On existence of triangles, clustering, and small-world property of random key graphs

Osman Yağan, *Member, IEEE* and Armand M. Makowski, *Fellow, IEEE*

***Abstract*—Random key graphs have been introduced about a decade ago in the context of key predistribution schemes for securing wireless sensor networks (WSNs). They have received much attention recently with applications spanning the areas of recommender systems, epidemics in social networks, cryptanalysis, and clustering and classification analysis. This paper is devoted to analyzing various properties of random key graphs including containment and number of triangles, clustering coefficient, and *small-world* behavior. In particular, we establish a zero-one law for the existence of triangles in random key graphs, and identify the corresponding critical scaling. This zero-one law exhibits significant differences with the corresponding result in Erdős-Rényi (ER) graphs. We also compute the clustering coefficient of random key graphs, and compare them with that of ER graphs in the many node regime when the expected average degrees are asymptotically equivalent. We show that on the parameter range of practical relevance in both wireless sensor network and social network applications, random key graphs are much more clustered than the corresponding ER graph. The suitability of random key graphs as small worlds in the sense of Watts and Strogatz is also demonstrated.**

Keywords: Wireless sensor networks, Key predistribution, Random graphs, Existence of triangles, Clustering coefficient.

I. Introduction

Random key graphs have been introduced about a decade ago in the context of the Eschenauer-Gligor random key predistribution scheme [8], which is a widely accepted solution for securing wireless sensor network (WSN) communications. Aside from the significant body of work [2], [6], [25], [29], [33], [34], that studies random key graphs in this context, they have recently received attention in application areas as diverse as classification analysis [11], recommender systems using collaborative filtering [18], trust networks [10], and epidemics in social networks [1]. Random key graphs belong to a larger class of random graphs known as *random intersection* graphs [27]; in fact, they are referred to as the *uniform* random intersection graphs by some authors [2], [11], [12].

A random key graph, denoted $\mathbb{K}(n; K, P)$, can be described as follows: Let $\mathcal{V} = \{1, \ldots, n\}$ denote the set of its and assume available a pool of P keys. Each vertex i is assigned K distinct keys that are selected uniformly at random from this key pool. Two distinct vertices i and j are deemed adjacent, i.e., have an undirected edge in between, as long as they share at least one key in common; see Section II for precise definitions. The terminology used here is borrowed from the field of WSNs, where the term "key" refers to a symmetric cryptographic key used for message encryption and authentication among sensors. In that context, the applicability of random key graph follows naturally as only those pairs of sensor who have a common key will be able to communicate securely. We remark that the terminology can be generalized to replace keys by any object such as activities, books, hobbies, movies, etc, making it possible to apply random key graphs in a wide range of areas.

Much efforts have been devoted to studying connectivity properties in random key graphs. A key motivation can be found in the need to obtain conditions under which the scheme of Eschenauer and Gligor guarantees secure connectivity with high probability in large WSNs. An interesting feature of this work lies in the following fact: Although random key graphs are *not* stochastically equivalent to the classical Erdős-Rényi (ER) graphs [7], it is possible to *formally* transfer well-known zero-one laws for connectivity (and k-connectivity) in ER graphs to random key graphs by asymptotically matching their edge probabilities. This approach, which was initiated by Eschenauer and Gligor in their original analysis [8], has now been validated rigorously; see the papers [2], [6], [25], [33], [34] for recent developments. Rybarczyk [25] has shown that this transfer from ER graphs also works for a number of issues related to the *giant component* and diameter.

In view of these developments, it is natural to wonder whether this (formal) transfer technique applies to other graph properties. In particular, in the literature on random graphs there is long standing interest [3], [15], [16], [24], [27] in the containment of certain (small) subgraphs, the simplest one being the *triangle*. This last case is also closely related to the *clustering* properties of a graph. Informally defined as the propensity of a node's neighbors to be neighbors as well, high clustering (or transitivity) has been observed in many real-world networks [26]. In addition, the clustering level is known to affect the dynamics of various network processes signif-

This work was supported in part by NSF Grant CCF-0729093 and parts of the material were presented in the 47th Annual Allerton Conference on Communication, Control and Computing, Monticello (IL), September 2009 [30], and in the First Workshop on Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks (GRAPH-HOC 2009), Chennai (India), December 2009 [31].

O. Yağan is with the Department of Electrical and Computer Engineering and CyLab, Carnegie Mellon University, Moffett Field, CA 94035 USA (e-mail: oyagan@ece.cmu.edu).

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

icantly; e.g., diffusion of information and epidemic diseases [20], [22], [9], propagation of influence [13], and cascading failures [14].

With these in mind, in the present paper we study the existence and number of triangles, and clustering coefficient in random key graphs. In particular, we establish a zero-one law for the existence of triangles in random key graphs and identify the corresponding threshold function. On the way to this result, we show that in the many node regime, the expected number of triangles in random key graphs is always at least as large as the corresponding quantity in asymptotically matched ER graphs. For the parameter range of practical relevance in both WSNs and social networks, this expected number of triangles is shown to be orders of magnitude larger in random key graphs than in ER graphs. Thus, we conclude that transferring results from ER graphs by matching their edge probabilities is not a valid approach in general, and can be quite misleading in the context of WSNs, reinforcing the call for a direct investigation of random key graphs.

With regard to the *clustering coefficient* of random key graphs, our main contributions are as follows. First, we establish a strong consistency result that validates using a simple definition of clustering based on the probability of a triangle conditioned on a *wedge*, in lieu of the widely adopted definition of *global* clustering coefficient [23]; see Section II-B for precise definitions. To the best of our knowledge, this is the first result in the literature that establishes the asymptotical equivalence of the two aforementioned definitions of clustering coefficient, in any random graph model. Next, we compare the clustering coefficients of random key graphs with that of ER graphs. We observe that the clustering coefficient of a random key graph is never smaller than the clustering coefficient of the corresponding ER graph with *identical* expected average degree. For the parameter range that is practically relevant for WSNs as well as social networks, we show that random key graphs are in fact *much more clustered* than ER graphs when expected average degrees are asymptotically equivalent. Recalling the fact that random key graphs also have a small *diameter* [25], we then conclude that random key graphs are *small-worlds* in the sense introduced by Watts and Strogatz [28]. This reinforces the possibility of using random key graphs in a wide range of applications including modeling social networks.

We believe that the results in this paper shed light on various interesting properties of a random graph model that is becoming increasingly popular across various disciplines. As such, we believe this work will facilitate further studies of random key graphs both in the existing application areas as well as new ones suggested in this paper; e.g., small-world networks.

We organize the rest of the paper as follows: In Section II we formally introduce the class of random key graphs and define its clustering coefficient. Section III is devoted to some preliminary findings that will be useful in presenting the main results of the paper. There we also provide definitions and facts concerning Erdős-Rényi graphs. The main results of the paper concerning the containment of triangles in random key graphs and its clustering coefficient are presented in Section IV. Section IV-C and Section IV-E are devoted to comparing random key graphs and Erdős-Rényi graphs in terms of their number of triangles and clustering coefficients, respectively. The main results of the paper are proved in Section V.

A word on the notation and conventions in use: All limiting statements, including asymptotic equivalences, are understood with n going to infinity. The random variables (rvs) under consideration are all defined on the same probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. Probabilistic statements are made with respect to this probability measure $\mathbb{P}$, and we denote the corresponding expectation operator by $\mathbb{E}$. The indicator function of an event E is denoted by $\mathbf{1}[E]$. For any discrete set S we write $|S|$ for its cardinality.

II. Model Definitions

Pick positive integers K and P such that $K \leq P$, and fix $n = 3, 4, \ldots$. We shall group the integers P and K into the ordered pair $\theta \equiv (K, P)$ in order to lighten the notation.

A. Random key graphs

The model is parametrized by the number n of nodes, the size P of the key pool and the size K of each key ring. For each node $i = 1, \ldots, n$, let $K_i(\theta)$ denote the random set of K distinct keys assigned to node i. Thus, under the convention that the P keys are labelled $1, \ldots, P$, the random set $K_i(\theta)$ is a subset of $\{1, \ldots, P\}$ with $|K_i(\theta)| = K$. The rvs $K_1(\theta), \ldots, K_n(\theta)$ are assumed to be *i.i.d.* rvs, each of which is *uniformly* distributed with

$$\mathbb{P}[K_i(\theta) = S] = \binom{P}{K}^{-1}, \qquad i = 1, \ldots, n \tag{1}$$

for any subset S of $\{1, \ldots, P\}$ with $|S| = K$. This corresponds to selecting keys randomly and *without* replacement from the key pool.

Distinct nodes $i, j = 1, \ldots, n$ are said to be adjacent if they share at least one key in their key rings, namely

$$K_i(\theta) \cap K_j(\theta) \neq \emptyset, \tag{2}$$

in which case an undirected link is assigned between nodes i and j. We find it convenient to introduce the event $E_{ij}(\theta)$ where (2) takes place, i.e.,

$$E_{ij}(\theta) = [K_i(\theta) \cap K_j(\theta) \neq \emptyset],$$

with indicator function $\xi_{ij}(\theta) = \mathbf{1}[E_{ij}(\theta)]$. The adjacency constraints (2) define a random graph on the vertex set $\{1, \ldots, n\}$, hereafter denoted $\mathbb{K}(n; \theta)$. We refer to it as the *random key graph*.

