Internet Mathematics Vol. 4, No. 2–3: 199–223

Infinite Limits and Adjacency Properties of a Generalized Copying Model

Anthony Bonato and Jeannette Janssen

Abstract. We present a new model for self-organizing networks such as the World Wide Web graph and analyze its limit behavior. In the model, new vertices are introduced over time that copy the neighborhood structure of existing vertices, and a certain number of extra edges may be added to the new vertex that randomly join to any of the existing vertices. A function ρ parameterizes the number of extra edges. We study the model by considering the infinite limit graphs it generates. The limit graphs satisfy with high probability certain adjacency properties similar to but not as strong as the ones satisfied by the infinite random graph. We prove that the strength of the adjacency properties satisfied by the limit are governed by the choice of ρ. We describe certain infinite deterministic graphs that arise naturally from our model and that embed in all graphs generated by the model.

1. Introduction

Many of the real-world networks that are the focus of study today, such as the World Wide Web graph or the network of protein–protein interactions in a living cell, are *self-organizing.* In self-organizing networks, each vertex acts as an independent agent, basing its decision on how to link to the existing network on local knowledge. As a result, the neighborhood of a new vertex will often be an imperfect copy of the neighborhood of an existing vertex. Both the *copying models* [Adler and Mitzenmacher 01, Kumar et al. 00, Krapivsky and Redner 05] of the web graph, and the *duplication models* [Chung et al. 03, Kim et al. 02] of biological networks incorporate this notion of copying in their definitions.

1542-7951/07 $0.50 per page

The graphs generated by these models, although different, share some similar properties, such as power-law degree distributions. For additional information on self-organizing networks, the reader is directed to [Chung and Lu 06].

We introduce a new model that, in a certain, sense generalizes and unifies the copying and duplication models. The three parameters of the model $\text{Copy}(p, \rho, H)$ are a *copying probability* $p \in [0, 1]$, an *extra edge function* $\rho : \mathbb{N} \to \mathbb{N}$, and a finite *initial graph* H. The model describes a random graph process over a countable sequence of discrete time-steps indexed by $t \in \mathbb{N}$. Let $t_0 = |V(H)|$.

1. At $t = t_0$, set $G_{t_0} = H$.
2. For a fixed $t > t_0$, assume that G_{t-1} has been defined, is finite, and contains G_{t_0} as an induced subgraph. To form G_t, add a vertex v_t to G_{t-1} and choose its neighbors as follows.
   (a) Choose an existing vertex u_t from G_{t-1} uniformly at random (u.a.r.). The vertex u_t is called the *copy vertex*.
      i. For each neighbor w of u_t, independently add an edge from v_t to w with probability p. These are called *copy edges*.
      ii. Choose a set of $\rho(t)$ vertices from $V(G_{t-1})$ u.a.r., and add edges from v_t to each of these vertices. The latter edges are called *extra edges*.
   (b) Make G_t simple by deleting any multiple edges.

If $\rho(t) = 0$, then the graphs G_t generated by the model $\text{Copy}(p, \rho, H)$ correspond exactly to the graphs generated by the duplication model. (The duplication model we refer to is the one from [Chung et al. 03], which is different than the one described in [Chung and Lu 06]. In the duplication model from [Chung and Lu 06], at each time-step an edge is added between the new vertex and the copy vertex.) If $\rho(t)$ is constant and $p > 0$, then the graphs G_t are undirected analogues of graphs generated by the copying model. Note that for all $t \in \mathbb{N}, |V(G_t)| = t$. To simplify the discussion, we require that $\rho(t)$ be an integer-valued, non-decreasing function such that, for some $\alpha < 1$, $\rho(t) \leq \alpha t$ for all $t \geq t_0$. Moreover, unless otherwise stated, we assume that $\rho(t) = \Theta(t^s)$ for some $s \in [0, 1]$.

