2021 60th IEEE Conference on Decision and Control (CDC)
December 13-15, 2021. Austin, Texas

Robustness of Random K-out Graphs

Eray Can Elumar Osman Yağan

2021 60th IEEE Conference on Decision and Control (CDC) | 978-1-6654-3659-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CDC45484.2021.9683492

***Abstract*— We consider a graph property known as r-robustness of the random K-out graphs. Random K-out graphs, denoted as $\mathbb{H}(n;K)$, are constructed as follows. Each of the n nodes select K distinct nodes uniformly at random, and then an edge is formed between these nodes. The orientation of the edges is ignored, resulting in an undirected graph. Random K-out graphs have been used in many applications including random (pairwise) key predistribution in wireless sensor networks, anonymous message routing in crypto-currency networks, and differentially-private federated averaging. r-robustness is an important metric in many applications where robustness of networks to disruptions is of practical interest, and r-robustness is especially useful in analyzing consensus dynamics. It was previously shown that consensus can be reached in an r-robust network for sufficiently large r even in the presence of some adversarial nodes. r-robustness is also useful for resilience against adversarial attacks or node failures since it is a stronger property than r-connectivity and thus can provide guarantees on the connectivity of the graph when up to $r-1$ nodes in the graph are removed. In this paper, we provide a set of conditions for K_n and n that ensure, *with high probability* (whp), the r-robustness of the random K-out graph.**

***Index Terms*— Robustness, random graphs, random K-out graphs, resilience, security, consensus dynamics**

I. Introduction

Random K-out graph is one of the earliest random graph models studied in literature [1], and has recently found a variety of applications such as analyzing the performance of the random *pairwise* key predistribution scheme in wireless sensor networks [2]–[6], decentralized learning [7], and anonymity preserving crypto-currency networks [8]. Random K-out graphs, denoted as $\mathbb{H}(n;K)$, are constructed over a set of n nodes as follows. Each node selects K distinct nodes uniformly at random, and then an undirected edge is formed between any pair of nodes if at least one selects the other.

One of the reasons that there are many applications envisioned for random K-out graphs is its ability to generate a connected topology even with a small number of nodes [9]; hence an important topic of interest is the connectivity and robustness of the network with respect to failures and disruption. A network is said to be connected if there exists a path of edges between every pair of vertices; and it was previously shown in [1], [2] that random K-out graphs are connected *whp* when $K \geq 2$ and not connected when $K = 1$; i.e.,

$$\lim_{n\to\infty} \mathbb{P}\left[\mathbb{H}(n;K) \text{ is connected}\right] = \begin{cases} 1 & \text{if} \quad K \geq 2, \\ 0 & \text{if} \quad K = 1. \end{cases} \tag{1}$$

E.C. Elumar and O. Yağan are with Department of Electrical and Computer Engineering and CyLab, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. Email: {eelumar@andrew.cmu.edu, oyagan@ece.cmu.edu}

A more generalized form of the connectivity property is r-connectivity. A graph is r-connected if it remains connected after the removal of any set of $r-1$ (or, fewer) nodes. It was previously shown in [1] that random K-out graphs are r-connected whp when $K \geq r$. Another graph property regarding robustness of a graph is r-robustness. A graph is r-robust if for every disjoint subset pair that partitions the graph, at least one node in one of these subsets is adjacent to at least r nodes in the other set; see Section II for a more formal definition for r-robustness. It was shown in [10] that if a graph is r-robust, it is at least r-connected. Thus, r-robustness is a stronger property than r-connectivity.

An important application where r-robustness property is useful is *consensus dynamics*, where some parameters of several agents get aligned after a sufficiently long period of local interactions. Consensus based approaches have been widely used in control theory and federated learning, which is an emerging field to train machine learning models in distributed systems [11], [12]. For example, in control theory, consensus dynamics have been used in a variety of applications in multi-agent systems such as flocking, swarming, synchronization of coupled oscillators and load balancing in networks [13]. There are also numerous studies involving consensus control under a variety of, application-specific, constraints. For example, in [14], a consensus control system was presented for a strongly connected network under the constraints of input saturation and communication time delay. In another example, [15], a consensus algorithm was presented for control of multi-agent systems in a fully connected network in the presence of adversaries.

To achieve resilience to adversaries in the control of a multi-agent system, network topology also plays an important role. For example, in [16], conditions on the connectivity of the network required to be resilient to misbehaving or faulty agents were derived for a linear consensus network. Furthermore, it was shown in [17] that there is a direct connection between the topology of the network and the performance/speed of the consensus dynamics, denoting the importance of comparing different graph models to find the topology most suitable for a given application. In [18], it was shown that network connectivity is not sufficient to characterize consensus when nodes use a certain class of local filtering rules. Instead, it was shown that consensus can be reached in graphs that are sufficiently robust. Thus, analyzing the r-robustness property of different graph models is particularly relevant for applications based on consensus dynamics. In federated learning applications of the consensus dynamics, random K-out graphs have received recent attention. In [7], random K-out graphs were used to construct

a communication graph in a differentially-private federated averaging scheme. From these examples, it is clear that the r-robustness of random K-out graphs will be of interest in future applications of random K-out graphs.