It is easy to check that

$$\mathbb{P}[K_i(\theta) \cap K_j(\theta) = \emptyset] = q(\theta) \tag{3}$$

with

$$q(\theta) = \frac{\binom{P-K}{K}}{\binom{P}{K}} \cdot \mathbf{1}[2K \leq P]. \tag{4}$$

The probability $p(\theta)$ of edge occurrence between any two nodes is therefore given by

$$p(\theta) = 1 - q(\theta). \tag{5}$$

If $P < 2K$ there exists an edge between any pair of nodes, and $\mathbb{K}(n;\theta)$ coincides with the complete graph on the vertex set $\{1,\dots,n\}$. While it is always the case that $0 \le q(\theta) < 1$, it is plain from (4) that $q(\theta) > 0$ if and only if $2K \le P$.

The expression (4) is a consequence of the fact

$$\mathbb{P}\left[S \cap K_i(\theta) = \emptyset\right] = \frac{\binom{P-|S|}{K}}{\binom{P}{K}}, \quad i = 1,\dots,n \tag{6}$$

valid for every subset S of $\{1,\dots,P\}$ with $|S| \le P-K$. For each $i = 1,\dots,n$, it is a simple matter to check with the help of (6) that the events $\{E_{ij}(\theta), \quad j \ne i,\ j = 1,\dots,n\}$ are mutually independent, or equivalently, that the rvs $\{\xi_{ij}(\theta), \quad j \ne i,\ j = 1,\dots,n\}$ form a collection of i.i.d. rvs.

B. Clustering coefficients

A formal definition of clustering in a network is given next. Consider an undirected graph G with no self-loops on the vertex set V. For each i in V, let $T_i(G)$ denote the number of distinct triangles in G that contain vertex i. The *local* clustering coefficient of *node* i is given by

$$C_i(G) = \begin{cases} \frac{T_i(G)}{\frac{1}{2}d_i(d_i-1)} & \text{if } d_i \ge 2 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where d_i is the degree of node i in G.

There are, however, several possible definitions for a graph-*wide* notion of clustering [23]: Inspired by (7), it is natural to consider the *average* of the local clustering coefficient $C_{\rm Avg}(G)$ over the graph G, i.e.,

$$C_{\rm Avg}(G) = \frac{1}{|V'|}\sum_{i\in V'} C_i(G) \tag{8}$$

where $V' = \{i \in V :\ d_i \ge 2\}$. This last definition, while a natural one, is often replaced by the notion of *global* clustering coefficient defined as the "fraction of transitive triples" over the whole graph G. Namely,

$$C^\star(G) = \frac{\sum_{i\in V} T_i(G)}{\frac{1}{2}\sum_{i\in V} d_i(d_i-1)} \tag{9}$$

provided $\sum_{i\in V} d_i(d_i-1) > 0$. It is convenient to set $C^\star(G) = 0$ otherwise.

Related (but simpler) definitions are possible in the context of random graphs with *exchangeable* link assignments (as is the case for the random graphs of interest here), e.g., [5]. In particular, a possible approach is to define the clustering coefficient of the random key graph $\mathbb{K}(n;\theta)$ by

$$C_{\rm K}(\theta) = \mathbb{P}\left[E_{12}(\theta) \mid E_{13}(\theta) \cap E_{23}(\theta)\right]. \tag{10}$$

Interest in this quantity stems from the fact that it is expected to provide a good approximation to (9) when n is large. One of our main contributions concerning the clustering in random key graphs formalizes this phenomenon with a strong consistency result in Theorem 4.4. This justifies using the simpler definition (10) for clustering for the rest of this paper.

III. Preliminaries

A. Counting triangles

Pick positive integers K and P such that $K \le P$, and fix $n = 3,4,\dots$ For distinct $i,j,k = 1,\dots,n$, we define the indicator function

$$\chi_{ijk}(\theta) = \mathbf{1}\left[\text{Nodes } i,j,k \text{ form a triangle in } \mathbb{K}(n;\theta)\right]. \tag{11}$$

The number of distinct triangles in $\mathbb{K}(n;\theta)$ is then simply given by

$$T_n(\theta) = \sum\nolimits_{1\le i<j<k\le n}^{n} \chi_{ijk}(\theta). \tag{12}$$

Of particular interest will be the event that there exists at least one triangle in $\mathbb{K}(n;\theta)$, namely $[T_n(\theta) > 0] = [T_n(\theta) = 0]^c$.

Key to the analysis presented here is our ability to evaluate the first two moments of the count variables (12). The first moment, computed next, will be conveniently expressed with the help of the quantity $\beta(\theta)$ given by

$$\beta(\theta) = (1-q(\theta))^3 + q(\theta)^3 - q(\theta)r(\theta) \tag{13}$$

with $r(\theta)$ defined by

$$r(\theta) = \frac{\binom{P-2K}{K}}{\binom{P}{K}} \cdot \mathbf{1}\left[3K \le P\right]. \tag{14}$$

Note that $r(\theta)$ corresponds to the probability (6) when $|S| = 2K$. The next result is established in Section V-A.

Proposition 3.1: Fix $n = 3,4,\dots$. For positive integers K and P such that $K \le P$, we have

$$\mathbb{E}\left[\chi_{123}(\theta)\right] = \beta(\theta) \tag{15}$$

with $\beta(\theta)$ defined at (13), so that

$$\mathbb{E}\left[T_n(\theta)\right] = \binom{n}{3}\beta(\theta). \tag{16}$$

Thus, $\beta(\theta)$ is simply the probability that three distinct vertices form a triangle in $\mathbb{K}(n;\theta)$. For future reference, we note

$$r(\theta) \le q(\theta)^2 \tag{17}$$

by direct inspection, whence

$$\beta(\theta) \ge (1-q(\theta))^3 > 0. \tag{18}$$

B. Asymptotic equivalences

For simplicity of exposition we refer to any pair of functions $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ as a *scaling* (for random key graphs) provided the natural condition $K_n \le P_n$ holds for all $n = 2,3,\dots$. The two asymptotic equivalence results presented next prove useful in a number of places. They provide easy asymptotic expression for the probability of a link and for the probability of a triangle, respectively in the random key graph. The first one was already obtained in [33], and is given here for easy reference.

Lemma 3.2: For any scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$, we have

$$\lim_{n\to\infty} q(\theta_n) = 1 \quad \textit{if and only if} \quad \lim_{n\to\infty} \frac{K_n^2}{P_n} = 0. \tag{19}$$

Under either condition at (19) we have the asymptotic equivalence

$$1 - q(\theta_n) \sim \frac{K_n^2}{P_n}. \tag{20}$$

The second asymptotic equivalence, whose proof is given in Appendix A, will be stated in terms of the quantity

$$\tau(\theta) = \frac{K^3}{P^2} + \left(\frac{K^2}{P}\right)^3, \quad K, P = 1, 2, \dots \tag{21}$$

Proposition 3.3: For any scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying (19), we have the asymptotic equivalence

$$\beta(\theta_n) \sim \tau(\theta_n). \tag{22}$$

C. Facts concerning Erdős-Rényi graphs

A little later in this paper, we shall compare random key graphs to related Erdős-Rényi (ER) graphs [7], but first some notation: For each $n = 2, 3, \dots$ and each p in $[0, 1]$, let $\mathbb{G}(n; p)$ denote the ER graph on the vertex set $\{1, \dots, n\}$ with link assignment probability p. In analogy with earlier notation let $E_{ij}(p)$ denote the event that there is an (undirected) link between nodes i and j in $\mathbb{G}(n; p)$. Thus, the ER graph $\mathbb{G}(n; p)$ is characterized by the events $\{E_{ij}(p),\ 1 \le i < j \le n\}$ that are mutually independent, each of probability p. We refer to any mapping $p : \mathbb{N}_0 \to [0, 1]$ as a scaling for ER graphs.

In analogy with (12) let $T_n(p)$ denote the number of distinct triangles in $\mathbb{G}(n; p)$. Under the enforced independence, we note that

$$\mathbb{E}\left[T_n(p)\right] = \binom{n}{3} \tau^\star(p), \quad n = 3, 4, \dots \tag{23}$$

with

$$\tau^\star(p) = p^3, \quad 0 \le p \le 1.$$

Link assignments being exchangeable in ER graphs, we again define the clustering coefficient in $\mathbb{G}(n; p)$ by

$$C_{\mathrm{ER}}(p) = \mathbb{P}\left[E_{12}(p) \mid E_{13}(p) \cap E_{23}(p)\right] \tag{24}$$

so that

$$C_{\mathrm{ER}}(p) = \frac{\mathbb{P}\left[E_{12}(p) \cap E_{13}(p) \cap E_{23}(p)\right]}{\mathbb{P}\left[E_{13}(p) \cap E_{23}(p)\right]} = p \tag{25}$$

by mutual independence. Here as well, we expect

$$\lim_{n\to\infty} C^\star(\mathbb{G}(n; p)) = C_{\mathrm{ER}}(p) \quad a.s. \tag{26}$$

This can be established by the same arguments as the ones provided in the proof of Proposition 4.4.

Random key graphs are *not* equivalent to ER graphs even when their edge probabilities are matched exactly: As graph-valued rvs, the random graphs $\mathbb{G}(n; p)$ and $\mathbb{K}(n; \theta)$ have different distributions under the *exact matching* condition

$$p = 1 - q(\theta) = p(\theta). \tag{27}$$

See [30] for a discussion of (dis)similarities. However, in order to meaningfully compare the asymptotic regime of random key graphs with that of ER graphs, we shall say that the scaling $p : \mathbb{N}_0 \to [0, 1]$ (for ER graphs) is *asymptotically matched* to the scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ (for random key graphs) if

$$p_n \sim p(\theta_n) = 1 - q(\theta_n). \tag{28}$$

Under the condition (19), the asymptotic matching condition (28) amounts to (see Lemma 3.2)

$$p_n \sim \frac{K_n^2}{P_n}. \tag{29}$$

Condition (27) (resp. (28)) is equivalent to requiring that the expected degrees in $\mathbb{K}(n; \theta_n)$ and $\mathbb{G}(n; p_n)$ coincide (resp. be asymptotically equivalent).

