m$ and the key sets $S_1, S_2, \dots, S_m$ on these m nodes, so all edges between these m nodes in graph $\mathbb{G}$ are determined. Let $\mathbb{C}_m, \mathbb{T}_m$ and $\mathbb{L}_m$ be the sets of all possible $\mathcal{C}_m, \mathcal{T}_m$ and $\mathcal{L}_m$, respectively.

Now we further introduce some notation to characterize how nodes $v_1, v_2, \dots, v_m$ have edges with nodes of $\overline{\mathcal{V}_m}$, where $\overline{\mathcal{V}_m}$ denotes $\{v_{m+1}, v_{m+2}, \dots, v_n\}$. Let N_i be the neighborhood set of node v_i, i.e., the set of nodes that have edges with v_i. We also define set $\overline{N_i}$ as the set $\{v_{m+1}, v_{m+2}, \dots, v_n\} \setminus N_i$. Then we are ready to define sets $M_{j_1 j_2 \dots j_m}$ for all $j_1, j_2, \dots, j_m \in \{0, 1\}$ which characterize the relationships between sets N_i for $i = 1, 2, \dots, m$. We define

$$M_{j_1 j_2 \dots j_m} := \Bigg(\bigcap_{i \in \{1,2,\dots,m\}: j_i = 1} N_i\Bigg) \cap \Bigg(\bigcap_{i \in \{1,2,\dots,m\}: j_i = 0} \overline{N_i}\Bigg). \tag{14}$$

In other words, for $i = 1, 2, \dots, m$, if N_i is not empty, each node in N_i belongs to $M_{j_1 j_2 \dots j_m}$ if $j_i = 1$ and does not belong to $M_{j_1 j_2 \dots j_m}$ if $j_i = 0$. Also, if $j_1 = j_2 = \dots = j_m = 0$, then $M_{j_1 j_2 \dots j_m} = \bigcap_{i=1}^m \overline{N_i}$. The sets $M_{j_1 j_2 \dots j_m}$ for $j_1, j_2, \dots, j_m \in \{0, 1\}$ are mutually disjoint, and constitute a partition of the set $\overline{\mathcal{V}_m}$ (a partition is allowed to contain empty sets here). By the definition of $M_{j_1 j_2 \dots j_m}$ for $j_1, j_2, \dots, j_m \in \{0, 1\}$, we have

$$\sum_{j_1, j_2, \dots, j_m \in \{0,1\}} |M^*_{j_1 j_2 \dots j_m}| = |\overline{\mathcal{V}_m}| = n - m, \tag{15}$$

and

$$\sum_{\substack{j_1, j_2, \dots, j_m \in \{0,1\}: \\ \sum_{i=1}^m j_i \geq 1.}} |M_{j_1 j_2 \dots j_m}| = \Bigg|\Bigg(\bigcup_{i=1}^m N_i\Bigg) \cap \overline{\mathcal{V}_m}\Bigg|. \tag{16}$$

We further define 2^m-tuple $\mathcal{M}_m$ through[1]

$$\mathcal{M}_m = \big(|M_{j_1 j_2 \dots j_m}| \;\big|\; j_1, j_2, \dots, j_m \in \{0, 1\}\big) \\ = \big(|M_{0^m}|, |M_{0^{m-1},1}|, |M_{0^{m-2}1,0}|, |M_{0^{m-2}1,1}|, \dots\big),$$

where $|M_{j_1 j_2 \dots j_m}|$ means the cardinality of $M_{j_1 j_2 \dots j_m}$.

Under event $\mathcal{E}_2$, the set $\mathcal{M}_m$ is determined and we denote its value by $\mathcal{M}_m^{(0)}$, which satisfies

$$\begin{cases} |M_{0^{i-1},1,0^{m-i}}| = h, & \text{for } i = 1, 2, \dots, m; \\ |M_{j_1 j_2 \dots j_m}| = 0, & \text{for } \sum_{i=1}^m j_i > 1; \\ |M_{0^m}| = n - m - hm. \end{cases} \tag{17}$$

To analyze event $\mathcal{E}_2$, we define $\mathbb{L}_m^{(0)}$ such that $\big(\mathcal{L}_m \in \mathbb{L}_m^{(0)}\big)$ is the event that no two of nodes $v_1, v_2, \dots, v_m$ have any common neighbor. In view of events $\big(\mathcal{L}_m \in \mathbb{L}_m^{(0)}\big)$, $\big(\mathcal{M}_m = \mathcal{M}_m^{(0)}\big)$ and $\mathcal{E}_2$, then $\mathcal{E}_2$ is the same as $\big(\mathcal{L}_m \in \mathbb{L}_m^{(0)}\big) \cap \big(\mathcal{M}_m = \mathcal{M}_m^{(0)}\big)$; i.e.,

$$\mathcal{E}_2 = \big[\big(\mathcal{L}_m \in \mathbb{L}_m^{(0)}\big) \cap \big(\mathcal{M}_m = \mathcal{M}_m^{(0)}\big)\big]. \tag{18}$$

We define $\mathbb{M}_m(\mathcal{L}_m)$ for $\mathcal{L}_m \in \mathbb{L}_m$ as the set of $\mathcal{M}_m$ under which each of $\mathcal{V}_m$ has degree h. Thus, the event that each of $\mathcal{V}_m$ has degree h is $\big(\mathcal{L}_m \in \mathbb{L}_m\big) \cap \big(\mathcal{M}_m \in \mathbb{M}_m(\mathcal{L}_m)\big)$, which together with (18) yields

$$\mathcal{E}_1 = \bigcup_{\substack{\mathcal{L}_m^* \in \mathbb{L}_m,\, \mathcal{M}_m^* \in \mathbb{M}_m(\mathcal{L}_m^*): \\ (\mathcal{L}_m^* \notin \mathbb{L}_m^{(0)}) \text{ or } (\mathcal{M}_m^* \neq \mathcal{M}_m^{(0)})}} \mathbb{P}\big[\big(\mathcal{L}_m = \mathcal{L}_m^*\big) \cap \big(\mathcal{M}_m = \mathcal{M}_m^*\big)\big]. \tag{19}$$

Now we prove Propositions 1 and 2 based on (18) and (19). The inequality below following from (7) will be applied often:

$$q_n \leq \frac{2 \ln n}{n} \text{ for all } n \text{ sufficiently large.} \tag{20}$$