Even though r-robustness has been studied in several random graph models such as the Erdős-Rényi graph and the Barabási-Albert graph model [19], to our knowledge, r-robustness of random K-out graphs has not been studied before. The aim of this paper is to fill this gap in the literature and provide results on the r-robustness of random K-out graphs. Specifically, we provide a set of conditions for K_n and n that ensure, *with high probability* (whp), the r-robustness of random K-out graph and show that $K_n = O(r\log(r))$ is needed for r-robustness. We also compare this result with those obtained for an Erdős-Rényi graph with same average node degree, and determine that the random K-out graph becomes r-robust with fewer edges compared to an Erdős-Rényi graph.

II. Notations and r-robustness of a Graph

All random variables are defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and probabilistic statements are given with respect to the probability measure $\mathbb{P}$. The complement of an event A is denoted by A^{c}. The cardinality of a discrete set A is denoted by $|A|$. We refer to any mapping $\gamma : \mathbb{N}_0 \to \mathbb{N}_0$ as a *scaling* if it satisfies the condition $2 \leq \gamma_n < n, \quad n = 2, 3, \ldots$. All limits are understood with n going to infinity. If the probability of an event tends to one as $n \to \infty$, we say that it occurs with high probability (whp). When comparing the asymptotic behavior of sequences $\{a_n\}$ and $\{b_n\}$, the statements $a_n = o(b_n)$, $a_n = \omega(b_n)$, $a_n = O(b_n)$, $a_n = \Theta(b_n)$, and $a_n = \Omega(b_n)$, are defined in the standard Landau notation. The asymptotic equivalence $a_n \sim b_n$ is used to denote the fact that $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$.

The random K-out graph is defined on the vertex set $V := \{v_1, \ldots, v_n\}$ as follows. Let $\mathcal{N} := \{1, 2, \ldots, n\}$ denote the set of vertex labels. For each $i \in \mathcal{N}$, let $\Gamma_{n,i} \subseteq \mathcal{N} \setminus i$ denote the set of K_n labels corresponding to the nodes selected by v_i. The choices made by each node are independent of the other nodes, hence $\Gamma_{n,1}, \ldots, \Gamma_{n,n}$ are mutually independent. Distinct nodes v_i and v_j are adjacent, denoted by $v_i \sim v_j$ if at least one of them picks the other (since edges are undirected in a random K-out graph, the outcome is still the same if both v_i and v_j choose each other). Namely,

$$v_i \sim v_j \qquad \text{if} \qquad [j \in \Gamma_{n,i}] \ \vee \ [i \in \Gamma_{n,j}]. \tag{2}$$

The set of neighbors of node i is defined as $\mathcal{V}_i := \{j \in \mathcal{N} \setminus i : v_i \sim v_j\}$, and the degree of node i is denoted as $d_i = |\mathcal{V}_i|$. The random graph defined on the vertex set V through the adjacency relation (2) is called as a random K-out graph [20], [21] and is denoted by $\mathbb{H}(n; K_n)$.

In this paper, we are concerned with the r-robustness of random K-out graphs. To formally define r-robustness property, we use the following definitions introduced in [18].

Definition 2.1 (r-reachable set): For a graph $\mathbb{H}$ and a subset $\mathcal{S}$ of nodes $\mathcal{S} \subset \mathcal{N}$, we say $\mathcal{S}$ is an r-reachable set if $\exists i \in \mathcal{S} : |\mathcal{V}_i \setminus \mathcal{S}| \geq r$, where $r \in \mathbb{Z}^+$. In other words, set $\mathcal{S}$ is an r-reachable set if it contains a node that has at least r neighbors outside that set.

Definition 2.2 (r-robust graph): A graph $\mathbb{H}$ is an r-robust graph if for every pair of nonempty, disjoint subsets of $\mathcal{N}$ that partition $\mathcal{N}$, at least one of the subsets is r-reachable, where $r \in \mathbb{Z}^+$.

III. Main Results and Discussion

Our main result is presented as Theorem 3.1 below.

Let $P(n, K_n, r) = \mathbb{P}\left[\mathbb{H}(n; K_n) \text{ is } r\text{-robust}\right]$.

Theorem 3.1: *Consider a scaling $K : \mathbb{N}_0 \to \mathbb{N}_0$ and let*

$$t(r) = \frac{2r\left(\log(r) + \log(\log(r) + 1)\right)}{\log(2) + 1/2 - \log\left(1 + \frac{\log(2)+1/2}{2\log(r)+5/2+\log(2)}\right)} \tag{3}$$

be the threshold function. Then, for all $r \in \mathbb{Z}^+$ satisfying $r \geq 3$, we have

$$\lim_{n\to\infty} P(n, K_n, r) = 1, \quad \text{if} \quad K_n > t(r), \ \forall n.$$

A. Discussion

In Theorem 3.1, we establish a threshold for one-law of r-robustness in random K-out graphs, more specifically we find that a random K-out graph is r-robust when $K_n \sim c \cdot r\log(r)$ for large n and r, where $c > 0$ is a constant. The proof of this result is given in Section IV. We were not able to validate this result through computer simulation, since it was shown in [22] that determining r-robustness of a graph is a co-NP-complete problem.