IV. Main Results and Discussion

A. Zero-one laws for the existence of triangles

A main result of the paper is a zero-one law for the existence of triangles in random key graphs. The zero-law is given first.

Theorem 4.1: For any scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$, the zero-law

$$\lim_{n\to\infty} \mathbb{P}\left[T_n(\theta_n) > 0\right] = 0$$

holds under the condition

$$\lim_{n\to\infty} n^3 \tau(\theta_n) = 0. \tag{30}$$

The one-law given next assumes a more involved form; its proof is given in Section V-D.

Theorem 4.2: For any scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ for which the limit $\lim_{n\to\infty} q(\theta_n) = q^\star$ exists, the one-law

$$\lim_{n\to\infty} \mathbb{P}\left[T_n(\theta_n) > 0\right] = 1$$

holds if either $0 \le q^\star < 1$ or $q^\star = 1$ with the additional condition

$$\lim_{n\to\infty} n^3 \tau(\theta_n) = \infty. \tag{31}$$

To facilitate an upcoming comparison with analogous results in ER graphs, we combine Theorem 4.1 and Theorem 4.2 into the symmetric statement.

Theorem 4.3: For any scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ for which $\lim_{n\to\infty} q(\theta_n)$ exists, we have

$$\lim_{n\to\infty} \mathbb{P}\left[T_n(\theta_n) > 0\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} n^3 \tau(\theta_n) = 0 \\ 1 & \text{if } \lim_{n\to\infty} n^3 \tau(\theta_n) = \infty. \end{cases}$$

By Lemma 3.2 we note that the condition $\lim_{n\to\infty} n^3 \tau(\theta_n) = 0$ implies $\lim_{n\to\infty} q(\theta_n) = 1$ (hence $q^\star = 1$).

B. Clustering in Random Key Graphs

Our first result concerning the clustering coefficient of random key graphs establishes a strong consistency between two aforementioned definitions of graph clustering; its proof is given in Appendix V-E.

Theorem 4.4: For positive integers K, P such that $K \leq P$, we have

$$\lim_{n\to\infty} C^\star(\mathbb{K}(n;\theta)) = C_{\mathrm{K}}(\theta) \quad a.s. \tag{32}$$

Simulation results given in Table I of Section IV-B illustrate the convergence (32) (and an analog one for ER graphs). The next result computes $C_{\mathrm{K}}(\theta)$. Its proof is given in Section V-B.

Proposition 4.5: For positive integers K, P such that $K \leq P$, we have

$$C_{\mathrm{K}}(\theta) = \frac{\beta(\theta)}{(1-q(\theta))^2} \tag{33}$$

with $\beta(\theta)$ given by (13).

C. Comparing the number of triangles in random key graphs and ER graphs

Fix p in $(0,1]$, and positive integers K and P such that $K \leq P$. From (16) and (23) it is plain that

$$\frac{\mathbb{E}\left[T_n(\theta)\right]}{\mathbb{E}\left[T_n(p)\right]} = \frac{\beta(\theta)}{\tau^\star(p)}, \quad n = 3,4,\ldots \tag{34}$$

Under the *exact* matching condition (27), this last expression becomes

$$\frac{\mathbb{E}\left[T_n(\theta)\right]}{\mathbb{E}\left[T_n(p(\theta))\right]} = \frac{\beta(\theta)}{\tau^\star(p(\theta))} = 1 + \frac{q(\theta)^2 - r(\theta)}{(1-q(\theta))^3}\cdot q(\theta) \tag{35}$$

for each $n = 3,4,\ldots$, whence $\mathbb{E}\left[T_n(p(\theta))\right] \leq \mathbb{E}\left[T_n(\theta)\right]$ by virtue of (17). Consequently, the expected number of triangles in a random key graph is always at least as large as the corresponding quantity in an ER graph matched to it. This was already suggested by Di Pietro et al. [6] with the help of limited simulations.

A similar result is available when the scalings are only *asymptoticlly* matched.

Corollary 4.6: Consider a scaling $K, P : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying (19), and a scaling $p : \mathbb{N}_0 \to [0,1]$. Under the asymptotic matching condition (28), we have the equivalence

$$\frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} \sim 1 + \frac{P_n}{K_n^3}. \tag{36}$$

In other words, for large n the expected number of triangles in random key graphs is always at least as large as the corresponding quantity in asymptotically matched ER graphs. In fact, the number of triangles in random key graphs can be several orders of magnitude larger than that of ER graphs, if $\lim_{n\to\infty} P_n/K_n^3 = \infty$. We argue in Section IV-D that this condition is likely to hold in many envisioned applications of random key graphs, particularly in the context of wireless sensor networks.

Proof. Replacing θ by θ_n and p by p_n according to the given scalings in the expression (34), we get

$$\frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} = \frac{\beta(\theta_n)}{\tau^\star(p_n)}, \quad n = 3,4,\ldots$$

Under (19), Proposition 3.3 yields

$$\frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} \sim \frac{\tau(\theta_n)}{\tau^\star(p_n)} \tag{37}$$

with

$$\frac{\tau(\theta_n)}{\tau^\star(p_n)} = \frac{1}{p_n^3}\cdot\left(\frac{K_n^3}{P_n^2}\right) + \frac{1}{p_n^3}\cdot\left(\frac{K_n^2}{P_n}\right)^3, \quad n = 2,3,\ldots$$

With the help of (29), we now conclude

$$\frac{\tau(\theta_n)}{\tau^\star(p_n)} \sim 1 + \frac{P_n}{K_n^3} \tag{38}$$

and the equivalence (36) follows from (37). ■

We conclude this section by comparing Theorem 4.3 with its analog for ER graphs. Fix $n = 2,3,\ldots$ and p in $[0,1]$. Consider the event that there exists at least one triangle in $\mathbb{G}(n;p)$, i.e., $[T_n(p) > 0]$. The following zero-one law for triangle containment in ER graphs is well known [3, Chp. 4], [15, Thm. 3.4, p. 56].

Theorem 4.7: For any scaling $p : \mathbb{N}_0 \to [0,1]$, we have

$$\lim_{n\to\infty} \mathbb{P}\left[T_n(p_n) > 0\right] = \begin{cases} 0 & \text{if } \lim_{n\to\infty} n^3\tau^\star(p_n) = 0 \\ 1 & \text{if } \lim_{n\to\infty} n^3\tau^\star(p_n) = \infty \end{cases}$$

This result was also established by the method of first and second moments, and its form is easily understood once we recall (23).

As we compare Theorem 4.3 and Theorem 4.7, we see a direct analogy since the terms $\tau(\theta_n)$ and $\tau^\star(p_n)$ both correspond to (asymptotic) probability that three arbitrary nodes form a triangle, in random key graphs and ER graphs, respectively. More interestingly, we see that under the asymptotic matching condition (28) (and (19)), triangles will start appearing *earlier* in the evolution of a random key graph as compared to an ER graph matched to it; this fact is evident from (38). Put differently, we can see triangles in a random key graph with smaller asymptotic edge probability (which corresponds to a smaller mean node degree) than in ER graph. As shall be clear from the previous discussions and (38), the larger is the quantity P_n/K_n^3, the more emphasized these differences will be.

We next discuss the behavior of P_n/K_n^3 that is likely to be observed in two popular applications of random key graphs.

D. Implications on random key graph applications: WSNs and Social Networks

In the context of a WSN with n nodes, it is natural to select the parameters K_n and P_n such that the induced random key graph is *connected*. However, there is a tradeoff between connectivity and security [6], requiring $\frac{K_n^2}{P_n}$ to be kept as close

as possible to the critical scaling $\frac{\log n}{n}$ for connectivity; see the papers [2], [6], [25], [29], [33]. The desired regime near the boundary can be achieved by taking

$$\frac{K_n^2}{P_n} \sim c \cdot \frac{\log n}{n} \tag{39}$$

with $c > 1$ but close to one. On the other hand, it follows from (36) that

$$\frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} \sim 1 \quad \text{if and only if} \quad \frac{P_n}{K_n^3} = o(1) \tag{40}$$

provided condition (19) holds. This obviously occurs when (39) holds, in which case the condition at (40) amounts to taking

$$\frac{1}{K_n} \sim o(1) \left(c \cdot \frac{\log n}{n} \right)^{-1}.$$

Thus, combining we see that under (39) it holds that

$$\frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} \sim 1 \quad \text{if and only if} \quad K_n \gg \frac{n}{\log n}. \tag{41}$$

In that case the expected number of triangles in random key graphs is of the same order as the corresponding quantity in asymptotically matched ER graphs with $\mathbb{E}\left[T_n(\theta_n)\right] \sim \mathbb{E}\left[T_n(p_n)\right] \sim \frac{c^3}{6}\left(\log n\right)^3$ – This is a direct consequence of (23), (29) and (39). This conclusion holds regardless of the value of c in (39).

However, given the limited memory and computational power of the sensor nodes, the key ring sizes at (41) are not practical. In addition, they will lead to *high* node degrees and this in turn will decrease network *resiliency* against node capture attacks. Indeed, it was proposed by Di Pietro et al. [6, Thm. 5.3] that resiliency in WSNs against node capture attacks can be ensured by selecting K_n and P_n such that $\frac{K_n}{P_n} \sim \frac{1}{n}$. Under (39) this additional requirement then leads to $K_n \sim c \cdot \log n$ so that $P_n \sim c \cdot n \log n$, and (36) now implies

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[T_n(\theta_n)\right]}{\mathbb{E}\left[T_n(p_n)\right]} = \lim_{n\to\infty} \left(1 + \frac{n}{(c \cdot \log n)^2} \right) = \infty. \tag{42}$$

This means that, for realistic WSN implementations the expected number of triangles in the induced random key graphs will be orders of magnitude larger than in ER graphs.