We study the infinite limits generated by this model, that is, the infinite graphs that result when time goes to infinity. Analyzing models by considering the infinite limit is a common technique in the natural sciences. Limit behavior can highlight certain properties of a model and point to significant differences and similarities among various models. In particular, the existence of a unique limit indicates coherent behavior of the model, while many distinct limits suggest a sensitivity to initial conditions that is an indicator of chaos. The use of infinite limits to study random graph processes was first proposed by the authors in [Bonato and Janssen 04a] and was studied in the context of the preferential attachment model in [Kleinberg and Kleinberg 05].

In [Bonato and Janssen 04a] the authors proved that limits of a deterministic version of the copying model satisfy a local version of the so-called n-existentially closed (n-e.c.) adjacency properties. A graph G is *n-e.c.* if for each pair of disjoint sets of vertices A and B such that $|A \cup B| = n$ (with one of A or B possibly empty), there is a vertex z not in $A \cup B$ joined to the vertices of A and not to the vertices of B. In other words, all extensions of n-element sets of vertices exist in the graph. We say that z is *correctly joined* or *c.j.* to A and B. The n-e.c. adjacency property and its variants have since been studied by many authors; see the survey [Bonato, to appear]. The unique isomorphism type of graph that is n-e.c. for all finite n is called the *infinite random graph* or *Rado graph*, written R.

The goal of the current paper is to study the generalized copy model $\text{Copy}(p, \rho, H)$ via its infinite limit. Our main tools are the n-e.c. adjacency properties, which measure in a certain sense how random an infinite graph is. One of our main results gives a "threshold" for the value of n for which the limit satisfies the n-e.c. property with probability 1. This value depends both on the copy probability p and on the order of the extra-edge function, and it suggests a subtle balance between the random behavior of the extra edges and the locally structured copy behavior. In the case when $\rho = 0$ and $p \in (0, 1)$, we define a new type of infinite graph, R_H, which with probability 1 is an induced subgraph of limits generated by $\text{Copy}(p, 0, H)$. In addition, we define another class of infinite graphs that are, with probability 1, induced subgraphs of limits generated by $\text{Copy}(p, \rho, H)$ for any choice of p and ρ. While our results are of mathematical interest in their own right, we think they serve as another step toward the use of infinite limits as tools for the study of models of self-organizing networks.

The main results, Theorems 2.2 and 2.3, are stated and proved in the next section. The proofs of these theorems follow from Lemmas 2.4 and 2.5, which give order bounds on the number of common neighbors of a set of given size. The lemmas are stated in the next section, while their rather technical proofs are deferred to Section 4. In Theorem 2.2, we prove that graphs generated by $\text{Copy}(p, \rho, H)$ are n-e.c. with probability 1, with n depending on the choice of ρ. In Theorem 2.3, we prove that, with positive probability, our limit graphs may not be n-e.c. for a suitable n. This result leads directly to Corollary 2.6, which proves that the limits may not be isomorphic depending on ρ. In Section 3, we study deterministic limit graphs that always embed in limits of graphs generated by our model. Every finite graph H gives rise to a limit R_H satisfying interesting properties. As discussed in Section 3, the study of R_H is connected to the theory of graph homomorphisms.

All the graphs we consider are countable, undirected, and simple. We use the notation $G \le H$ if G is an induced subgraph of H. For an event A in a probability space, we denote the probability of A by $\mathbb{P}(A)$; the negation of A is written $\overline{A}$. If X is a random variable, then $\mathbb{E}(X)$ is its expectation.

2. Main Results: Adjacency and Isomorphism Properties of the Limits

If $(G_t : t \ge t_0)$ is a sequence of graphs with $G_t \le G_{t+1}$, then define the *limit* of the G_t, written $G = \lim_{t\to\infty} G_t$, by

$$V(G) = \bigcup_{t\in\mathbb{N}} V(G_t),\;\; E(G) = \bigcup_{t\in\mathbb{N}} E(G_t).$$

We say that a vertex $x \in V(G)$ is *born* at time t if $x \in V(G_t) \setminus V(G_{t-1})$ if $t > 0$, or $x \in V(G_{t_0})$ if $t = t_0$. A finite set $X \subseteq V(G)$ of vertices of G is *born* at time t if all vertices of X were born at time t or earlier, and some vertex of X is born at time t. If y is a vertex of a graph G, then $N(y) = \{z \in V(G) : yz \in E(G)\}$ is the *neighbor set* of y in G. In the context of $(G : t \ge t_0)$, we will use $N_t(y)$ to denote the neighbor set of y in the graph G_t.