1) The Proof of Proposition 1

In view of (19) and considering the disjointness of events $\big(\mathcal{L}_m = \mathcal{L}_m^*\big) \cap \big(\mathcal{M}_m = \mathcal{M}_m^*\big)$ for $\mathcal{L}_m^* \in \mathbb{L}_m$ and $\mathcal{M}_m^* \in \mathbb{M}_m(\mathcal{L}_m^*)$, we express $\mathbb{P}[\mathcal{E}_1]$ as

$$\sum_{\substack{\mathcal{L}_m^* \in \mathbb{L}_m,\, \mathcal{M}_m^* \in \mathbb{M}_m(\mathcal{L}_m^*): \\ (\mathcal{L}_m^* \notin \mathbb{L}_m^{(0)}) \text{ or } (\mathcal{M}_m^* \neq \mathcal{M}_m^{(0)})}} \mathbb{P}\big[\big(\mathcal{L}_m = \mathcal{L}_m^*\big) \cap \big(\mathcal{M}_m = \mathcal{M}_m^*\big)\big] \tag{21}$$

We evaluate (21) by computing

$$\mathbb{P}\big[\big(\mathcal{M}_m = \mathcal{M}_m^*\big) \mid \mathcal{L}_m = \mathcal{L}_m^*\big]. \tag{22}$$

With $\mathcal{C}_m^*$ and $\mathcal{T}_m^*$ defined such that $\mathcal{L}_m^* = (\mathcal{C}_m^*, \mathcal{T}_m^*)$, event $(\mathcal{L}_m = \mathcal{L}_m^*)$ is the union of events $(\mathcal{C}_m = \mathcal{C}_m^*)$ and $(\mathcal{T}_m = \mathcal{T}_m^*)$. Since $(\mathcal{C}_m = \mathcal{C}_m^*)$ and $(\mathcal{M}_m = \mathcal{M}_m^*)$ are independent, we get

$$(22) = \mathbb{P}\big[\big(\mathcal{M}_m = \mathcal{M}_m^*\big) \mid \big(\mathcal{T}_m = \mathcal{T}_m^*\big)\big].$$

For each $j_1, j_2, \dots, j_m \in \{0, 1\}$, for any distinct nodes $w_1, w_2 \in \overline{\mathcal{V}_m}$, events $(w_1 \in M_{j_1 j_2 \dots j_m})$ and $(w_2 \in M_{j_1 j_2 \dots j_m})$ are conditionally independent given $(\mathcal{T}_m = \mathcal{T}_m^*)$, where $\mathcal{T}_m^*$ specifies the key sets $S_1, S_2, \dots, S_m$ as $S_1^*, S_2^*, \dots, S_m^*$, respectively). Thus, with $\mathcal{M}_m^*$ being $\big(|M^*_{0^m}|, |M^*_{0^{m-1},1}|, |M^*_{0^{m-2}1,0}|, |M^*_{0^{m-2}1,1}|, \dots\big)$, we obtain

$$(22) = f(n-m, \mathcal{M}_m^*) \mathbb{P}[w \in M_{0^m} \,|\, \mathcal{T}_m = \mathcal{T}_m^*]^{|M^*_{0^m}|} \times \\ \prod_{\substack{j_1, j_2, \dots, j_m \in \{0,1\}: \\ \sum_{i=1}^m j_i \geq 1.}} \mathbb{P}[w \in M_{j_1 j_2 \dots j_m} \,|\, \mathcal{T}_m = \mathcal{T}_m^*]^{|M^*_{j_1 j_2 \dots j_m}|}, \tag{23}$$

where $f(n - m, \mathcal{M}_m^*)$ is the number of ways assigning the $(n - m)$ nodes from $\overline{\mathcal{V}_m}$ to $M_{j_1 j_2 \dots j_m}$ such that $|M_{j_1 j_2 \dots j_m}|$ equals $|M^*_{j_1 j_2 \dots j_m}|$, for $j_1, j_2, \dots, j_m \in \{0, 1\}$. Then

$$f(n-m, \mathcal{M}_m^*) = \frac{(n-m)!}{\prod_{j_1, j_2, \dots, j_m \in \{0,1\}} (|M^*_{j_1 j_2 \dots j_m}|!)}, \tag{24}$$

which along with (15) yields

$$f(n-m, \mathcal{M}_m^*) \leq [(n-m)!]/(|M^*_{0^m}|!) \\ \leq n^{\sum_{\substack{j_1, j_2, \dots, j_m \in \{0,1\}: \\ \sum_{i=1}^m j_i \geq 1.}} |M^*_{j_1 j_2 \dots j_m}|}. \tag{25}$$

For any $j_1, j_2, \dots, j_m \in \{0, 1\}$ with $\sum_{i=1}^m j_i \geq 1$, there exists $t \in \{0, 1, \dots, m\}$ such that $j_t = 1$, so

$$\mathbb{P}\big[w \in M_{j_1 j_2 \dots j_m} \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \\ \leq \mathbb{P}[E_{wv_t} \mid \mathcal{T}_m = \mathcal{T}_m^*] = \mathbb{P}[E_{wv_t}] = q_n, \tag{26}$$

where E_{wv_t} is the event that an edge exists between nodes w and v_t. Substituting (25) and (26) into (23), and denoting $\sum_{\substack{j_1, j_2, \dots, j_m \in \{0,1\}: \\ \sum_{i=1}^m j_i \geq 1.}} |M^*_{j_1 j_2 \dots j_m}|$ by Λ, we obtain

$$(22) < (nq_n)^{\Lambda} \times \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{|M^*_{0^m}|}. \tag{27}$$

To further evaluate (22) based on (27), we will prove below that if $\big(\mathcal{L}_m^* \notin \mathbb{L}_m^{(0)}\big)$ or $\big(\mathcal{M}_m^* \neq \mathcal{M}_m^{(0)}\big)$, then

$$\Lambda \leq hm - 1. \tag{28}$$

[1] For a non-negative integer x, the term 0^x is short for $\underbrace{00\dots0}_{\text{“}x\text{” number of “}0\text{”}}$. Also, for clarity, we add commas in the subscript of $M_{0^{m-2}1,0}$ etc.

On the one hand, if $\mathcal{L}_m^* \notin \mathbb{L}_m^{(0)}$, there exist i_1 and i_2 with $1 \leq i_1 < i_2 \leq m$ such that nodes v_{i_1} and v_{i_2} are neighbors. Hence, $\{v_{i_1}, v_{i_2}\} \subseteq [(\bigcup_{i=1}^m N_i) \bigcap \mathcal{V}_m]$ holds. Then from (16), we have $\Lambda = \big|\bigcup_{i=1}^m N_i\big| - \big|\big(\bigcup_{i=1}^m N_i\big) \cap \mathcal{V}_m\big| \leq hm - 2$. On the other hand, if $\mathcal{M}_m^* \neq \mathcal{M}_m^{(0)}$, there exist i_3 and i_4 with $1 \leq i_3 < i_4 \leq m$ such that $N_{i_3} \cap N_{i_4} \neq \emptyset$. Then from (16), $\Lambda \leq \big|\bigcup_{i=1}^m N_i\big| \leq \big(\sum_{i=1}^m |N_i|\big) - |N_{i_3} \cap N_{i_4}| \leq hm-1$ follows. Thus, we have proved (28), which along with (15) leads to

$$|M_{0^m}^*| = n - m - \Lambda > n - m - hm. \quad (29)$$

From (7), it is true that $nq_n \sim \ln n$, implying $nq_n > 1$ for all n sufficiently large. Then substituting (28) and (29) into (27), we obtain that if $\big(\mathcal{L}_m^* \notin \mathbb{L}_m^{(0)}\big)$ or $\big(\mathcal{M}_m^* \neq \mathcal{M}_m^{(0)}\big)$, then for all n sufficiently large, it holds that

$$(22) < (nq_n)^{hm-1} \times \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm}. \quad (30)$$