As mentioned earlier, random key graphs have been considered in various application areas including social network modeling; e.g., see [1] and our discussion in Section IV-F. To that end, we now discuss the parameter ranges for the random key graph that are likely to be observed in applications related to social networks. We will do so with an eye towards understanding the behavior of the expression appearing at (36).

To begin with, most real-world social networks are known [21] to be *sparse* in the sense that the mean number of edges per node is equal to a constant $c > 0$. In the case of random key graphs, the mean degree of a node is given by $(n-1)(1-q(\theta_n))$ so that sparsity implies $1 - q(\theta_n) \sim \frac{c}{n}$, or equivalently that

$$\frac{K_n^2}{P_n} \sim \frac{c}{n} \tag{43}$$

by virtue of Lemma 3.2. In view of Corollary 4.6 and Corollary 4.8, we then see that in the sparse regime random key graphs will have many more triangles and will be much more clustered than the asymptotically matched ER graphs unless we have

$$K_n = \Omega(n). \tag{44}$$

We remark that this condition is even more stringent than (41) that is derived for WSN applications. More importantly, under (43) and (44) generating the random key graph will require each of the n nodes selecting $\Omega(n)$ keys/objects from a pool with size $\Omega(n^3)$. The computational complexity of this will quickly become prohibitively high as the number of individuals in the social network gets large. Moreover, a natural interpretation of the random key graph as a social network model consists of assigning K_n hobbies/interests to each individual and assuming that two of them are *friends* if they have a common hobby or interest. Hence, we would expect the realistic values for one's number of interests K_n to be much smaller than the network size n, in contrast to (44).

Collecting, we see that (44) is naturally eliminated in realistic and feasible applications of random key graphs as social network models. Put differently, random key graphs will naturally have very high clustering and contain very large number of triangles when utilized in social network modeling.

E. Comparing the clustering coefficients of random key graphs and ER graphs

Fix p in $(0, 1]$, and positive integers K and P such that $K \leq P$. Combining (33) and (25) we get

$$\frac{C_K(\theta)}{C_{\mathrm{ER}}(p)} = \frac{\beta(\theta)}{p(1-q(\theta))^2}. \tag{45}$$

Under the exact matching condition (27) we find

$$\frac{C_K(\theta)}{C_{\mathrm{ER}}(p(\theta))} = 1 + \frac{q(\theta)^2 - r(\theta)}{(1-q(\theta))^3} \cdot q(\theta) \tag{46}$$

as we recall (5). Thus, $C_{\mathrm{K}}(\theta) \geq C_{\mathrm{ER}}(p(\theta))$ by virtue of (17) – The clustering coefficient of a random key graph is at least as large as that of the ER graph exactly matched to it.

In fact, the lower bound mentioned above is achieved only when $P < 2K$, i.e., from (4) we get

$$\frac{C_{\mathrm{K}}(\theta)}{C_{\mathrm{ER}}(p(\theta))} = 1 \quad \text{whenever} \quad P < 2K.$$

It is a simple matter to check that for $K = 1$, (48) yields

$$\frac{C_{\mathrm{K}}(1, P)}{C_{\mathrm{ER}}(p(\theta))} = P$$

for each $P = 1, 2, \ldots$, while for $K = 2$ we have

$$\frac{C_{\mathrm{K}}(2, P)}{C_{\mathrm{ER}}(p(\theta))} = \frac{P}{2} \cdot \frac{2P^3 - 4P^2 - P + 3}{(2P-3)^3} \geq \frac{P}{8}.$$

It is also straightforward to show that

$$1 \leq \frac{C_{\mathrm{K}}(\theta)}{C_{\mathrm{ER}}(p(\theta))} \leq P.$$

K	P	$C_K(\theta)$	$\widehat{C}_n^\star(\theta)$	$C_{ER}(p)$	$\widehat{C}_n^\star(p)$
4	10^3	0.2590	0.2587	0.0160	0.0159
8	5×10^3	0.1348	0.1349	0.0127	0.0128
16	2×10^4	0.0737	0.0736	0.0127	0.0128
20	4×10^4	0.0590	0.0590	0.0100	0.0100
24	10^5	0.0469	0.0468	0.0057	0.0057
32	10^5	0.0408	0.0408	0.0102	0.0102
40	5×10^5	0.0280	0.0280	0.0032	0.0031
64	10^6	0.0196	0.0196	0.0041	0.0041

TABLE I
CLUSTERING COEFFICIENTS WITH FIXED θ AND $p = 1 - q(\theta)$

Since P can be made very large, we understand that the parameters of the random key graph can be selected so that it has a much larger clustering coefficient than the ER graph matched to it. This will especially be so for WSNs where the size of the key pool P is expected to be in the range $2^{17} - 2^{20}$ [8].

In Table I we compare the clustering coefficients of random key graphs and ER graphs for several realistic parameter values. The numerical values of $C_{\rm K}(\theta)$ and $C_{\rm ER}(p)$ are obtained directly from the expressions (10) and (24), respectively. On the other hand, $\widehat{C}_n^\star(\theta)$ and $\widehat{C}_n^\star(p)$ stand for the clustering coefficient of $\mathbb{K}(n;\theta)$ and $\mathbb{G}(n;p)$, respectively, calculated through (9) and averaged over 1000 realizations; the number of nodes is set to $n = 1000$ in all simulations. The data support the claim that the definition (9) captures essentially the same feature as the quantities (10) and (24), i.e., the results given in Table I can be taken as an indication of the validity of (32) and (26).

Next we compare the clustering coefficients of random key graphs and ER graphs when the parameters θ and p are scaled with n.

Corollary 4.8: Consider a scaling $K, P : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying (19) and a scaling $p : \mathbb{N}_0 \to [0,1]$. Under the asymptotic matching condition (28), we have the equivalence

$$\frac{C_{\rm K}(\theta_n)}{C_{\rm ER}(p_n)} \sim 1 + \frac{P_n}{K_n^3}. \tag{47}$$

Proof. As we replace θ by θ_n and p by p_n according to these scalings in the expression (45), we get

$$\frac{C_{\rm K}(\theta_n)}{C_{\rm ER}(p_n)} = \frac{\beta(\theta_n)}{p_n(1-q(\theta_n))^2}, \quad n = 1, 2, \ldots \tag{48}$$

Note that

$$\frac{C_{\rm K}(\theta_n)}{C_{\rm ER}(p_n)} \sim \frac{\beta(\theta_n)}{(1-q(\theta_n))^3} \sim \frac{\tau(\theta_n)}{(1-q(\theta_n))^3}. \tag{49}$$

The first equivalence is a consequence of (28) and the second one is validated by Proposition 3.3 under (19). With (29) being still valid here, we easily conclude (47) by the same arguments as the ones used to obtain (36). ∎

Thus, under (19) and (28), we conclude that

$$\lim_{n\to\infty} \frac{C_{\rm K}(\theta_n)}{C_{\rm ER}(p_n)} = 1 \quad \text{if} \quad \lim_{n\to\infty} \frac{K_n^3}{P_n} = \infty, \tag{50}$$

and

$$\lim_{n\to\infty} \frac{C_{\rm K}(\theta_n)}{C_{\rm ER}(p_n)} = \infty \quad \text{if} \quad \lim_{n\to\infty} \frac{K_n^3}{P_n} = 0. \tag{51}$$

It is now easy to see that asymptotically matched random key graphs and ER graphs can have vastly different clustering coefficients: Under the condition (39), (50) can hold only if the key ring size K is much larger than $n/\log n$. As already discussed, this condition can not be satisfied in a practical WSN scenario due to limitations on the sensor nodes and security constraints. In fact, we see from (36) and (42) that, in a realistic WSN scenario, the condition (51) is always in effect and the clustering coefficient of the random key graph is much larger than that of the asymptotically matched ER graph. This provides yet another property of the random key graphs that ER graph models can not adequately capture, further indicating the non-equivalence of the two models for practical WSN implementations.

F. Random Key Graphs as Small Worlds?

Since random key graphs can be highly clustered, a natural question arises as to their suitability to model the *small world effect*. This notion is linked to a well-known series of experiments conducted by Milgram [19] in the late sixties. The results, commonly known as six degrees of separation, suggest that the social network of people in the United States is *small* in the sense that the path lengths between pairs of individuals are short. As a way to capture Milgram's experiments, Watts and Strogatz [28] defined *small worlds* as network models that are highly clustered and yet have a small average path length. More precisely, a random graph is considered to be a small world if its average path length is of the same order as that of an ER graph with the same expected average degree, but with a much larger clustering coefficient.

The results of this paper already show that random key graphs can satisfy the high clustering coefficient requirement of a small world. Recently Rybarczyk [25] has shown under (39) that

$$\text{diam}[\mathcal{K}(n;\theta_n)] \sim \frac{\log n}{\log\log n}$$

with high probability where $\mathcal{K}(n;\theta_n)$ is the largest connected component of $\mathbb{K}(n;\theta_n)$. This suggests that the diameter, hence the average path length, in random key graphs is *small* as was the case with ER graphs [4]. We also note [32, Corollary 5.2] that random key graphs have *very small* (e.g., ≤ 2) diameter under certain parameter ranges (e.g., with $P_n = n^{1/2-\delta}$ for any $\delta > 0$). Therefore, random key graphs may indeed be considered as good candidate models for small worlds!