Fix a real number $p \in (0, 1)$. If we generate a countable infinite random graph G as a limit of a random graph process where new vertices are joined to existing ones independently and with fixed probability p, then, with probability 1, the graph G will be isomorphic to the infinite random graph R. The deterministic graph R is the unique isomorphism type of countable graph satisfying the *e.c.* adjacency property, which is the logical conjunction of all the n-e.c. properties.

A new adjacency property introduced in the context of limits of copying models graphs in [Bonato and Janssen 04a] is *locally e.c.* (In [Bonato and Janssen 04a], locally e.c. is referred to as the less-descriptive Property B.) A graph G is *locally e.c.* if for each vertex y of G, for each finite $X \subseteq N(y)$, and each finite $Y \subseteq V(G) \setminus X$, there exists a vertex $z \neq y$ which is c.j. to X and Y. The locally e.c. property is a variant of the e.c. property that applies only to sets contained in the neighbor set of a vertex. Further, it plays a critical role in the model Copy(p, ρ, H). For example, as the next theorem demonstrates, the model Copy(p, ρ, H) almost surely generates limits satisfying the locally e.c. property.

We first define a useful function $p_\rho(i, j, t)$, which is exactly the probability that a new vertex v_t is joined by extra edges to each vertex of an existing set X of cardinality i and no vertices from a set Y of cardinality j. For all non-negative integers i, j, t, define

$$p_\rho(i,j,t) = \frac{\binom{t-1-i-j}{\rho(t)-i}}{\binom{t-1}{\rho(t)}}. \tag{2.1}$$

If $\rho(t) \geq 1$, then by estimating (2.1), we have that

$$\left(1 - \frac{\rho(t)}{t-j}\right)^j \left(\frac{\rho(t)-i+1}{t-i+1}\right)^i \leq p_\rho(i,j,t) \leq \left(\frac{\rho(t)}{t-1}\right)^i. \tag{2.2}$$

If i and j are constants, then (2.2) implies the useful fact that

$$p_\rho(i,j,t) = \Theta(\rho(t)^i t^{-i}).$$

Note also that $p_\rho(0,j,t)$ is increasing in t.

Theorem 2.1. *Fix $p \in (0,1)$, H a finite graph, and let $\rho(t) = \Theta(t^s)$ for $s \in [0,1]$. With probability 1, the limit $G = \lim_{t\to\infty} G_t$ of graphs generated by the model* Copy(p, ρ, H) *is locally e.c.*

Proof. Since a countable union of measure 0 subsets has measure 0, it suffices to show that for a fixed $y \in V(G)$ and finite disjoint $X \subseteq N(y), Y \subseteq V(G)$, the probability that there is no vertex correctly joined to all of X, Y is 0 (since there are only countably many choices for y and X, Y in G). Fix a vertex y and disjoint finite sets $X \subseteq N(y)$ and Y in $V(G)$, and let t_1 be the time that $X \cup Y \cup \{y\}$ is born. Let $Y_1 = Y \cap N_{t_1}(y)$. Let $|X| = k$, $|Y| = j$, and $|Y_1| = j_1$.

We will show that the probability that none of the new nodes $\{v_t : t > t_1\}$ is c.j. to X and Y in the limit equals zero. Let $B_{X,Y}(t)$ be the event that v_t is c.j. to X and Y. Let $B'_{X,Y}(t)$ be the event that all of the following occur:

1. The copy node u_t in time-step t equals y.
2. Every edge from y to vertices in X is copied.
3. None of the edges from y to vertices in Y_1 are copied.
4. No vertex in Y receives an extra edge in time-step t.

Note that $B'_{X,Y}(t)$ contains $B_{X,Y}(t)$. Also, the events $B'_{X,Y}(t)$ for different values of t are independent.