To put this result in perspective, we compare it with the result from an Erdős-Rényi graph $G(n, p)$, one of the most commonly studied random graph models. In [22], it was shown that an Erdős-Rényi graph $G(n, p)$ is r-robust *whp* if $p_n = \frac{\log(n)+(r-1)\log(\log(n))+\omega(1)}{n}$, which translates to an average node degree of $< k_n > \sim \log(n) + (r-1)\log(\log(n))$. This shows that random K-out graphs can be ensured to be r-robust *whp* at an average node degree significantly smaller than the average node degree required for an Erdős-Rényi graph to be r-robust. Hence we can conclude that random K-out graphs are much more robust than Erdős-Rényi graphs in terms of the r-robustness property when both graphs have similar average node degree.

Another point worth mentioning is the comparison of the thresholds for r-robustness and r-connectivity. It was shown in [1] that random K-out graph $\mathbb{H}(n; K_n)$ is r-connected *whp* for large n when $K_n \geq r$. Since the threshold for r-robustness is $K_n \sim c \cdot r\log(r)$ for large n and r where $c > 0$ is a constant, there is only a factor of $\log(r)$ difference between them for large n and r. This result is as expected since r-robustness is a stronger property than r-connectivity.

B. Further Applications

Random K-out graphs have been useful for constructing distributed networks since they become connected efficiently (with as few edges as possible, compared to Erdős-Rényi graphs) [23]. As such, they have a wide variety of applications such as wireless sensor networks [2], decentralized

learning [7], and anonymity preserving crypto-currency networks [8].

Since we have shown that they also satisfy the r-robustness property with fewer edges compared to Erdős-Rényi graphs, they can also be used in applications that require r-robustness with as few edges as possible such as robust control of multi-agent systems, robust and differentially-private federated learning, as discussed in the introduction.

IV. A Proof of Theorem 3.1

To prove Theorem 3.1, we need to show that the random K-out graph $\mathbb{H}(n; K_n)$ is r-robust *whp* when $K_n > t(r)$, $\forall n$, for $r \in \mathbb{Z}^+$ satisfying $r \geq 3$, and with $t(r)$ defined as in (3).

First, let $\mathcal{E}_n(K_n, r; \mathcal{S})$ denote the event that $\mathcal{S} \subset V$ is an r-reachable set as per Definition 2.1. The event $\mathcal{E}_n(K_n, r; \mathcal{S})$ occurs if there exists at least one node in $\mathcal{S}$ that is adjacent to at least r nodes in $R = V \setminus \mathcal{S}$, the subset comprised of nodes outside the subset $\mathcal{S}$. Thus, we have

$$\mathcal{E}_n(K_n, r; \mathcal{S}) = \bigcup_{i \in \mathcal{N}_\mathcal{S}} \left\{ \left(\sum_{j \in \mathcal{N}_R} \mathbb{1}\{v_i \sim v_j\} \right) \geq r \right\}$$

with $\mathcal{N}_\mathcal{S}$, $\mathcal{N}_R$ denoting the set of labels of the vertices in $\mathcal{S}$ and R, respectively, and $\mathbb{1}\{\}$ denoting the indicator function. We are also interested in the complement of this event $\mathcal{E}_n(K_n, r; \mathcal{S})$, denoted as $(\mathcal{E}_n(K_n, r; \mathcal{S}))^{\mathrm{c}}$, which occurs if all nodes in $\mathcal{S}$ are adjacent to less than r nodes in $R = V \setminus \mathcal{S}$. This can be written as

$$(\mathcal{E}_n(K_n, r; \mathcal{S})^{\mathrm{c}}) = \bigcap_{i \in \mathcal{N}_\mathcal{S}} \left\{ \left(\sum_{j \in \mathcal{N}_R} \mathbb{1}\{v_i \sim v_j\} \right) < r \right\}.$$

Note that at least one subset in every disjoint subset pairs that partition V needs to be r-reachable per the definition of r-robustness, hence it is sufficient to show that every subset of V with size up to $\lfloor \frac{n}{2} \rfloor$ is r-reachable *with high probability* to prove r-robustness. Thus, let $\mathcal{Z}(K_n, r)$ denote the event that all non-empty subsets of the vertices with size up to $\lfloor \frac{n}{2} \rfloor$ is r-reachable. Namely, $\mathcal{Z}(K_n, r)$ is the event that for all subset pairs $\mathcal{S}$ and $V \setminus \mathcal{S}$ such that $\mathcal{S} \subset V$ and $|\mathcal{S}| \leq \lfloor \frac{n}{2} \rfloor$, the subset $\mathcal{S}$ is r-reachable. Thus, we have

$$\mathcal{Z}(K_n, r) = \bigcap_{S \in \mathcal{P}_n:\ |S| \leq \lfloor \frac{n}{2} \rfloor} \mathcal{E}_n(K_n, r; S),$$

where $\mathcal{P}_n$ is the collection of all non-empty subsets of V. Complementing both sides and using union bound, we get

$$\begin{aligned} \mathbb{P}\left[(\mathcal{Z}(K_n, r))^{\mathrm{c}}\right] &\leq \sum_{|S| \leq \lfloor \frac{n}{2} \rfloor} \mathbb{P}[(\mathcal{E}_n(K_n, r; \mathcal{S}))^{\mathrm{c}}] \\ &= \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \sum_{S \in \mathcal{P}_{n,m}} \mathbb{P}[(\mathcal{E}_n(K_n, r; \mathcal{S}))^{\mathrm{c}}], \end{aligned} \tag{4}$$

where $\mathcal{P}_{n,m}$ denotes the collection of all subsets of V with exactly m elements.