Applying (22) and (30) to (21),we get

$$(21) < \sum_{\mathcal{L}_m^* \in \mathbb{L}_m} \Big\{ |\mathbb{M}_m(\mathcal{L}_m^*)| \times \mathbb{P}\big[\mathcal{L}_m = \mathcal{L}_m^*\big] \times \text{R.H.S. of (30)} \Big\}. \quad (31)$$

To bound $|\mathbb{M}_m(\mathcal{L}_m^*)|$, note that $\mathcal{M}_m$ is a 2^m-tuple. Among the 2^m elements of the tuple, each of $|M_{j_1 j_2 \ldots j_m}|\big|_{j_1, j_2, \ldots, j_m \in \{0,1\}: \sum_{i=1}^m j_i \geq 1.}$ is at least 0 and at most h; and the remaining element $|M_{0^m}|$ can be determined by (15). Then it's straightforward that $|\mathbb{M}_m(\mathcal{L}_m^*)| \leq (h+1)^{2^m - 1}$. Using this result in (31), and considering $\big(\mathcal{L}_m = \mathcal{L}_m^*\big)$ is the union of independent events $\big(\mathcal{T}_m = \mathcal{T}_m^*\big)$ and $\big(\mathcal{C}_m = \mathcal{C}_m^*\big)$, and $\sum_{\mathcal{C}_m^* \in \mathbb{C}_m} \mathbb{P}\big[\mathcal{C}_m = \mathcal{C}_m^*\big] = 1$, we derive

$$(21) < (h+1)^{2^m-1} (nq_n)^{hm-1} \times \sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \Big\{ \mathbb{P}\big[\mathcal{T}_m = \mathcal{T}_m^*\big] \times \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \Big\}. \quad (32)$$

From (32) and $nq_n \sim \ln n \to \infty$ as $n \to \infty$ by (7), the proof of Proposition 1 is completed once we show

$$\sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \leq e^{-mnq_n} \cdot [1 + o(1)]. \quad (33)$$

C. Establishing (33)

From (61) and (62) (Lemma 4 in the Appendix), we get

$$\begin{aligned} &\mathbb{P}[w \in M_{0^m}^* \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \\ &= \mathbb{P}[w \in M_{0^m}^* \mid \mathcal{T}_m = \mathcal{T}_m^*]^n \mathbb{P}[w \in M_{0^m}^* \mid \mathcal{T}_m = \mathcal{T}_m^*]^{-m-hm} \\ &\leq e^{-mnq_n + m^2 n {q_n}^2 + \frac{nq_n p_n}{K_n} \sum_{1 \leq i < j \leq m} |S_{ij}^*|} (1 - mq_n)^{-m-hm} \quad (34) \end{aligned}$$

for all n sufficiently large, where $S_{ij}^* := S_i^* \cap S_j^*$. With (7) (i.e., $q_n \sim \frac{\ln n}{n}$), we have $m^2 n {q_n}^2 = o(1)$ and $mq_n = o(1)$, which are substituted into (34) to induce (33) once we prove

$$\sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] e^{\frac{nq_n p_n}{K_n} \sum_{1 \leq i < j \leq m} |S_{ij}^*|} \leq 1 + o(1). \quad (35)$$

L.H.S. of (35) is denoted by $H_{n,m}$ and evaluated below. For each fixed and sufficiently large n, we consider: a) $p_n < n^{-\delta}(\ln n)^{-1}$ and b) $p_n \geq n^{-\delta}(\ln n)^{-1}$, where δ is an arbitrary constant with $0 < \delta < 1$.

a) $p_n < n^{-\delta}(\ln n)^{-1}$

From $p_n < n^{-\delta}(\ln n)^{-1}$, $|S_{ij}^*| \leq K_n$ for $1 \leq i < j \leq m$ and (20), then for all n sufficiently large, it holds that $e^{\frac{nq_n p_n}{K_n} \sum_{1 \leq i < j \leq m} |S_{ij}^*|} < e^{2n^{-\delta} \cdot \binom{m}{2}} < e^{m^2 n^{-\delta}}$, which is used in $H_{n,m}$ so that $H_{n,m} < e^{m^2 n^{-\delta}} \sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] = e^{m^2 n^{-\delta}}$.

b) $p_n \geq n^{-\delta}(\ln n)^{-1}$

We relate $H_{n,m}$ to $H_{n,m-1}$ and assess $H_{n,m}$ iteratively. First, with $\mathcal{T}_m^* = (S_1^*, S_2^*, \ldots, S_m^*)$, event $(\mathcal{T}_m = \mathcal{T}_m^*)$ is the intersection of independent events: $(\mathcal{T}_{m-1} = \mathcal{T}_{m-1}^*)$ and $(S_m = S_m^*)$. Then we have

$$\begin{aligned} H_{n,m} &= \sum_{\substack{\mathcal{T}_{m-1}^* \in \mathbb{T}_{m-1}, \\ S_m^* \in \mathbb{S}_m}} \Big(\mathbb{P}[(\mathcal{T}_{m-1} = \mathcal{T}_{m-1}^*) \cap (S_m = S_m^*)] \times e^{\frac{nq_n p_n}{K_n} \sum_{1 \leq i < j \leq m-1} |S_{ij}^*|} e^{\frac{nq_n p_n}{K_n} \sum_{i=1}^{m-1} |S_{im}^*|} \Big) \\ &= H_{n,m-1} \cdot \sum_{S_m^* \in \mathbb{S}_m} \mathbb{P}[S_m = S_m^*] e^{\frac{nq_n p_n}{K_n} \sum_{i=1}^{m-1} |S_{im}^*|}. \quad (36) \end{aligned}$$

By $\sum_{i=1}^{m-1} |S_{im}^*| \leq m \big|S_m^* \cap \big(\bigcup_{i=1}^{m-1} S_i^*\big)\big|$ and (20), we have $e^{\frac{nq_n p_n}{K_n} \sum_{i=1}^{m-1} |S_{im}^*|} \leq e^{\frac{2mp_n \ln n}{K_n} |S_m^* \cap (\bigcup_{i=1}^{m-1} S_i^*)|}$, which is used in (36) to induce

$$\frac{H_{n,m}}{H_{n,m-1}} \leq \sum_{u=0}^{K_n} \mathbb{P}\bigg[\bigg| S_m^* \cap \bigg(\bigcup_{i=1}^{m-1} S_i^* \bigg) \bigg| = u \bigg] e^{\frac{2ump_n \ln n}{K_n}}. \quad (37)$$