The fact that random key graphs can exhibit *small world* properties can be taken as a further motivation to pursue other application areas for them, particularly in the field of social networks. We provide one concrete possibility below, where random key graphs appear naturally in modeling *common-interest* relationships in a population: Suppose that there exists an *object pool* consisting of P objects and that each of the n users picks K distinct objects uniformly at random from this object pool; objects may represent hobbies, books, movies, etc. Two *friends* are said to have a common-interest relation if

they have at least *one* common object in their object rings. This naturally suggests modeling the common interest relationship by a random key graph $\mathbb{K}(n; K, P)$. Various properties of random key graphs would then be of interest in applications such as implementation of large-scale, distributed publish-subscribe services [34].

V. Proofs of Main Results

A. A proof of Proposition 3.1

Fix positive integers K and P such that $K \leq P$. As exchangeability yields (16), we need only show the validity of (15). In the discussion that follows we omit the explicit dependence on θ when no confusion arises from doing so. Also, we make repeated use of the fact that for any pair of events E and F in $\mathcal{F}$, we have

$$\mathbb{P}[E \cap F] = \mathbb{P}[E] - \mathbb{P}[E \cap F^c]. \tag{52}$$

Thus, by repeated application of (52) we find

$$\begin{aligned}
&\mathbb{E}[\chi_{123}(\theta)] \\
&\quad = \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 \neq \emptyset, K_2 \cap K_3 \neq \emptyset] \\
&\quad = \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 \neq \emptyset] \\
&\qquad - \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 \neq \emptyset, K_2 \cap K_3 = \emptyset] \\
&\quad = \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 \neq \emptyset] \\
&\qquad - \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_2 \cap K_3 = \emptyset] \\
&\qquad + \mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset].
\end{aligned}$$

By independence, with the help of (6), we readily obtain the expressions

$$\mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 \neq \emptyset] = (1 - q(\theta))^2 \tag{53}$$

and

$$\mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_2 \cap K_3 = \emptyset] = (1 - q(\theta))\, q(\theta).$$

Next, as we use (52) one more time, we get

$$\begin{aligned}
&\mathbb{P}[K_1 \cap K_2 \neq \emptyset, K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset] \\
&\quad = \mathbb{P}[K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset] \\
&\qquad - \mathbb{P}[K_1 \cap K_2 = \emptyset, K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset].
\end{aligned}$$

Again, by independence, with the help of (6) we conclude that

$$\mathbb{P}[K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset] = q(\theta)^2 \tag{54}$$

and

$$\begin{aligned}
&\mathbb{P}[K_1 \cap K_2 = \emptyset, K_1 \cap K_3 = \emptyset, K_2 \cap K_3 = \emptyset] \\
&\quad = \mathbb{P}[K_1 \cap K_2 = \emptyset, K_3 \cap (K_1 \cup K_2) = \emptyset] \\
&\quad = q(\theta) r(\theta)
\end{aligned} \tag{55}$$

since $|K_1 \cup K_2| = 2K$ when $K_1 \cap K_2 = \emptyset$. Collecting these facts we find

$$\begin{aligned}
&\mathbb{E}[\chi_{123}(\theta)] \\
&\quad = (1 - q(\theta))^2 - (1 - q(\theta))\, q(\theta) + q(\theta)^2 - q(\theta) r(\theta)
\end{aligned}$$

and the conclusion (15) follows by elementary algebra. ∎

B. A proof of Proposition 4.5

The definitions of $C_{\mathrm{K}}(\theta)$ and $\chi_{123}(\theta)$ yield

$$\begin{aligned}
C_{\mathrm{K}}(\theta) &= \frac{\mathbb{P}[E_{12}(\theta) \cap E_{13}(\theta) \cap E_{23}(\theta)]}{\mathbb{P}[E_{13}(\theta) \cap E_{23}(\theta)]} \\
&= \frac{\mathbb{E}[\chi_{123}(\theta)]}{(1 - q(\theta))^2}
\end{aligned} \tag{56}$$

since the events $E_{13}(\theta)$ and $E_{23}(\theta)$ are independent, with

$$\begin{aligned}
&\mathbb{P}[E_{13}(\theta) \cap E_{23}(\theta)] \\
&\quad = \mathbb{P}[K_1(\theta) \cap K_3(\theta) \neq \emptyset, K_2(\theta) \cap K_3(\theta) \neq \emptyset] \\
&\quad = (1 - q(\theta))^2
\end{aligned} \tag{57}$$

by virtue of (6) (and comments following it). The conclusion (33) is immediate upon substituting (15) into (56). ∎

C. A proof of Theorem 4.1

Consider a scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$. For each $n = 3, 4, \ldots$, the elementary bound $\mathbb{P}[T_n(\theta_n) > 0] \leq \mathbb{E}[T_n(\theta_n)]$ implies

$$\mathbb{P}[T_n(\theta_n) > 0] \leq \binom{n}{3} \beta(\theta_n)$$

by virtue of Proposition 3.1. Theorem 4.1 thus follows if under (30) we show that $\lim_{n\to\infty} \binom{n}{3} \beta(\theta_n) = 0$. By Proposition 3.3 this convergence is equivalent to the assumed condition $\lim_{n\to\infty} n^3 \tau(\theta_n) = 0$, and the proof of Theorem 4.1 is now complete. ∎

D. A proof of Theorem 4.2

The second moment of the count variables (12) is computed next and will play a crucial role in the proof of Theorem 4.2.

Proposition 5.1: For positive integers K and P such that $K \leq P$, we have for all $n = 3, 4, \ldots$

$$\begin{aligned}
\mathbb{E}\left[T_n(\theta)^2\right] &= \mathbb{E}[T_n(\theta)] + \left(\frac{\binom{n-3}{3}}{\binom{n}{3}} + 3 \frac{\binom{n-3}{2}}{\binom{n}{3}} \right) \left(\mathbb{E}[T_n(\theta)]\right)^2 \\
&\quad + 3(n-3) \binom{n}{3} \cdot \mathbb{E}[\chi_{123}(\theta) \chi_{124}(\theta)].
\end{aligned} \tag{58}$$

We defer the proof of Proposition 5.1 to the end of this section.

We now turn to proving Theorem 4.2. Assume first that $q^\star$ satisfies $0 \leq q^\star < 1$. Fix $n = 3, 4, \ldots$ and partition the n nodes into the $k_n + 1$ non-overlapping groups $(1, 2, 3)$, $(4, 5, 6)$, $\ldots$, $(3k_n + 1, 3k_n + 2, 3k_n + 3)$ with $k_n = \lfloor \frac{n-3}{3} \rfloor$. If $\mathbb{K}(n; \theta_n)$ contains no triangle, then *none* of these $k_n + 1$ groups of nodes forms a triangle. With this in mind we get

$$\begin{aligned}
&\mathbb{P}[T_n(\theta_n) = 0] \\
&\quad \leq \mathbb{P}\left[\bigcap_{\ell=0}^{k_n} \left[\begin{array}{c} \text{Nodes } 3\ell+1, 3\ell+2, 3\ell+3 \text{ do not} \\ \text{form a triangle in } \mathbb{K}(n; \theta_n) \end{array} \right] \right] \\
&\quad = \prod_{\ell=0}^{k_n} \mathbb{P}\left[\begin{array}{c} \text{Nodes } 3\ell+1, 3\ell+2, 3\ell+3 \text{ do not} \\ \text{form a triangle in } \mathbb{K}(n; \theta_n) \end{array} \right]
\end{aligned} \tag{59}$$

$$
\begin{aligned}
&= (1-\beta(\theta_n))^{k_n+1} \\
&\le \left(1-(1-q(\theta_n))^3\right)^{k_n+1} \qquad (60)\\
&\le e^{-(k_n+1)(1-q(\theta_n))^3}. \qquad (61)
\end{aligned}
$$

Note that (59) follows from the fact that the events

$$
\left[\begin{array}{c}\text{Nodes } 3\ell+1, 3\ell+2, 3\ell+3 \text{ do not}\\ \text{form a triangle in } \mathbb{K}(n;\theta_n)\end{array}\right],\ \ell=0,\ldots,k_n
$$

are mutually independent due to the non-overlap condition, while the inequality (60) is justified with the help of (18). Let n go to infinity in the inequality (61). From the constraint $q^\star < 1$ we conclude that $\lim_{n\to\infty}\mathbb{P}\left[T_n(\theta_n)=0\right]=0$ since $k_n \sim \frac{n}{3}$ so that $\lim_{n\to\infty}(k_n+1)(1-q(\theta_n))^3=\infty$. This establishes the one law in the case $q^\star<1$.

To handle the case $q^\star = 1$, we use a standard bound which forms the basis of the method of second moment [15, Remark 3.1, p. 55]. Here this bound takes the form

$$
\frac{\left(\mathbb{E}\left[T_n(\theta_n)\right]\right)^2}{\mathbb{E}\left[T_n(\theta_n)^2\right]} \le \mathbb{P}\left[T_n(\theta_n)>0\right], \quad n=3,4,\ldots \qquad (62)
$$

Theorem 4.2 then will be established in the case $q^\star=1$ if we show under (19) that the condition (31) implies

$$
\lim_{n\to\infty}\frac{\mathbb{E}\left[T_n(\theta_n)^2\right]}{\left(\mathbb{E}\left[T_n(\theta_n)\right]\right)^2}=1. \qquad (63)
$$

As pointed earlier, the conditions (19) imply $3K_n < P_n$ for all n sufficiently large in $\mathbb{N}_0$. On that range, with θ replaced by θ_n, Proposition 5.1 yields

$$
\begin{aligned}
\frac{\mathbb{E}\left[T_n(\theta_n)^2\right]}{\left(\mathbb{E}\left[T_n(\theta_n)\right]\right)^2} &= \frac{1}{\mathbb{E}\left[T_n(\theta_n)\right]}+\left(\frac{\binom{n-3}{3}}{\binom{n}{3}}+3\frac{\binom{n-3}{2}}{\binom{n}{3}}\right)\\
&\quad+\frac{3(n-3)}{\binom{n}{3}}\cdot\frac{\mathbb{E}\left[\chi_{123}(\theta_n)\chi_{124}(\theta_n)\right]}{\left(\mathbb{E}\left[\chi_{123}(\theta_n)\right]\right)^2}
\end{aligned}
$$

as we make use of (16) in the last term.