Let $\mathcal{S}_m$ be a subset of the vertex set V with size m, i.e. $\mathcal{S}_m \subset V$ and $|\mathcal{S}_m| = m$. Also define $R_m = V \setminus \mathcal{S}_m$, and define $\mathcal{E}_{n,m}(K_n, r)$ as the event that a random subset of V of size m is an r-reachable set. From the exchangeability of the node labels and associated random variables, we have

$$\mathbb{P}[\mathcal{E}_n(K_n, r; \mathcal{S}_m)] = \mathbb{P}[\mathcal{E}_{n,m}(K_n, r)], \quad \mathcal{S}_m \in \mathcal{P}_{n,m}.$$

$|\mathcal{P}_{n,m}| = \binom{n}{m}$, since there are $\binom{n}{m}$ subsets of V with m elements. Thus, we have

$$\sum_{\mathcal{S}_m \in \mathcal{P}_{n,m}} \mathbb{P}[(\mathcal{E}_n(K_n, r; \mathcal{S}_m))^{\mathrm{c}}] = \binom{n}{m} \mathbb{P}[(\mathcal{E}_{n,m}(K_n, r))^{\mathrm{c}}].$$

Substituting this into (4), we obtain

$$\mathbb{P}\left[(\mathcal{Z}(K_n, r))^{\mathrm{c}}\right] \leq \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{m} \mathbb{P}[(\mathcal{E}_{n,m}(K_n, r))^{\mathrm{c}}] \tag{5}$$

We will abbreviate

$$P_m := \binom{n}{m} \mathbb{P}[(\mathcal{E}_{n,m}(K_n, r))^{\mathrm{c}}],$$

$$P_Z := \mathbb{P}\left[(\mathcal{Z}(K_n, r))^{\mathrm{c}}\right].$$

Then, we can write

$$P_Z \leq \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} P_m \tag{6}$$

To find P_m, we first need to find the probability of a node in $\mathcal{S}_m$ being adjacent to d nodes in R_m, denoted as P_d. A node v in the subset $\mathcal{S}_m$ can be adjacent to d nodes in R_m if it chooses i nodes out of its K_n selections from the subset R_m, and at the same time $d - i$ nodes in R_m that were not selected by the node v select v, where $i \leq d$ and $d \leq K_n$. Then, P_d can be expressed as the sum of the probabilities of such events over all possible i values, leading to

$$\begin{aligned} P_d = \sum_{i=0}^{d} \frac{\binom{n-m}{i}\binom{m-1}{K_n-i}}{\binom{n-1}{K_n}} \left(\frac{K_n}{n-1}\right)^{d-i} \\ \cdot \left(1 - \frac{K_n}{n-1}\right)^{n-m-d} \binom{n-m-i}{d-i} \end{aligned} \tag{7}$$

Using this, the probability of a node in $\mathcal{S}_m$ being adjacent to less than r nodes in R_m, denoted as P_r can be calculated as $P_r = \sum_{d=0}^{r-1} P_d$. Furthermore, the probability that all nodes in $\mathcal{S}_m$ are adjacent to less than r nodes in R_m, denoted as $P_{m,r}$ can be found as $P_{m,r} \leq (P_r)^m$. Hence, $P_m = \binom{n}{m} P_{m,r} \leq \binom{n}{m} (\sum_{d=0}^{r-1} P_d)^m$. Using (7) we obtain

$$\begin{aligned} P_m \leq \binom{n}{m} \Bigg(\sum_{d=0}^{r-1} \sum_{i=0}^{d} \frac{\binom{n-m}{i}\binom{m-1}{K_n-i}}{\binom{n-1}{K_n}} \left(\frac{K_n}{n-1}\right)^{d-i} \\ \cdot \left(1 - \frac{K_n}{n-1}\right)^{n-m-d} \binom{n-m-i}{d-i} \Bigg)^m \end{aligned} \tag{8}$$

To prove Theorem 3.1, we need to show that $\lim_{n \to \infty} P_Z = 0$ when $K_n > t(r)$, $\forall n$, $r \in \mathbb{Z}^+$ satisfying $r \geq 3, r^2 = o(n)$, and $t(r)$ defined as in (3). Depending on the value of m, we need to consider three cases when evaluating P_m, which are the (i) $m \leq K_n + 1 - r$, (ii) $K_n + 2 \leq m$, and (iii) $K_n - r + 2 \leq m \leq K_n + 1$ case. We start with the first case.