Denoting $\big|\bigcup_{i=1}^{m-1} S_i^*\big|$ by v, then for u satisfying $0 \leq u \leq |S_m^*| = K_n$ and $S_m^* \cup \big(\bigcup_{i=1}^{m-1} S_i^*\big) = K_n + v - u \leq P_n$ (i.e., for $u \in [\max\{0, K_n + v - P_n\}, K_n]$), we obtain

$$\mathbb{P}\bigg[\bigg| S_m^* \cap \bigg(\bigcup_{i=1}^{m-1} S_i^* \bigg) \bigg| = u \bigg] = \binom{v}{u} \binom{P_n - v}{K_n - u} \Big/ \binom{P_n}{K_n}, \quad (38)$$

which together with $K_n \leq v \leq mK_n$ yields

$$\begin{aligned} \text{L.H.S. of (38)} &\leq \frac{(mK_n)^u}{u!} \cdot \frac{(P_n - K_n)^{K_n - u}}{(K_n - u)!} \cdot \frac{K_n!}{(P_n - K_n)^{K_n}} \\ &\leq \frac{1}{u!} \bigg(\frac{m{K_n}^2}{P_n - K_n} \bigg)^u. \quad (39) \end{aligned}$$

For $u \notin [\max\{0, K_n + v - P_n\}, K_n]$, L.H.S. of (38) equals 0. Then from (37) and (39),

$$\begin{aligned} \text{R.H.S. of (37)} &\leq \sum_{u=0}^{K_n} \frac{1}{u!} \bigg(\frac{m{K_n}^2}{P_n - K_n} \cdot e^{\frac{2mp_n \ln n}{K_n}} \bigg)^u \\ &\leq e^{\frac{m{K_n}^2}{P_n - K_n} \cdot e^{\frac{2mp_n \ln n}{K_n}}}. \quad (40) \end{aligned}$$

By [15, Fact 5] and $1 - x \leq e^{-x}$ for any real x, it holds that

$$s_n \geq 1 - \big(1 - K_n/P_n\big)^{K_n} \geq 1 - e^{-{K_n}^2/P_n}, \quad (41)$$

For n sufficiently large, from $p_n \geq n^{-\delta}(\ln n)^{-1}$ and (20) (i.e., $q_n = p_n s_n \leq \frac{2 \ln n}{n}$), we have

$$s_n = {p_n}^{-1} q_n \leq {p_n}^{-1} \cdot 2n^{-1} \ln n \leq 2n^{\delta - 1}(\ln n)^2. \quad (42)$$

Hence, for n sufficiently large, we apply (41) (42) and $P_n > 2K_n$ (which holds from the condition $\frac{K_n}{P_n} = o(1)$) to produce

$$\begin{aligned} {K_n}^2/(P_n - K_n) &< 2{K_n}^2/P_n \leq -2\ln(1 - s_n) \\ &\leq -2\ln(1 - 2n^{\delta-1}(\ln n)^2) \leq 2\sqrt{2} n^{\frac{\delta - 1}{2}} \ln n, \quad (43) \end{aligned}$$

where the last step uses $-\ln(1-y) \leq \sqrt{y}$ for $0 < y < 1$. From (7) and condition $P_n = \Omega(n)$, we obtain from [15, Lemma 7] that $K_n = \omega\big(\sqrt{\ln n}\big) = \omega(1)$. Then for an arbitrary constant $c > 2$, it holds that $\frac{K_n}{p_n} \geq K_n \geq \frac{4c \cdot m}{(c-2)(1-\delta)}$ holds for all n sufficiently large. Hence,

$$e^{\frac{2mp_n \ln n}{K_n}} \leq e^{\frac{(c-2)(1-\delta)}{2c} \ln n} = n^{\frac{(c-2)(1-\delta)}{2c}}. \quad (44)$$

The use of (40) (43) and (44) in (37) yields

$$\begin{aligned} &H_{n,m}/H_{n,m-1} \leq \text{R.H.S. of (37)} \\ &\leq e^{2\sqrt{2} m n^{\frac{\delta-1}{2}} \cdot n^{\frac{(c-2)(1-\delta)}{2c}} \cdot \ln n} \leq \Big(e^{3n^{\frac{\delta-1}{c}} \ln n} \Big)^m. \quad (45) \end{aligned}$$

To derive $H_{n,m}$ iteratively based on (45), we compute $H_{n,2}$ below. Setting $m = 2$ in L.H.S. of (35) and considering the independence between $(S_1 = S_1^*)$ and $(S_2 = S_2^*)$, we gain

$$H_{n,2} = \sum_{S_1^* \in \mathbb{S}_m} \mathbb{P}[S_1 = S_1^*] \sum_{S_2^* \in \mathbb{S}_m} \mathbb{P}[S_2 = S_2^*] e^{\frac{nq_n p_n}{K_n}|S_1^* \cap S_2^*|}. \quad (46)$$

Clearly, $\sum_{S_2^* \in \mathbb{S}_m} \mathbb{P}[S_2 = S_2^*] e^{\frac{nq_n p_n}{K_n}|S_1^* \cap S_2^*|}$ equals R.H.S. of (37) with $m = 2$. Then from (45) and (46),

$$H_{n,2} \leq \sum_{S_1^* \in \mathbb{S}_m} \mathbb{P}[S_1 = S_1^*] e^{6n^{\frac{\delta-1}{c}} \ln n} = e^{6n^{\frac{\delta-1}{c}} \ln n}. \quad (47)$$

Therefore, it holds via (45) and (47) that

$$H_{n,m} \leq \left(e^{3n^{\frac{\delta-1}{c}} \ln n}\right)^{m+(m-1)+\ldots+3} e^{6n^{\frac{\delta-1}{c}} \ln n} \leq e^{3m^2 n^{\frac{\delta-1}{c}} \ln n}.$$

Finally, from cases a) and b), for n sufficiently large, $H_{n,m}$ is at most $\max\left\{e^{m^2 n^{-\delta}},\, e^{3m^2 n^{\frac{\delta-1}{c}} \ln n}\right\}$. Then (35) follows.