Let n go to infinity in the resulting expression: Under condition (31), we have $\lim_{n\to\infty} n^3\beta(\theta_n)=\infty$ by Proposition 3.3, whence $\lim_{n\to\infty}\mathbb{E}\left[T_n(\theta_n)\right]=\infty$ by virtue of (16). Since

$$
\lim_{n\to\infty}\left(\frac{\binom{n-3}{3}}{\binom{n}{3}}+3\frac{\binom{n-3}{2}}{\binom{n}{3}}\right)=1 \qquad (64)
$$

and

$$
\frac{\binom{n}{3}}{3(n-3)}\sim\frac{n^2}{18}, \qquad (65)
$$

the convergence (63) will hold if we show that

$$
\lim_{n\to\infty}\frac{1}{n^2}\frac{\mathbb{E}\left[\chi_{123}(\theta_n)\chi_{124}(\theta_n)\right]}{\left(\mathbb{E}\left[\chi_{123}(\theta_n)\right]\right)^2}=0 \qquad (66)
$$

under the foregoing conditions on the scaling.

This is shown as follows: Given positive integers K and P such that $K\le P$, fix $n=3,4,\ldots$. It is immediate that

$$
\begin{aligned}
&\mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right]\\
&\quad\le \mathbb{E}\left[\chi_{123}(\theta)\mathbf{1}\left[K_1(\theta)\cap K_4(\theta)\neq\emptyset\right]\right]. \qquad (67)
\end{aligned}
$$

From (6) it follows that the rvs $\chi_{123}(\theta)$ and $\mathbf{1}\left[K_1(\theta)\cap K_4(\theta)\neq\emptyset\right]$ are independent conditionally on $K_1(\theta)$, and an easy conditioning argument yields

$$
\mathbb{E}\left[\chi_{123}(\theta)\mathbf{1}\left[K_1(\theta)\cap K_4(\theta)\neq\emptyset\right]\right]=\beta(\theta)(1-q(\theta)) \qquad (68)
$$

as we recall (4) and (73). Using (67) together with (15) and (68) we readily obtain the inequalities

$$
\frac{\mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right]}{\left(\mathbb{E}\left[\chi_{123}(\theta)\right]\right)^2}\le\frac{\beta(\theta)(1-q(\theta))}{\beta(\theta)^2}\le\beta(\theta)^{-2/3} \qquad (69)
$$

where in the last step we noted that $1-q(\theta)\le\beta(\theta)^{1/3}$ by appealing to (18).

Returning to the convergence (66) we see from (69) that we need only show

$$
\lim_{n\to\infty} n^2\beta(\theta_n)^{2/3}=\infty. \qquad (70)
$$

As Proposition 3.3 yields $n^2\beta(\theta_n)^{2/3}\sim n^2\tau(\theta_n)^{2/3}=\left(n^3\tau(\theta_n)\right)^{2/3}$, the desired conclusion (70) follows under the condition (31). ∎

Proof of Proposition 5.1. Fix positive integers K and P such that $K\le P$, and $n=3,4,\ldots$. By exchangeability and the binary nature of the rvs involved we readily obtain

$$
\begin{aligned}
&\mathbb{E}\left[T_n(\theta)^2\right]\\
&\quad=\mathbb{E}\left[T_n(\theta)\right]+\binom{n}{3}\binom{3}{2}\binom{n-3}{1}\mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right]\\
&\qquad+\binom{n}{3}\binom{3}{1}\binom{n-3}{2}\mathbb{E}\left[\chi_{123}(\theta)\chi_{145}(\theta)\right]\\
&\qquad+\binom{n}{3}\binom{n-3}{3}\mathbb{E}\left[\chi_{123}(\theta)\chi_{456}(\theta)\right]. \qquad (71)
\end{aligned}
$$

Under the enforced independence assumptions the rvs $\chi_{123}(\theta)$ and $\chi_{456}(\theta)$ are independent and identically distributed. As a result,

$$
\mathbb{E}\left[\chi_{123}(\theta)\chi_{456}(\theta)\right]=\mathbb{E}\left[\chi_{123}(\theta)\right]\mathbb{E}\left[\chi_{456}(\theta)\right]=\beta(\theta)^2
$$

so that using the relation (16) we obtain

$$
\binom{n}{3}\binom{n-3}{3}\mathbb{E}\left[\chi_{123}(\theta)\chi_{456}(\theta)\right]=\frac{\binom{n-3}{3}}{\binom{n}{3}}\left(\mathbb{E}\left[T_n(\theta)\right]\right)^2. \qquad (72)
$$

On the other hand, with the help of (6) we readily check that the indicator rvs $\chi_{123}(\theta)$ and $\chi_{145}(\theta)$ are independent and identically distributed *conditionally* on $K_1(\theta)$ with

$$
\mathbb{P}\left[\chi_{123}(\theta)=1|K_1(\theta)\right]=\mathbb{P}\left[\chi_{123}(\theta)=1\right]=\beta(\theta). \qquad (73)
$$

As a similar statement applies to $\chi_{145}(\theta)$, we conclude that the rvs $\chi_{123}(\theta)$ and $\chi_{145}(\theta)$ are (unconditionally) independent and identically distributed with

$$
\mathbb{E}\left[\chi_{123}(\theta)\chi_{145}(\theta)\right]=\mathbb{E}\left[\chi_{123}(\theta)\right]\mathbb{E}\left[\chi_{145}(\theta)\right]=\beta(\theta)^2.
$$

Again by virtue of (16), this last observation yields

$$
\begin{aligned}
&\binom{n}{3}\binom{3}{1}\binom{n-3}{2}\mathbb{E}\left[\chi_{123}(\theta)\chi_{145}(\theta)\right]\\
&\quad=3\frac{\binom{n-3}{2}}{\binom{n}{3}}\cdot\left(\mathbb{E}\left[T_n(\theta)\right]\right)^2. \qquad (74)
\end{aligned}
$$

Substituting (72) and (74) into (71) establishes Proposition 5.1. ∎

E. A proof of Theorem 4.4

Throughout P and K are positive integers such that $K \leq P$, and fix $n = 3, 4, \ldots$. For each $i = 1, \ldots, n$, we introduce the index set

$$\mathcal{P}_{n,i} = \{(j,k) : \ 1 \leq j < k \leq n, \ j \neq i, \ k \neq i\}.$$

Next, define the count rvs $T_{n,i}(\theta)$ and $T^{\star}_{n,i}(\theta)$ by

$$T_{n,i}(\theta) = \sum_{(j,k) \in \mathcal{P}_{n,i}} \xi_{ij}(\theta)\xi_{ik}(\theta)\xi_{jk}(\theta)$$

and

$$T^{\star}_{n,i}(\theta) = \sum_{(j,k) \in \mathcal{P}_{n,i}} \xi_{ij}(\theta)\xi_{ik}(\theta).$$

The rv $T_{n,i}(\theta)$ counts the number of distinct triangles in $\mathbb{K}(n;\theta)$ which have node i as a vertex, while $T^{\star}_{n,i}(\theta)$ counts the number of distinct pairs of nodes which are both connected to node i in $\mathbb{K}(n;\theta)$. We also introduce the rv $D_i(\theta)$ as the degree of node i in $\mathbb{K}(n;\theta)$ given by

$$D_{n,i}(\theta) = \sum_{k=1, \ k \neq i}^{n} \xi_{ik}(\theta).$$

Observe that we have

$$\sum_{i=1}^{n} T_{n,i}(\theta) = 3T_n(\theta)$$

while

$$D_{n,i}(\theta)\left(D_{n,i}(\theta) - 1\right) = 2T^{\star}_{n,i}(\theta).$$

Under the condition

$$\sum_{i=1}^{n} D_{n,i}(\theta)\left(D_{n,i}(\theta) - 1\right) > 0,$$

the definition of $C^{\star}(\mathbb{K}(n;\theta))$ yields

$$\begin{aligned} C^{\star}(\mathbb{K}(n;\theta)) &= \frac{\sum_{i=1}^{n} T_{n,i}(\theta)}{\frac{1}{2}\sum_{i=1}^{n} D_{n,i}(\theta)\left(D_{n,i}(\theta) - 1\right)} \\ &= \frac{\sum_{i=1}^{n} T_{n,i}(\theta)}{\sum_{i=1}^{n} T^{\star}_{n,i}(\theta)} \end{aligned}$$

so that

$$C^{\star}(\mathbb{K}(n;\theta)) = \frac{3T_n(\theta)}{\sum_{i=1}^{n} T^{\star}_{n,i}(\theta)} \mathbf{1}\left[\sum_{i=1}^{n} T^{\star}_{n,i}(\theta) > 0\right]. \quad (75)$$

The desired conclusion (32) is now immediate from Lemma 5.2 and Lemma 5.3 established below. They deal with the a.s. convergence of the numerator and denominator (properly normalized) appearing in the ratio (75), respectively.