(i). Case 1 ($m \leq K_n + 1 - r$):

In this case it can easily be seen that even if a node $v \in \mathcal{S}_m$ chooses all the other $m-1$ nodes in $\mathcal{S}_m$, it still needs to make $K_n + 1 - m$ more selections, which need to be chosen from the subset R_m. Since $r \leq K_n + 1 - m$, every node in $\mathcal{S}_m$ will be adjacent to at least r nodes in R_m, and hence $\mathcal{S}_m$ will be guaranteed to be r-reachable. Thus, we have

$$P_m = 0, \quad m \leq K_n + 1 - r \tag{9}$$

(ii). Case 2 ($K_n - r + 2 \leq m \leq K_n + 1$):

When $K_n - r + 2$, any node $v \in \mathcal{S}_m$ is guaranteed to select at least $K_n + 1 - m$ nodes in R_m, hence the summation variable i (which tracks the number of selections made by v in R_m), in (8) should start from $K_n + 1 - m$ in this case. Thus, for this case we have

$$\begin{aligned} P_m &\leq \binom{n}{m} \Bigg(\sum_{d=K_n+1-m}^{r-1} \sum_{i=K_n+1-m}^{d} \frac{\binom{n-m}{i}\binom{m-1}{K_n-i}}{\binom{n-1}{K_n}} \\ &\quad \cdot \left(\frac{K_n}{n-1}\right)^{d-i} \left(1 - \frac{K_n}{n-1}\right)^{n-m-d} \binom{n-m-i}{d-i} \Bigg)^m \\ &\leq \binom{n}{m} \Bigg(\sum_{d=0}^{r-1} \sum_{i=0}^{d} \frac{\binom{n-m}{i}\binom{m-1}{K_n-i}}{\binom{n-1}{K_n}} \left(\frac{K_n}{n-1}\right)^{d-i} \\ &\quad \cdot \left(1 - \frac{K_n}{n-1}\right)^{n-m-d} \binom{n-m-i}{d-i} \Bigg)^m \end{aligned} \tag{10}$$

Hence, we can still use (8) to find an upper bound on P_m.

(iii). Case 3 ($K_n + 2 \leq m$):

In this case it is not guaranteed that a node $v \in \mathcal{S}_m$ will select a node in R_m, so we can directly use (8) to find an upper bound on P_m.

In light of these three cases, we can see that P_m only needs to be evaluated for $m \geq K_n - r + 2$. Following standard upper bounds will be used in the rest of the proof.

$$\binom{n}{l} \leq \left(\frac{n}{l}\right)^l \left(\frac{n}{n-l}\right)^{n-l}, \quad \forall l = 1, \ldots, n \tag{11}$$

$$\binom{n}{l} \leq \left(\frac{en}{l}\right)^l, \quad \forall l = 1, \ldots, n \tag{12}$$

$$1 \pm x \leq e^{\pm x}, \quad \forall x \tag{13}$$

Using (11), we have

$$\begin{aligned} \frac{\binom{n-m}{i}\binom{m-1}{K_n-i}\binom{n-m-i}{d-i}}{\binom{n-1}{K_n}} &\leq \left(\frac{n-m}{i}\right)^i \left(\frac{n-m}{n-m-i}\right)^{n-m-i} \\ &\cdot \left(\frac{m-1}{K_n-i}\right)^{K_n-i} \left(\frac{m-1}{m-1+i-K_n}\right)^{m-1+i-K_n} \\ &\cdot \left(\frac{n-m-i}{d-i}\right)^{d-i} \left(\frac{n-m-i}{n-m-d}\right)^{n-m-d} \left(\frac{K_n}{n-1}\right)^{K_n} \\ &\cdot \left(1 - \frac{K_n}{n-1}\right)^{n-1-K_n} \end{aligned} \tag{14}$$

Substituting (14) into (8), we have

$$\begin{aligned} P_m &\leq \binom{n}{m} \Bigg[\sum_{d=0}^{r-1} \sum_{i=0}^{d} \left(\frac{(n-1)(d-i)(K_n-i)}{(m-1+i-K_n)iK_n} \right)^i \\ &\cdot \left(\frac{(n-m-d)K_n}{(n-1-K_n)(d-i)} \right)^d \left(\frac{n-1-K_n}{n-1} \right)^{2n-m-K_n-1} \\ &\cdot \left(\frac{(m-1)K_n}{(n-1)(K_n-i)} \right)^{K_n} \left(\frac{n-m}{n-m-d} \right)^{n-m} \\ &\cdot \left(\frac{m-1}{m-1+i-K_n} \right)^{m-1-K_n} \Bigg]^m \end{aligned} \tag{15}$$