V. The Proof of Proposition 2

We define $\mathcal{C}_m^{(0)}$ and $\mathbb{T}_m^{(0)}$ by $\mathcal{C}_m^{(0)} = (\underbrace{0, 0, \ldots, 0}_{m(m-1)/2 \text{ number of “0”}})$ and $\mathbb{T}_m^{(0)} = \{\mathcal{T}_m \mid S_i \cap S_j = \emptyset,\ \forall 1 \leq i < j \leq m.\}$. Clearly, $(\mathcal{C}_m = \mathcal{C}_m^{(0)})$ or $(\mathcal{T}_m \in \mathbb{T}_m^{(0)})$ each implies $(\mathcal{L}_m \in \mathbb{L}_m^{(0)})$. Also, $(\mathcal{C}_m = \mathcal{C}_m^{(0)})$ and $(\mathcal{M}_m = \mathcal{M}_m^{(0)})$ are independent of each other. Thus, with $\mathcal{P}_2 = \mathbb{P}\big[(\mathcal{L}_m \in \mathbb{L}_m^{(0)}) \cap (\mathcal{M}_m = \mathcal{M}_m^{(0)})\big]$, we derive

$$\mathcal{P}_2 \geq \mathbb{P}\big[\mathcal{C}_m = \mathcal{C}_m^{(0)}\big] \mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big], \quad (48)$$

and

$$\mathcal{P}_2 \geq \mathbb{P}\big[\mathcal{T}_m \in \mathbb{T}_m^{(0)}\big] \mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \,|\, (\mathcal{T}_m \in \mathbb{T}_m^{(0)})\big]. \quad (49)$$

Given that event $(\mathcal{C}_m = \mathcal{C}_m^{(0)})$ is $\overline{\bigcup_{1 \leq i < j \leq m} C_{ij}}$ and event $(\mathcal{T}_m \in \mathbb{T}_m^{(0)})$ is $\overline{\bigcup_{1 \leq i < j \leq m} \Gamma_{ij}}$, using the union bound, we get

$$\mathbb{P}\big[\mathcal{C}_m = \mathcal{C}_m^{(0)}\big] \geq 1 - \sum_{1 \leq i < j \leq m} \mathbb{P}[C_{ij}] \geq 1 - m^2 p_n/2, \quad (50)$$

and

$$\mathbb{P}\big[\mathcal{T}_m \in \mathbb{T}_m^{(0)}\big] \geq 1 - \sum_{1 \leq i < j \leq m} \mathbb{P}[\Gamma_{ij}] \geq 1 - m^2 s_n/2. \quad (51)$$

Denoting $(h!)^{-m}(nq_n)^{hm} e^{-mnq_n}$ by Λ, we will prove

$$\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big] \sim \Lambda, \quad (52)$$

and

$$\mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \mid (\mathcal{T}_m \in \mathbb{T}_m^{(0)})\big] \geq \Lambda \cdot [1 - o(1)]. \quad (53)$$

Substituting (50) and (52) into (48), and applying (51) and (53) to (49), we get (i) $\mathcal{P}_2/\Lambda \geq (1 - \min\{s_n, p_n\} \cdot m^2/2)[1 - o(1)]$. From (52), we get (ii) $\mathcal{P}_2 \leq \mathbb{P}\big[\mathcal{M}_m \in \mathbb{M}_m^{(0)}\big] \leq \Lambda[1 + o(1)]$. Combining (i) and (ii) above and using $\min\{s_n, p_n\} \leq \sqrt{s_n p_n} = \sqrt{q_n} = o(1)$ which holds from $q_n = s_n p_n$ and (7), Proposition 2 follows. Below we establish (52) and (53).

A. Establishing (52)

We write $\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big]$ as

$$\sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \Big\{\mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \mid (\mathcal{T}_m = \mathcal{T}_m^*)\big]\Big\},$$

where $\mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \mid (\mathcal{T}_m = \mathcal{T}_m^*)\big]$ equals

$$f\big(n - m, \mathcal{M}_m^{(0)}\big) \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \times \prod_{i=1}^{m} \mathbb{P}[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*]^h,$$

where $f\big(n - m, \mathcal{M}_m^{(0)}\big)$ is the number of ways assigning the $(n - m)$ nodes from $\overline{\mathcal{V}_m}$ to $M_{j_1 j_2 \ldots j_m}$ such that $|M_{j_1 j_2 \ldots j_m}|$ is given by $\mathcal{M}_m^{(0)}$ (see (17)). Hence, it holds from (24) that

$$f\big(n - m, \mathcal{M}_m^{(0)}\big) = \frac{(n - m)!}{(n - m - hm)!(h!)^m} \sim (h!)^{-m} n^{hm}. \quad (54)$$

We will establish

$$\sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \Big\{\mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \prod_{i=1}^{m} \{\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \,|\, \mathcal{T}_m = \mathcal{T}_m^*\big]^h\}\Big\} \geq {q_n}^{hm} \cdot [1 - o(1)]. \quad (55)$$

We use (54) and (55) as well as (61) (viz., Lemma 4 in the Appendix) in evaluating $\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big]$ above. Then

$$\begin{aligned}
&\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big] \\
&\geq (h!)^{-m} n^{hm} \cdot [1 - o(1)] \cdot (1 - mq_n)^n \times \\
&\quad \sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \prod_{i=1}^{m} \{\mathbb{P}[w \in M_{0^{i-1},1,0^{m-i}} \,|\, \mathcal{T}_m = \mathcal{T}_m^*]^h\} \\
&\geq (h!)^{-m} (nq_n)^{hm} e^{-mnq_n} \cdot [1 - o(1)]. \quad (56)
\end{aligned}$$

Substituting (33) (54) above and (63) in Lemma 4 into the computation of $\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big]$ yields

$$\begin{aligned}
&\mathbb{P}\big[\mathcal{M}_m = \mathcal{M}_m^{(0)}\big] \\
&\leq (h!)^{-m} n^{hm} {q_n}^{hm} \times [1 + o(1)] \times \\
&\quad \sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \\
&\sim (h!)^{-m} (nq_n)^{hm} e^{-mnq_n}. \quad (57)
\end{aligned}$$

Then (52) follows from (56) and (57). Namely, (52) holds upon the establishment of (55). From (64) in Lemma 4 and $q_n = o(1)$ by (7), we obtain (55) once proving

$$\frac{p_n}{K_n} \sum_{\mathcal{T}_m^* \in \mathbb{T}_m} \Bigg(\mathbb{P}[\mathcal{T}_m = \mathcal{T}_m^*] \sum_{1 \leq i < j \leq m} |S_{ij}^*|\Bigg) = o(1). \quad (58)$$

If $\mathcal{T}_m^* \in \mathbb{T}_m^{(0)}$, then $|S_{ij}^*| = 0$. Then from (51), we get (58) by

$$\begin{aligned}
\text{L.H.S. of (58)} &\leq p_n \cdot m(m-1)/2 \cdot \mathbb{P}[\mathcal{T}_m^* \in \mathbb{T}_m \setminus \mathbb{T}_m^{(0)}] \\
&\leq p_n \cdot m^2/2 \cdot m^2 s_n/2 \leq m^4 n^{-1} \ln n/2 = o(1).
\end{aligned}$$