Lemma 5.2: For positive integers P and K such that $K \leq P$, we have

$$\lim_{n \to \infty} \frac{T_n(\theta)}{\binom{n}{3}} = \beta(\theta) \quad a.s. \quad (76)$$

Proof. Fix $n = 3, 4, \ldots$ and $\varepsilon > 0$. Markov's inequality already gives

$$\mathbb{P}\left[\left|\frac{T_n(\theta)}{\binom{n}{3}} - \beta(\theta)\right| > \varepsilon\right] \leq \varepsilon^{-2}\mathrm{Var}\left[\frac{T_n(\theta)}{\binom{n}{3}}\right]$$

as we recall (16). It is now plain from (58) that

$$\begin{aligned} &\mathrm{Var}\left[\frac{T_n(\theta)}{\binom{n}{3}}\right] \\ &= \mathbb{E}\left[\left(\frac{T_n(\theta)}{\binom{n}{3}}\right)^2\right] - \left(\frac{\mathbb{E}\left[T_n(\theta)\right]}{\binom{n}{3}}\right)^2 \\ &= \frac{\mathbb{E}\left[T_n(\theta)\right]}{\binom{n}{3}^2} + \left(\frac{\binom{n-3}{3}}{\binom{n}{3}} + 3\frac{\binom{n-3}{2}}{\binom{n}{3}} - 1\right) \cdot \left(\frac{\mathbb{E}\left[T_n(\theta)\right]}{\binom{n}{3}}\right)^2 \\ &\quad + 3(n-3)\binom{n}{3} \cdot \frac{\mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right]}{\binom{n}{3}^2} \\ &= \frac{\beta(\theta)}{\binom{n}{3}} + \left(\frac{\binom{n-3}{3}}{\binom{n}{3}} + 3\frac{\binom{n-3}{2}}{\binom{n}{3}} - 1\right) \cdot \beta(\theta)^2 \\ &\quad + \frac{3(n-3)}{\binom{n}{3}} \cdot \mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right] \end{aligned} \quad (77)$$

as we again make use of the expression (16).

With the help of (64) and (65), it is easy to see that

$$\lim_{n \to \infty} \mathrm{Var}\left[\frac{T_n(\theta)}{\binom{n}{3}}\right] = 0, \quad (78)$$

a fact which would readily imply a weaker form of (76) with a.s. convergence replaced by convergence in probability. However, elementary algebra on (77) shows that (78) takes place according to

$$\lim_{n \to \infty} n^2 \mathrm{Var}\left[\frac{T_n(\theta)}{\binom{n}{3}}\right] = C$$

with

$$C = 18\left(\mathbb{E}\left[\chi_{123}(\theta)\chi_{124}(\theta)\right] - \beta(\theta)^2\right) > 0.$$

As a result, for every $\varepsilon > 0$, we have

$$\sum_{n=3}^{\infty} \mathbb{P}\left[\left|\frac{T_n(\theta)}{\binom{n}{3}} - \beta(\theta)\right| > \varepsilon\right] \leq \frac{C'}{\varepsilon^2} \sum_{n=3}^{\infty} n^{-2} < \infty$$

for some $C' > C$, and the conclusion (76) follows by the Borel-Cantelli Lemma. ■

Lemma 5.3: For positive integers P and K such that $K \leq P$, we have

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^{\star}_{n,i}(\theta)}{\binom{n}{3}} = 3p(\theta)^2 \quad a.s. \quad (79)$$

Proof. Fix $n = 3, 4, \ldots$. Note that

$$\begin{aligned} T^{\star}_{n,1}(\theta) &= \sum_{j=2}^{n-1} \sum_{k=j+1}^{n} \xi_{1j}(\theta)\xi_{1k}(\theta) \\ &= \Phi_n(\xi_{12}(\theta), \ldots, \xi_{1n}(\theta)) \end{aligned} \quad (80)$$

where the mapping $\Phi_n : [0,1]^{n-1} \to \mathbb{R}_+$ is given by

$$\Phi_n(x_2, \ldots, x_n) = \sum_{\ell=2}^{n-1} \sum_{k=\ell+1}^{n} x_\ell x_k$$

$$= \sum_{\ell=2}^{n-1} x_\ell \left(\sum_{k=\ell+1}^{n} x_k \right) \qquad (81)$$

with $(x_2, \ldots, x_n)$ arbitrary in $[0,1]^{n-1}$. For each $j = 2, \ldots, n$, consider pairs of elements $(x_2, \ldots, x_n)$ and $(y_2, \ldots, y_n)$ in $[0,1]^{n-1}$ which differ only in the j^{th} component, i.e.,

$$x_\ell = y_\ell, \quad \ell \neq j, \quad \ell = 2, \ldots, n.$$

Under such conditions, it is easy to check that

$$\begin{aligned} & |\Phi_n(x_2, \ldots, x_n) - \Phi_n(y_2, \ldots, y_n)| \\ & \leq |x_j - y_j| \cdot \sum_{\ell=2,\ \ell \neq j}^{n-1} x_\ell \\ & \leq n-1. \end{aligned} \qquad (82)$$

Recall that the $(n-1)$ rvs $\{\xi_{1j}(\theta),\ j = 2, \ldots, n\}$ are i.i.d. Bernoulli rvs. In view of the constraints (82) we can now apply McDiarmid's inequality [17] (with $c_j = (n-1)$ for all $j = 2, \ldots, n-1$); see also Corollary 2.17 and Remark 2.28 in the monograph [15, p. 38]. Thus, for every $t > 0$ we find

$$\mathbb{P}\left[\left|T^\star_{n,1}(\theta) - \mathbb{E}\left[T^\star_{n,1}(\theta)\right]\right| > t\right] \leq 2e^{-\frac{2t^2}{(n-1)^3}} \qquad (83)$$

with

$$\begin{aligned} \mathbb{E}\left[T^\star_{n,1}(\theta)\right] &= \sum_{j=2}^{n-1} \sum_{k=j+1}^{n} \mathbb{E}\left[\xi_{1j}(\theta)\xi_{1k}(\theta)\right] \\ &= \sum_{j=2}^{n-1} (n-j) p(\theta)^2 \\ &= \frac{(n-1)(n-2)}{2} \cdot p(\theta)^2 \end{aligned} \qquad (84)$$

under the independence noted earlier.

With $\varepsilon > 0$ we now substitute

$$t = \frac{(n-1)(n-2)}{2}\varepsilon$$

into (83). Since

$$\frac{2t^2}{(n-1)^3} = \frac{(n-2)^2}{2(n-1)} \cdot \varepsilon^2 \sim \frac{n}{2} \cdot \varepsilon^2,$$

we obtain from (83) and (84) that

$$\mathbb{P}\left[\left| \frac{T^\star_{n,1}(\theta)}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \leq 2e^{-\frac{n}{2}(1+o(1))\varepsilon^2}. \qquad (85)$$

Since

$$\begin{aligned} \left| \frac{\sum_{i=1}^n T^\star_{n,i}}{\frac{n(n-1)(n-2)}{2}} - p(\theta)^2 \right| &= \left| \frac{1}{n} \sum_{i=1}^{n} \left(\frac{T^\star_{n,i}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right) \right| \\ &\leq \frac{1}{n} \sum_{i=1}^{n} \left| \frac{T^\star_{n,i}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right|, \end{aligned}$$

it is plain that

$$\begin{aligned} & \mathbb{P}\left[\left| \frac{\sum_{i=1}^n T^\star_{n,i}}{\frac{n(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \\ & \leq \mathbb{P}\left[\frac{1}{n} \sum_{i=1}^{n} \left| \frac{T^\star_{n,i}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \\ & \leq \mathbb{P}\left[\bigcup_{i=1}^{n} \left[\left| \frac{T^\star_{n,i}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \right] \\ & = \sum_{i=1}^{n} \mathbb{P}\left[\left| \frac{T^\star_{n,i}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \\ & = n\mathbb{P}\left[\left| \frac{T^\star_{n,1}}{\frac{(n-1)(n-2)}{2}} - p(\theta)^2 \right| > \varepsilon \right] \end{aligned}$$

where the equality before last follows by a union bound argument and the last equality is a consequence of exchangeability.

Invoking (85) (with $\frac{\varepsilon}{3}$ instead of ε) we get

$$\mathbb{P}\left[\left| \frac{\sum_{i=1}^n T^\star_{n,i}}{\binom{n}{3}} - 3p(\theta)^2 \right| > \varepsilon \right] \leq 2ne^{-\frac{n}{18}(1+o(1))\varepsilon^2}$$

with

$$\sum_{n=3}^{\infty} ne^{-\frac{n}{18}(1+o(1))\varepsilon^2} < \infty$$

for every $\varepsilon > 0$. The a.s. convergence (79) now follows by the Borel-Cantelli Lemma. ∎

ACKNOWLEDGEMENT

This work was supported in part by NSF Grant CCF-0729093 and in part by the Department of Electrical and Computer Engineering at Carnegie Mellon University. The paper was completed during Fall 2014 while A.M. Makowski was a Visiting Professor with the Department of Statistics of the Hebrew University of Jerusalem with the support of a fellowship from the Lady Davis Trust. The authors also thank a colleague (who wishes to remain anonymous) for suggestions that lead to a shorter proof of the one law in Theorem 4.2.

REFERENCES

[1] F. G. Ball, D. J. Sirl, and P. Trapman, "Epidemics on random intersection graphs," *The Annals of Applied Probability* **24**:3, pp. 1081-1128, 2014.

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

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

[4] F. Chung and L. Lu, "The diameter of sparse random graphs," *Advances in Applied Mathematics* **26** (2001), pp. 257-279.

[5] M. Deijfen and W. Kets, "Random intersection graphs with tunable degree distribution and clustering," *Probability in the Engineering and Informational Sciences* **23** (2009), pp. 661-674.

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

[7] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1960), pp. 17-61.

[8] 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 (2002), Washington (DC), November 2002.

[9] J. P. Gleeson, S. Melnik, and A. Hackett, A, "How clustering affects the bond percolation threshold in complex networks," *Physical Review E*, 066114, 2010.

[10] V. Gligor, A. Perrig, and J. Zhao, "Brief encounters with a random key graph," In *Proc. of the 17th Security Protocols Workshop (SPW 17)*, Cambridge, U.K, April 2009. Lecture Notes in Computer Science (LNCS), volume 7028. Springer Verlag.