We now use upper or lower bounds on the summation variables i and d. Note that the i and $d-i$ terms in the denominator become zero when $i = 0$ or $d = i$. Since these terms were obtained from the upper bound on combination, and since $\binom{n}{0} = 1$, lower bound of i and $d-i$ terms on the denominator is taken as 1. It can still be shown that the upper bound found below using this correction still holds for all possible i or d values. After this, we have

$$\begin{aligned} P_m &\leq \binom{n}{m} \Bigg[\sum_{d=0}^{r-1} \sum_{i=0}^{d} \left(\frac{d(n-1)}{m-1-K_n} \right)^i \\ &\cdot \left(\frac{K_n(n-m-d)}{n-1-K_n} \right)^d \left(\frac{n-1-K_n}{n-1} \right)^{2n-m-K_n-1} \\ &\cdot \left(\frac{(m-1)K_n}{(n-1)(K_n+1-r)} \right)^{K_n} \left(\frac{n-m}{n-m-d} \right)^{n-m} \\ &\cdot \left(\frac{m-1}{m-1-K_n} \right)^{m-1-K_n} \Bigg]^m \\ &\leq \binom{n}{m} \Bigg[\sum_{d=0}^{r-1} (d+1) \left(\frac{dK_n(n-1)(n-m-d)}{(m-1-K_n)(n-1-K_n)} \right)^d \\ &\cdot \left(\frac{n-1-K_n}{n-1} \right)^{2n-m-K_n-1} \left(\frac{m-1}{m-1-K_n} \right)^{m-1-K_n} \\ &\cdot \left(\frac{(m-1)K_n}{(n-1)(K_n+1-r)} \right)^{K_n} \left(\frac{n-m}{n-m-d} \right)^{n-m} \Bigg]^m \end{aligned} \tag{16}$$

Noting that $d \leq r-1$ and $\left(\frac{n-m}{n-m-d}\right)^{n-m-d} \leq e^d \leq e^{r-1}$, we have

$$\begin{aligned} P_m &\leq \binom{n}{m} \Bigg[\sum_{d=0}^{r-1} r \left(\frac{(r-1)K_n(n-1)(n-m)}{(m-1-K_n)(n-1-K_n)} \right)^d \\ &\cdot \left(\frac{n-1-K_n}{n-1} \right)^{2n-m-K_n-1} \left(\frac{m-1}{m-1-K_n} \right)^{m-1-K_n} \\ &\cdot \left(\frac{(m-1)K_n}{(n-1)(K_n+1-r)} \right)^{K_n} e^{r-1} \Bigg]^m \end{aligned} \tag{17}$$

$$\leq \binom{n}{m}\left[e^{r-1}r^2\left(\frac{(r-1)K_n(n-m)}{(m-1-K_n)}\right)^{r-1}\right.$$
$$\cdot\left(1-\frac{K_n}{n-1}\right)^{2n-m-K_n-1}\left(1+\frac{K_n}{m-1-K_n}\right)^{m-1-K_n}$$
$$\left.\cdot\left(1+\frac{K_n}{n-1-K_n}\right)^{r-1}\left(\frac{mK_n}{n(K_n+1-r)}\right)^{K_n}\right]^m \tag{18}$$

$$\leq \binom{n}{m}\left[e^{r-1}r^2K_n^{r-1}(r-1)^{r-1}e^{\frac{K_n(r-1)}{n-1-K_n}}e^{\frac{K_n(m+K_n-1)}{n-1}}\right.$$
$$\left.\cdot\left(\frac{n-m}{m-1-K_n}\right)^{r-1}\left(\frac{mK_n}{n(K_n+1-r)}\right)^{K_n}e^{-K_n}\right]^m \tag{19}$$

$$\leq \left[e^r r^2K_n^{r-1}(r-1)^{r-1}e^{\frac{K_n(r-1)}{n-1-K_n}}e^{\frac{mK_n}{n}}e^{\frac{K_n^2}{n-1}}e^{-K_n}\right.$$
$$\left.\cdot\left(\frac{n-m}{m-1-K_n}\right)^{r-1}\left(\frac{K_n}{K_n+1-r}\right)^{K_n}\left(\frac{m}{n}\right)^{K_n-1}\right]^m \tag{20}$$

$$\leq \exp\left(m\left[r+\frac{K_n(r-1)}{n-1-K_n}+\frac{mK_n}{n}+\frac{K_n^2}{n-1}\right.\right.$$
$$+2\log(r)+(r-1)\log(K_n)+(r-1)\log(r-1)$$
$$+(K_n-1)\log\left(\frac{m}{n}\right)+(r-1)\log\left(\frac{n-m}{m-1-K_n}\right)$$
$$\left.\left.+K_n\log\left(\frac{K_n}{K_n+1-r}\right)-K_n\right]\right) \tag{21}$$

We will show that the right side of the above expression goes to zero as n goes to infinity. Let

$$A_{n,r,m} := \exp\left(m\left[r+\frac{K_n(r-1)}{n-1-K_n}+\frac{mK_n}{n}+\frac{K_n^2}{n-1}\right.\right.$$
$$+2\log(r)+(r-1)\log(K_n)+(r-1)\log(r-1)$$
$$+(K_n-1)\log\left(\frac{m}{n}\right)+(r-1)\log\left(\frac{n-m}{m-1-K_n}\right)$$
$$\left.\left.+K_n\log\left(\frac{K_n}{K_n+1-r}\right)-K_n\right]\right)$$