B. Establishing (53)

Let Δ denote $\mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \mid (\mathcal{T}_m \in \mathbb{T}_m^{(0)})\big]$. Clearly, Δ is equivalent to $\mathbb{P}\big[(\mathcal{M}_m = \mathcal{M}_m^{(0)}) \mid (\mathcal{T}_m = \mathcal{T}_m^*)\big]$ for any $\mathcal{T}_m^* \in \mathbb{T}_m^{(0)}$, so it follows that

$$\Delta = f\big(n - m, \mathcal{M}_m^{(0)}\big) \mathbb{P}[w \in M_{0^m} \,|\, \mathcal{T}_m = \mathcal{T}_m^*]^{n-m-hm} \times \prod_{i=1}^{m} \{\mathbb{P}[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*]^h\}, \quad (59)$$

with $f\big(n - m, \mathcal{M}_m^{(0)}\big)$ given by (54). For $\mathcal{T}_m^* \in \mathbb{T}_m^{(0)}$, from $|S_{ij}^*| = 0$ and (64) in Lemma 4, we derive

$$\prod_{i=1}^{m} \Big\{\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \,|\, \mathcal{T}_m = \mathcal{T}_m^*\big]\Big\}^h \geq {q_n}^{hm}(1 - 2hm^2 q_n). \quad (60)$$

Substituting (54) (60) above and (61) in Lemma 4 into (59), we conclude that Δ is at least

$$\begin{aligned}
&(h!)^{-m} n^{hm} \cdot [1 - o(1)] \\
&\times {q_n}^{hm}(1 - 2hm^2 q_n) \cdot (1 - mq_n)^{n-m-hm} = \Lambda \cdot [1 - o(1)].
\end{aligned}$$

VI. Numerical Experiments

To confirm our analytical results, we now provide numerical experiments in the non-asymptotic regime.

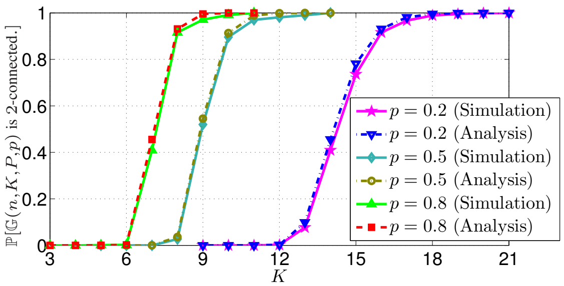

Fig. 1. A plot generated from the simulation and the analysis for the probability that $\mathbb{G}(n, K, P, p)$ is 2-connected versus K with $n = 2,000$, $P = 10,000$ and $p = 0.2, 0.5, 0.8$.

In Figure 1, we depict the probability that graph $\mathbb{G}(n, K, P, p)$ is 2-connected from both the simulation and the analysis, as elaborated below. In all set of experiments, we fix the number of nodes at $n = 2,000$ and the key pool size at $P = 10,000$. For the probability p of a communication channel being *on*, we consider $p = 0.2, 0.5, 0.8$, while varying the parameter K from 3 to 21. For each pair (K, p), we generate $1,000$ independent samples of $\mathbb{G}(n, K, P, p)$ and count the number of times that the obtained graphs are 2-connected. Then the counts divided by $1,000$ become the empirical probabilities.

The curves in Figure 1 corresponding to the analysis are determined as follows. We use the asymptotical result to approximate the probability of 2-connectivity in $\mathbb{G}(n, K, P, p)$; specifically, given n, K, P, p and $k = 2$, we determine α by considering $p \cdot \left[1 - \binom{P-K}{K} \big/ \binom{P}{K}\right] = \frac{\ln n + (k-1) \ln \ln n + \alpha}{n}$, a condition stemming from (4) and the computation of q_n in Section II, and then use $e^{-\frac{e^{-\alpha}}{(k-1)!}}$ as the analytical reference of $\mathbb{P}[\mathbb{G}(n, K, P, p)$ is 2-connected$]$ for a comparison with the empirical probabilities. Figure 1 indicates that the experimental results are in agreement with our analysis.

VII. Related Work

Random key graphs. For a random key graph $G(n, K_n, P_n)$ (viz., Section II) which models the topology induced by the EG scheme, Rybarczyk [7] derives the asymptotically exact probability of connectivity, covering a weaker form of the result – a zero-one law which is also obtained in [1], [10]. Rybarczyk [8] further establishes a zero-one law for k-connectivity, and we [14] obtain the asymptotically exact probability of k-connectivity. Under $P_n = \Theta(n^c)$ for some constant $c > 1$ and $\frac{{K_n}^2}{P_n} = \frac{\ln n + (k-1) \ln \ln n + \alpha_n}{n}$, Rybarczyk's result [8] is that the probability of k-connectivity in graph $G(n, K_n, P_n)$ is asymptotically converges to 1 (resp. 0) if $\lim_{n \to \infty} \alpha_n$ equals ∞ (resp., $-\infty$), while we [14] prove that such probability asymptotically approaches to $e^{-\frac{e^{-\alpha^*}}{(k-1)!}}$ if $\lim_{n \to \infty} \alpha_n = \alpha^* \in (-\infty, \infty)$.

Erdős–Rényi graphs. For an Erdős–Rényi graph $G(n, p_n)$ where any two nodes have an edge in between independently with probability p_n, Erdős and Rényi consider connectivity in [2] and k-connectivity in [3], where the latter result is that if $p_n = \frac{\ln n + (k-1) \ln \ln n + \alpha_n}{n}$ and $\lim_{n \to \infty} \alpha_n = \alpha^* \in [-\infty, \infty]$, graph $G(n, p_n)$ is k-connected with a probability asymptotically tending to $e^{-\frac{e^{-\alpha^*}}{(k-1)!}}$.

Random key graphs $\cap$ Erdős–Rényi graphs. As given in Section II, our studied graph $\mathbb{G}$ is the intersection of a random key graph $G(n, K_n, P_n)$ and an Erdős–Rényi graph $G(n, p_n)$. For graph $\mathbb{G}$, Yağan [9] establishes a zero-one law for connectivity, and we [15], [12] extend Yağan's result to k-connectivity and show that with $P_n = \Omega(n)$, $\frac{K_n}{P_n} = o(1)$ and q_n set as $\frac{\ln n + (k-1) \ln \ln n + \alpha_n}{n}$, graph $\mathbb{G}$ is (resp., is not) k-connected with high probability if $\lim_{n \to \infty} \alpha_n = \infty$ (resp., $\lim_{n \to \infty} \alpha_n = -\infty$). Compared with this result in [15], [12], our result on the asymptotically exact probability of k-connectivity is stronger and more challenging to derive.