[11] E. Godehardt and J. Jaworski "Two models of random intersection graphs for classification," in *Studies in Classification, Data Analysis and Knowledge Organization* **22**, Eds. O. Optiz and M. Schwaiger, Springer, Berlin (2003), pp. 67-82.

[12] E. Godehardt, J. Jaworski and K. Rybarczyk, "Random intersection graphs and classification," in *Studies in Classification, Data Analysis and Knowledge Organization* **33**, Eds. H.J. Lens and R,. Decker, Springer, Berlin (2007), pp. 67-74.

[13] A. Hackett, S. Melnik, and J. P. Gleeson, "Cascades on a class of clustered random networks," *Physical Review E*, 056107, 2011.

[14] X. Huang, S. Shao, H. Wang, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "The robustness of interdependent clustered networks," *Europhysics Letters*, 101(1), 18002, 2013.

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

[16] M.K. Karoński, E.R. Scheinerma, and K.B. Singer-Cohen, "On random intersection graphs: The subgraph problem," *Combinatorics, Probability and Computing* **8** (1999), pp. 131-159.

[17] C. McDiarmid, "On the method of bounded differences," in *Surveys in Combinatorics*, Cambridge University Press, Cambridge (UK), 1989. pp. 148-188.

[18] P. Marbach, "A lower-bound on the number of rankings required in recommender systems using collaborative filtering," in Proceedings of the 42nd Annual Conference on Information Sciences and Systems (CISS 2008), Princeton University, Princeton (NJ), March 2008.

[19] S. Milgram, "The small world problem," *Psychology Today* **2** (1967), pp. 60-67.

[20] J. C. Miller, "Percolation and epidemics in random clustered networks," *Physical Review E*, 020901, 2009.

[21] A. Mislove, M. Marcon, K.P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pp. 29-42, 2007.

[22] M. E. J. Newman, "Random Graphs with Clustering," *Phys. Rev. Lett.*, 058701, 2009.

[23] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review* **45** (2003), pp. 167-256.

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

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

[26] M. A. Serrano and M. Boguna, "Clustering in complex networks. i. general formalism, *Physical Review E* **74**, Nov 2006, p. 056114.

[27] K.B. Singer, *Random Intersection Graphs*, Ph.D. Thesis, The Johns Hopkins University, Baltimore (MD), 1995.

[28] D. Watts and S. Strogatz, "Collective dynamics of "small-world" networks," *Nature* **393** (1998), pp. 440-442.

[29] O. Yağan and A. M. Makowski, "Connectivity results for random key graphs," in Proceedings of the IEEE International Symposium on Information Theory (ISIT 2009), Seoul (Korea), June 2009.

[30] O. Yağan and A. M. Makowski, "On the existence of triangles in random key graphs," in Proceedings of the 47th Annual Allerton Conference on Communication, Control and Computing, Monticello (IL), September 2009.

[31] O. Yağan and A. M. Makowski, "Random key graphs – Can they be small worlds?," in Proceedings of the First Workshop on Applications of Graph Theory in Wireless Ad hoc Networks and Sensor Networks (GRAPH-HOC 2009), Chennai (India), December 2009.

[32] O. Yağan and A. M. Makowski, "Connectivity in random graphs induced by a key predistribution scheme: Small key pools," in Proceedings of the 44th Annual Conference on Information Sciences and Systems (CISS 2010), March 2010.

[33] 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.

[34] 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. 38103836.

APPENDIX
A PROOF OF PROPOSITION 3.3

Because $1 \leq K_n \leq {K_n}^2$ for all $n = 1, 2, \ldots$, the condition (19) implies both

$$\lim_{n\to\infty} \frac{1}{P_n} = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{K_n}{P_n} = 0. \tag{A.1}$$

Therefore, $\lim_{n\to\infty} P_n = \infty$, and for any $c > 0$, we have $cK_n < P_n$ for all n sufficiently large in $\mathbb{N}_0$ (dependent on c). Thus, we have $3K_n < P_n$ for all n sufficiently large in $\mathbb{N}_0$. On that range we can use the expression (13) to write

$$\beta(\theta_n) = (1 - q(\theta_n))^3 + q(\theta_n)^3 \left(1 - \frac{r(\theta_n)}{q(\theta_n)^2}\right).$$

As Lemma 3.2 already implies $q(\theta_n)^3 \sim 1$ and $(1 - q(\theta_n))^3 \sim \left(\frac{K_n^2}{P_n}\right)^3$, the asymptotic equivalence $\beta(\theta_n) \sim \tau(\theta_n)$ will be established if we show that

$$1 - \frac{r(\theta_n)}{q(\theta_n)^2} \sim \frac{K_n^3}{P_n^2}. \tag{A.2}$$

We proceed as follows: With positive integers K, P such that $3K \leq P$, we note that

$$\begin{aligned}
\frac{r(\theta)}{q(\theta)^2} &= \left(\frac{(P-2K)!}{(P-K)!}\right)^2 \cdot \frac{(P-2K)!}{(P-3K)!} \cdot \frac{P!}{(P-K)!} \\
&= \prod_{\ell=0}^{K-1} \left(\frac{P-2K-\ell}{P-K-\ell}\right) \cdot \prod_{\ell=0}^{K-1} \left(\frac{P-\ell}{P-K-\ell}\right) \\
&= \prod_{\ell=0}^{K-1} \left(1 - \left(\frac{K}{P-K-\ell}\right)^2\right),
\end{aligned}$$

and elementary bounding arguments yield

$$\begin{aligned}
1 - \left(1 - \left(\frac{K}{P-K}\right)^2\right)^K &\leq 1 - \frac{r(\theta)}{q(\theta)^2} \\
&\leq 1 - \left(1 - \left(\frac{K}{P-2K}\right)^2\right)^K.
\end{aligned}$$

Pick a scaling $P, K : \mathbb{N}_0 \to \mathbb{N}_0$ satisfying the equivalent conditions (19) and consider n sufficiently large in $\mathbb{N}_0$ so that $3K_n < P_n$. On that range, we replace θ by θ_n in the last chain of inequalities according to this scaling. A standard sandwich argument will yield the desired equivalence (A.2) if we show that

$$1 - \left(1 - \left(\frac{K_n}{P_n - cK_n}\right)^2\right)^{K_n} \sim \frac{K_n^3}{P_n^2}, \quad c = 1, 2. \tag{A.3}$$

To do so we proceed as follows: Fix $c = 1, 2$. With

$$A_n(c) = \left(\frac{K_n}{P_n - cK_n}\right), \quad n = 1, 2, \ldots$$

standard calculus yields

$$\begin{aligned}
&1 - \left(1 - \left(\frac{K_n}{P_n - cK_n}\right)^2\right)^{K_n} \\
&= K_n A_n(c)^2 \int_0^1 \left(1 - A_n(c)^2 t\right)^{K_n - 1} dt
\end{aligned}$$

on the appropriate range. The asymptotics

$$A_n(c)^2 = \left(\frac{K_n}{P_n - cK_n}\right)^2 \sim \left(\frac{K_n}{P_n}\right)^2 \tag{A.4}$$

and

$$K_n A_n(c)^2 \sim \frac{K_n^3}{P_n^2} \tag{A.5}$$

follow from (A.1), so that (A.3) will hold if we show that

$$\lim_{n\to\infty} \int_0^1 \left(1 - A_n(c)^2 t\right)^{K_n - 1} dt = 1. \tag{A.6}$$

In view of (A.4) we conclude from (A.1) that for all n sufficiently large in $\mathbb{N}_0$ we have $\sup_{0\le t\le 1} \left|1 - A_n(c)^2 t\right| \le 1$. Therefore, the Bounded Convergence Theorem will yield (A.6) as soon as we establish

$$\lim_{n\to\infty} \left(1 - A_n(c)^2 t\right)^{K_n - 1} = 1, \quad 0 \le t \le 1. \tag{A.7}$$

To that end, recall the decomposition

$$\log(1-x) = -\int_0^x \frac{1}{1-t} dt = -x - \Psi(x) \tag{A.8}$$

where

$$\Psi(x) = \int_0^x \frac{t}{1-t} dt, \quad 0 \le x < 1.$$

It is easy to check that

$$\lim_{x\downarrow 0} \frac{\Psi(x)}{x} = 0. \tag{A.9}$$

Fix n sufficiently large in $\mathbb{N}_0$ as required above. For each t in the interval $(0, 1]$, with the help of (A.8) we can write

$$\begin{aligned} &\left(1 - A_n(c)^2 t\right)^{K_n - 1} \\ &= e^{(K_n-1)\log\left(1 - A_n(c)^2 t\right)} \\ &= e^{-(K_n-1)A_n(c)^2 t - (K_n-1)\Psi(A_n(c)^2 t)}. \end{aligned} \tag{A.10}$$

Returning to (A.5), we use (19) and (A.1) to find

$$\lim_{n\to\infty} K_n A_n(c)^2 = \lim_{n\to\infty} \left(\frac{K_n^2}{P_n} \cdot \frac{K_n}{P_n}\right) = 0.$$

It is then plain that $\lim_{n\to\infty}(K_n - 1)A_n(c)^2 = 0$, whence

$$\begin{aligned} &\lim_{n\to\infty} (K_n - 1)\Psi(A_n(c)^2 t) \\ &= \lim_{n\to\infty} (K_n - 1)A_n(c)^2 t \cdot \frac{\Psi(A_n(c)^2 t)}{A_n(c)^2 t} = 0 \end{aligned}$$

with the help of (A.9) in the last step. Finally, letting n go to infinity in (A.10), we readily get (A.7) as desired. ∎