We write

$$P_Z \leq \sum_{m=1}^{K_n+1-r}P_m+\sum_{m=K_n+2-r}^{\log(n)}A_{n,r,m}+\sum_{m=\log(n)}^{\lfloor\frac{n}{2}\rfloor}A_{n,r,m}$$
$$\leq \sum_{m=K_n+2-r}^{\log(n)}A_{n,r,m}+\sum_{m=\log(n)}^{\lfloor\frac{n}{2}\rfloor}A_{n,r,m} := Q_1+Q_2,$$

since the first summation $\sum_{m=1}^{K_n+1-r}P_m$ is zero as determined in Case 1. Thus, we need to show that both Q_1 and Q_2 go to zero as $n\to\infty$. We start with the first summation Q_1. Assume as in the statement of Theorem 3.1 that

$$K_n > \frac{2r\left(\log(r)+\log(\log(r)+1)\right)}{\log(2)+1/2-\log\left(1+\frac{\log(2)+1/2}{2\log(r)+5/2+\log(2)}\right)}, \quad \forall n \tag{22}$$

Taking $n\to\infty$ and using the fact that $m\leq\log(n)$ from the limits of the summation, and also using the given condition $r^2=o(n)$, we have

$$A_{n,r,m} \leq \exp\left(m\left[o(1)+r+2\log(r)+(r-1)\log(K_n)\right.\right.$$
$$+(r-1)\log(r-1)-K_n+K_n\log\left(\frac{K_n}{K_n+1-r}\right)$$
$$\left.\left.+(K_n-1)\log(m)+(r-1)\log(m-1-K_n)\right]\right)$$
$$\cdot\exp[-m(K_n-r)\log(n)]$$
$$A_{n,r,m} \leq o\left(e^{m(K_n-r)\log(n)}\right)e^{-m(K_n-r)\log(n)} \tag{23}$$

It can easily be seen that the $e^{-m(K_n-r)\log(n)}$ term dominates this expression for large n, hence $\lim_{n\to\infty}A_{n,r,m}=0$. Thus, for large n, we have

$$Q_1 \leq \sum_{m=K_n+2-r}^{\log(n)}(A_{n,r,m})^r \leq \sum_{m=1}^{\infty}(A_{n,r,m})^r = \frac{A_{n,r,m}}{1-A_{n,r,m}} \tag{24}$$

where the geometric sum converges by virtue of $\lim_{n\to\infty}A_{n,r,m}=0$. Using this once again, it is clear from the last expression that $\lim_{n\to\infty}Q_1=0$.

Now, similarly consider the second summation S_2. Again assume K_n is larger than the threshold given in (22), and use the given condition $r^2=o(n)$. For large n, we have

$$Q_2 \leq \sum_{m=\log(n)}^{\lfloor\frac{n}{2}\rfloor}\exp\left(m\left[r+\frac{K_n(r-1)}{n-1-K_n}+\frac{mK_n}{n}\right.\right.$$
$$+2\log(r)+(r-1)\log(K_n)+(r-1)\log(r-1)$$
$$+(K_n-r)\log\left(\frac{m}{n}\right)+(r-1)\log\left(\frac{m(n-m)}{n(m-1-K_n)}\right)$$
$$\left.\left.+\frac{K_n^2}{n-1}+K_n\log\left(\frac{K_n}{K_n+1-r}\right)-K_n\right]\right) \tag{25}$$

Since the sequence $a_m=\frac{mK_n}{n}+(K_n-r)\log\left(\frac{m}{n}\right)+(r-1)\log\left(\frac{m(n-m)}{n(m-1-K_n)}\right)$ is monotonically increasing with m in the range $\log(n)\leq m\leq\lfloor\frac{n}{2}\rfloor$, we can use the term with $m=\frac{n}{2}$ as an upper bound for this sequence. Hence, we have

$$Q_2 \leq \sum_{m=\log(n)}^{\lfloor\frac{n}{2}\rfloor}\exp\left(m\left[o(1)+\log(2)+r-\frac{K_n}{2}\right.\right. \tag{26}$$
$$+(r-1)\log(K_n)+(r+1)\log(r)$$
$$\left.\left.-K_n\log(2)+K_n\log\left(\frac{K_n}{K_n+1-r}\right)\right]\right)$$

Since $K_n\geq\frac{2r\log r+2r}{\log(2)+1/2}$, we have

$$\log\left(\frac{K_n}{K_n+1-r}\right)\leq\log\left(1+\frac{\log(2)+1/2}{2\log(r)+5/2+\log(2)}\right)$$

Plugging thus back to (26), we have

$$A_{n,m,r} \leq \exp\Big([o(1) + \log(2) + r + (r-1)\log(K_n) + (r+1)\log(r) - K_n\Big[\frac{1}{2} + \log(2) - \log\Big(1 + \frac{\log(2) + \frac{1}{2}}{2\log(r) + \frac{5}{2} + \log(2)}\Big)\Big]\Big]\Big) \tag{27}$$

Substituting for K_n via (22) and taking the limit as $n \to \infty$ it can be seen that $\lim_{n\to\infty} A_{n,m,r} = 0$ upon noting that $\log(K_n) = \log(r) + \log(2) + \log(\frac{K_n}{2r})$ and $(r-1)\log\left(\frac{t(r)}{2r}\right) \leq 2r\log(\log(r)) + r - \log(2)$.
Similar to the case in Q_1, we have:

$$Q_2 \leq \sum_{m=\log(n)}^{\lfloor \frac{n}{2} \rfloor} (A_{n,r,m})^r \leq \frac{(A_{n,r,m})^{\log(n)}}{1 - A_{n,r,m}} \tag{28}$$

where the geometric sum converges by virtue of $\lim_{n\to\infty} A_{n,r,m} = 0$, leading to Q_2 converging to zero as n gets large.