Random key graphs $\cap$ random geometric graphs. Connectivity properties have also been studied in secure sensor networks employing the EG scheme under the disk model, where any two nodes need to be within a certain distance r_n to have a link in between. When nodes are assumed to be uniformly and independently deployed in some region $\mathcal{A}$, the topology of such a network is represented by the intersection of a random key graph $G(n, K_n, P_n)$ and a random geometric graph, where a random geometric graph denoted by $G(n, r_n, \mathcal{A})$ is defined on n nodes independently and uniformly distributed in $\mathcal{A}$ such that an edge exists between two nodes if and only if their distance is at most r_n. Krzywdziński and Rybarczyk [6], Krishnan *et al.* [5], and we [13] present connectivity results in graph $G(n, K_n, P_n) \cap G(n, r_n, \mathcal{A})$. With the network region $\mathcal{A}$ being a square of unit area, Krzywdziński and Rybarczyk [6] show that $G(n, K_n, P_n) \cap G(n, r_n, \mathcal{A})$ is connected with high probability if $\pi {r_n}^2 \cdot \frac{{K_n}^2}{P_n} \sim \frac{c \ln n}{n}$ for any constant $c > 8$. Krishnan *et al.* [5] improves the condition on c to $c > 2\pi$. Later we [13] derive the critical value c^* of c as $\max\{1 + \lim_{n \to \infty} \left(\ln \frac{P_n}{{K_n}^2} / \ln n\right), 4 \lim_{n \to \infty} \left(\ln \frac{P_n}{{K_n}^2} / \ln n\right)\}$; namely, graph $G(n, K_n, P_n) \cap G(n, r_n, \mathcal{A})$ is (resp., is not) connected with high probability for any constant $c > c^*$ (resp., $c < c^*$). There has not been any analogous result for k-connectivity reported in the literature.

VIII. Conclusion and Future Work

In this paper, we consider secure WSNs under the Eschenauer–Gligor (EG) key predistribution scheme with unreliable links and obtain the asymptotically exact probability of k-connectivity. A future direction is to consider k-connectivity in WSNs employing the EG scheme under the disk model [9], [5] in which two nodes have to be within a certain distance for communication in addition to sharing at least one key.

References

[1] S. R. Blackburn and S. Gerke. Connectivity of the uniform random intersection graph. *Discrete Mathematics*, 309(16), August 2009.

[2] P. Erdős and A. Rényi. On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.

[3] P. Erdős and A. Rényi. On the strength of connectedness of random graphs. *Acta Math. Acad. Sci. Hungar.*, pages 261–267, 1961.

[4] L. Eschenauer and V. Gligor. A key-management scheme for distributed sensor networks. In *ACM Conference on Computer and Communications Security (CCS)*, 2002.

[5] B. Krishnan, A. Ganesh, and D. Manjunath. On connectivity thresholds in superposition of random key graphs on random geometric graphs. In *IEEE International Symposium on Information Theory (ISIT)*, 2013.

[6] K. Krzywdziński and K. Rybarczyk. Geometric graphs with randomly deleted edges — connectivity and routing protocols. *Mathematical Foundations of Computer Science*, 6907:544–555, 2011.

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

[8] K. Rybarczyk. Sharp threshold functions for the random intersection graph via a coupling method. *The Electronic Journal of Combinatorics*, 18:36–47, 2011.

[9] O. Yağan. Performance of the Eschenauer-Gligor key distribution scheme under an on/off channel. *IEEE Transactions on Information Theory*, 58(6):3821–3835, 2012.

[10] O. Yağan and A. M. Makowski. Zero-one laws for connectivity in random key graphs. *IEEE Transactions on Information Theory*, 58(5):2983–2999, May 2012.
[11] 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 *IEEE Global Telecommunications Conference (GLOBECOM)*, 2006.
[12] J. Zhao, O. Yağan, and V. Gligor. Secure k-connectivity in wireless sensor networks under an on/off channel model. In *IEEE International Symposium on Information Theory (ISIT)*, 2013.
[13] J. Zhao, O. Yağan, and V. Gligor. Connectivity in secure wireless sensor networks under transmission constraints. In *Allerton Conference on Communication, Control, and Computing*, 2014.
[14] J. Zhao, O. Yağan, and V. Gligor. On the strengths of connectivity and robustness in general random intersection graphs. In *IEEE Conference on Decision and Control (CDC)*, 2014.
[15] J. Zhao, O. Yağan, and V. Gligor. k-Connectivity in random key graphs with unreliable links. *IEEE Transactions on Information Theory*, 2015.

APPENDIX

A. Useful Lemmas

We present below Lemmas 4 and 5, which are proved in the next subsections. Lemma 4 is used in establishing Propositions 1 and 2 in Section IV-B. The condition $P_n \geq 3K_n$ in Lemma 4 follows for all n sufficiently large given $K_n/P_n = o(1)$ in Propositions 1 and 2. Lemma 5 is used in proving Lemma 4.

Lemma 4. *Given $P_n \geq 3K_n$ and any $\mathcal{T}_m^* = (S_1^*, S_2^*, \ldots, S_m^*)$, with S_{ij}^* denoting $S_i^* \cap S_j^*$, for any node $w \in \overline{\mathcal{V}_m}$, we obtain*

$$\mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*] \geq 1 - mq_n, \quad \textit{and} \tag{61}$$

$$\mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*] \leq e^{-mq_n + m^2{q_n}^2 + {K_n}^{-1} q_n p_n \sum_{1 \leq i < j \leq m} |S_{ij}^*|}; \tag{62}$$

and for any $i = 1, 2, \ldots, m$, we have

$$\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \leq q_n, \quad \textit{and} \tag{63}$$

$$\prod_{i=1}^m \left\{\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big]\right\}^h \geq {q_n}^{hm}\big(1 - 2hm^2 q_n - \tfrac{2hp_n}{K_n} \textstyle\sum_{1 \leq i < j \leq m} |S_{ij}^*|\big). \tag{64}$$

Lemma 5. *With Γ_{ij} denoting the event that an edge exists between distinct nodes v_i and v_j in random key graph $G(n, K_n, P_n)$, if $P_n \geq 3K_n$, then for three distinct nodes v_i, v_j and v_t, we have $\mathbb{P}[(\Gamma_{it} \cap \Gamma_{jt} \mid (|S_{ij}| = u)] \leq {K_n}^{-1} s_n u + 2{s_n}^2$ for $u = 0, 1, \ldots, K_n$.*

B. The Proof of Lemma 4

For any node $w \in \overline{\mathcal{V}_m}$, event $(w \in M_{0^m})$ equals $\overline{\bigcup_{i=1}^m E_{wv_i}}$, where E_{wv_i} is the event that there exists an edge between nodes w and v_i in $\mathbb{G}$. By a union bound, L.H.S. of (61) is at least $1 - \sum_{i=1}^m \mathbb{P}[E_{wv_i} \mid \mathcal{T}_m = \mathcal{T}_m^*] = 1 - mq_n$ so that (61) is proved. And to prove (62), by the inclusion–exclusion principle, we get

$$\mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*] \leq 1 - \sum_{i=1}^m \mathbb{P}[E_{wv_i} \mid \mathcal{T}_m = \mathcal{T}_m^*] + \sum_{1 \leq i < j \leq m} \mathbb{P}[E_{wv_i} \cap E_{wv_j} \mid \mathcal{T}_m = \mathcal{T}_m^*].$$

Then we use Lemma 5 to further derive

$$\begin{aligned}
&\mathbb{P}[w \in M_{0^m} \mid \mathcal{T}_m = \mathcal{T}_m^*] \\
&\leq 1 - mq_n + {p_n}^2 \sum_{1 \leq i < j \leq m} \big({K_n}^{-1} s_n |S_{ij}^*| + 2{s_n}^2\big) \\
&\leq e^{-mq_n + m^2{q_n}^2 + {K_n}^{-1} q_n p_n \sum_{1 \leq i < j \leq m} |S_{ij}^*|},
\end{aligned}$$

where the last step uses $1 + x \leq e^x$ for any real x.

For any node $w \in \overline{\mathcal{V}_m}$, event $w \in M_{0^{i-1},1,0^{m-i}}$ means that node w has an edge with node v_i, but has no edge with any node in $\mathcal{V}_m \setminus \{v_i\} = \{v_j \mid j \in \{1, 2, \ldots, m\} \setminus \{i\}\}$. Then (63) follows since $\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big]$ is at most $\mathbb{P}[E_{wv_i} \mid \mathcal{T}_m = \mathcal{T}_m^*] = \mathbb{P}[E_{wv_i}] = q_n$. where the last step uses the independence between event E_{wv_i} and event $(\mathcal{T}_m = \mathcal{T}_m^*)$.

We now demonstrate (64). From the above, we have

$$\begin{aligned}
&\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \\
&= \mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcap_{j \in \{1,2,\ldots,m\} \setminus \{i\}} \overline{E_{wv_j}}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \\
&= \mathbb{P}[E_{wv_i}] - \mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big],
\end{aligned} \tag{65}$$

where the last step uses $\mathbb{P}[E_{wv_i} \mid \mathcal{T}_m = \mathcal{T}_m^*] = \mathbb{P}[E_{wv_i}]$ since event E_{wv_i} is independent of event $(\mathcal{T}_m = \mathcal{T}_m^*)$.

From (65) and $\mathbb{P}[E_{wv_i}] = q_n$, we obtain

$$\begin{aligned}
&{q_n}^{-1} \mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \\
&= 1 - {q_n}^{-1} \mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big],
\end{aligned}$$

so that

$$\begin{aligned}
&{q_n}^{-hm} \cdot \prod_{i=1}^m \left\{\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}}^{(0)} \mid \mathcal{T}_m = \mathcal{T}_m^*\big]\right\}^h \\
&= \prod_{i=1}^m \left\{1 - {q_n}^{-1} \mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big]\right\}^h \\
&\geq 1 - h \sum_{i=1}^m \left\{{q_n}^{-1} \mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big]\right\},
\end{aligned} \tag{66}$$

where the last step uses the following inequality easily proved by mathematical induction: $\prod_{\ell=1}^r (1 - x_\ell) \geq 1 - \sum_{\ell=1}^r x_\ell$ for any positive integer r and any x_ℓ with $0 \leq x_\ell \leq 1$ for $\ell = 1, 2, \ldots, r$ (we set $r = mh$, with the mh number of x_l as m groups, where the group i for $i = 1, 2, \ldots, m$ has m members all being ${q_n}^{-1} \mathbb{P}[E_{wv_i} \cap (\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}) \mid \mathcal{T}_m = \mathcal{T}_m^*]$.)
To analyze (66), we use the union bound and Lemma 5 to get

$$\begin{aligned}
&\mathbb{P}\big[E_{wv_i} \cap \big(\textstyle\bigcup_{j \in \{1,2,\ldots,m\} \setminus \{i\}} E_{wv_j}\big) \mid \mathcal{T}_m = \mathcal{T}_m^*\big] \\
&\leq \textstyle\sum_{j \in \{1,2,\ldots,m\} \setminus \{i\}} \mathbb{P}[E_{wv_i} \cap E_{wv_j} \mid \mathcal{T}_m = \mathcal{T}_m^*] \\
&\leq \textstyle\sum_{j \in \{1,2,\ldots,m\} \setminus \{i\}} {p_n}^2 \big({K_n}^{-1} s_n |S_{ij}^*| + 2{s_n}^2\big) \\
&\leq 2m{q_n}^2 + {K_n}^{-1} p_n q_n \textstyle\sum_{j \in \{1,2,\ldots,m\} \setminus \{i\}} |S_{ij}^*|,
\end{aligned}$$

which is substituted into (66) to establish (64) by

$$\begin{aligned}
&{q_n}^{-hm} \cdot \textstyle\prod_{i=1}^m \left\{\mathbb{P}\big[w \in M_{0^{i-1},1,0^{m-i}} \mid \mathcal{T}_m = \mathcal{T}_m^*\big]\right\}^h \\
&\geq 1 - h \textstyle\sum_{i=1}^m \left\{2mq_n + {K_n}^{-1} p_n \sum_{j \in \{1,2,\ldots,m\} \setminus \{i\}} |S_{ij}^*|\right\} \\
&\geq 1 - 2hm^2 q_n - \tfrac{2hp_n}{K_n} \textstyle\sum_{1 \leq i < j \leq m} |S_{ij}^*|.
\end{aligned} \tag{67}$$

C. The Proof of Lemma 5

We use the inclusion–exclusion principle to obtain

$$\begin{aligned}
&\mathbb{P}[\Gamma_{it} \cap \Gamma_{jt} \mid (|S_{ij}| = u)] \\
&= \mathbb{P}[\Gamma_{it} \mid (|S_{ij}| = u)] + \mathbb{P}[\Gamma_{jt} \mid (|S_{ij}| = u)] \\
&\quad - \mathbb{P}[\Gamma_{it} \cup \Gamma_{jt} \mid (|S_{ij}| = u)] \\
&= 2s_n - 1 + \tbinom{P_n - (2K_n - u)}{K_n} \big/ \tbinom{P_n}{K_n},
\end{aligned} \tag{68}$$

in view that event $(|S_{ij}| = u)$ is independent of each of Γ_{it} and Γ_{jt}, and event $\Gamma_{it} \cup \Gamma_{jt}$ means $S_t \cap (S_i \cup S_j) \neq \emptyset$.

By [9, Lemma 5.1] and [15, Fact 2], we derive

$$\begin{aligned}
(1 - s_n)^{\frac{2K_n - u}{K_n}} &\leq 1 - \tfrac{s_n(2K_n - u)}{K_n} + \tfrac{1}{2}\big(\tfrac{s_n(2K_n - u)}{K_n}\big)^2 \\
&\leq 1 - 2s_n + {K_n}^{-1} s_n u + 2{s_n}^2,
\end{aligned}$$

which is substituted into (68) to complete the proof.