We establish that P_Z converges to zero as n goes to infinity, since $P_Z \leq Q_1 + Q_2$, and both Q_1 and Q_2 converge to zero when n is large. Since $P_Z = \mathbb{P}\left[(\mathcal{Z}(K_n, r))^{\mathrm{c}}\right] = 1 - P(n, K_n, r)$, this result yields the desired conclusion $\lim_{n\to\infty} P(n, K_n, r) = 1$ when $K_n > t(r)$, $\forall n$ and $r \in \mathbb{Z}^+$ satisfying $r \geq 3$ and $r^2 = o(n)$ as in Theorem 3.1. This concludes the proof of Theorem 3.1.

V. Conclusions

In this work, we provide a set of results for K_n and n that ensure, *with high probability*, the r-robustness of the random K-out graph $\mathbb{H}(n; K_n)$. Using our result, we compare the mean node degree of a random K-out graph with the mean node degree of an Erdős-Rényi graph at the threshold value required to ensure r-robustness *whp*, and determine that random K-out graphs attain r-robustness at a significantly lower mean node degree value compared to Erdős-Rényi graphs. This result reinforces the usefulness of random K-out graphs in applications that require a certain degree of robustness such as federated learning, consensus dynamics, robust control of multi-agent systems and wireless sensor networks.

Acknowledgements

This work has been supported in part by the Office of Naval Research through Grant #N00014-21-1-2547, a Pennsylvania Infrastructure Technology Alliance (PITA) grant, and by the CyLab@IoT Initiative.

References

[1] T. I. Fenner and A. M. Frieze, "On the connectivity of random m-orientable graphs and digraphs," *Combinatorica*, vol. 2, no. 4, pp. 347–359, Dec 1982.

[2] O. Yağan and A. M. Makowski, "On the connectivity of sensor networks under random pairwise key predistribution," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5754–5762, Sept 2013.

[3] ——, "Modeling the pairwise key predistribution scheme in the presence of unreliable links," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1740–1760, March 2013.

[4] F. Yavuz, J. Zhao, O. Yağan, and V. Gligor, "Designing secure and reliable wireless sensor networks under a pairwise key predistribution scheme," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 6277–6283.

[5] R. Eletreby and O. Yağan, "Connectivity of inhomogeneous random K-out graphs," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 7067–7080, 2020.

[6] M. Sood and O. Yağan, "On the size of the giant component in inhomogeneous random k-out graphs," *arXiv preprint arXiv:2009.01610*, 2020, to appear in 59th IEEE Conference on Decision and Control (CDC 2020), December 2020.

[7] C. Sabater, A. Bellet, and J. Ramon, "Distributed differentially private averaging with improved utility and robustness to malicious parties," 2020.

[8] G. Fanti, S. B. Venkatakrishnan, S. Bakshi, B. Denby, S. Bhargava, A. Miller, and P. Viswanath, "Dandelion++: Lightweight cryptocurrency networking with formal anonymity guarantees," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 2, no. 2, pp. 29:1–29:35, Jun. 2018.

[9] M. Sood and O. Yağan, "Tight bounds for the probability of connectivity in random k-out graphs," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.

[10] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 4, pp. 766–781, 2013.

[11] S. Savazzi, M. Nicoli, and V. Rampa, "Federated learning with cooperating devices: A consensus approach for massive iot networks," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4641–4654, 2020.

[12] Y. Li, C. Chen, N. Liu, H. Huang, Z. Zheng, and Q. Yan, "A blockchain-based decentralized federated learning framework with committee consensus," *IEEE Network*, vol. 35, no. 1, pp. 234–241, 2021.

[13] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[14] J. Wang, X. Li, Y. Li, X. Li, and X. Luo, "Event-based consensus control of multi-agent systems with input saturation constraint," in *2018 Chinese Control And Decision Conference (CCDC)*, 2018, pp. 4694–4699.

[15] H. LeBlanc and X. Koutsoukos, "Consensus in networked multi-agent systems with adversaries," 01 2011, pp. 281–290.

[16] F. Pasqualetti, A. Bicchi, and F. Bullo, "On the security of linear consensus networks," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, 2009, pp. 4894–4901.

[17] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[18] H. Zhang and S. Sundaram, "Robustness of information diffusion algorithms to locally bounded adversaries," in *2012 American Control Conference (ACC)*, 2012, pp. 5855–5861.

[19] H. Zhang and S. Sundaram, "Robustness of complex networks with implications for consensus and contagion," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 3426–3432.

[20] B. Bollobás, *Random graphs*. Cambridge university press, 2001.

[21] A. Frieze and M. Karoński, *Introduction to random graphs*. Cambridge University Press, 2016.

[22] H. Zhang, E. Fata, and S. Sundaram, "A notion of robustness in complex networks," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 310–320, 2015.

[23] E. C. Elumar, M. Sood, and O. Yağan, "On the connectivity and giant component size of random k-out graphs under randomly deleted nodes," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 2572–2577.