If $t > 0$, then the probability that u_t equals y is $\frac{1}{t-1}$, since u_t is chosen u.a.r. from G_{t-1}. The probability that all edges to X and no edges to Y are copied, given that $u_t = y$, equals $p^k(1-p)^{j_1}$. For all $t > 0$ so that $j < t - \rho(t)$, the probability that no vertex in Y receives an extra edge equals

$$p_\rho(0,j,t) = \Theta(1). \tag{2.3}$$

Hence,

$$\mathbb{P}(B'_{X,Y}(t)) \geq \frac{d}{t},$$

where $d \in (0,1)$ does not depend on t. Then

$$\begin{aligned}\mathbb{P}(\text{no vertex of } G \text{ is c.j. to } X, Y) &\leq \mathbb{P}\left(\bigcap_{t \geq t_2} \overline{B'_{X,Y}(t)}\right) \\ &\leq \prod_{t \geq t_2} \left(1 - \frac{d}{t}\right) = 0,\end{aligned}$$

where the last inequality follows by elementary properties of infinite products. □

The problem of determining whether the limits of graphs generated by copying models converge to R was left open in [Bonato and Janssen 04a]. In the following two theorems we address this question by studying the following adjacency property. For a non-negative integer n, a graph is *strongly n-e.c.* if, for each pair of disjoint, finite sets of vertices A and B with $|A| = n$, there is a vertex z not in $A \cup B$ c.j. to A and B. Hence, no restriction is put on $|B|$ in terms of n. Note that a graph G is *strongly* 0*-e.c.* if and only if for each finite set $B \subseteq V(G)$, there is a vertex not in B that is not joined to any vertex of B. For notational consistency, we say that a graph is *strongly ∞-e.c.* if it is e.c.

We will establish a precise threshold for the values of n for which an infinite limit generated by $\text{Copy}(p, \rho, H)$ has the strongly n-e.c. property. In particular, for an extra-edge function $\rho(t) = \Theta(t^s)$, we define a value $n_{p,s}$ below and show that any limit of $\text{Copy}(p, \rho, H)$ is strongly n-e.c. if $n \leq n_{p,s}$, but, with positive probability, not n-e.c. if $n > n_{p,s}$.

For $p, s \in (0,1)$, define the integer

$$n_{p,s} = \max\left(\left\lfloor \frac{1}{1-s} \right\rfloor, \lfloor \log_p(1-s) \rfloor + 1\right).$$

For all $s \in (0,1)$, we define $n_{0,s} = \lfloor \frac{1}{1-s} \rfloor$, and for all $p \in (0,1)$, $n_{p,1} = \infty$. Note that $n_{p,s}$ is the maximum of two quantities, one of which depends only on s, while the other depends on p and s. If the maximum is attained by $\lfloor \frac{1}{1-s} \rfloor$, the limit is primarily determined by the randomly chosen extra edges (such as when $s > 1 - p$). In the other case (for example when $s < 1/2$), there is a subtle interplay between copy edges and extra edges.

Theorem 2.2. *Let $p \in [0,1)$, $\rho(t) = \Theta(t^s)$ for some $s \in [0,1]$, and H be a finite graph. Let $G = \lim_{t\to\infty} G_t$ be generated by* $\text{Copy}(p, \rho, H)$*. With probability* 1*, G is strongly $(n_{p,s})$-e.c. In particular, if $s = 1$, then G is isomorphic to R.*

In the case $p = 0$, no copying occurs and only $\rho(t)$ extra edges are added to the new vertex in each time-step. This includes the *growing m-out* model (see [Bollobás and Riordan 05]), where $p = 0$ and $\rho(t) = m$. In the case $p = 0$, we have a sharp threshold for the strongly n-e.c. property at $n_{0,s} = \lfloor \frac{1}{1-s} \rfloor$. For values of $p > 0$ and s such that $n_{p,s} = \lfloor \frac{1}{1-s} \rfloor$, the infinite limit behaves similarly to the case $p = 0$. This provides additional evidence that in this case the copy behavior plays a secondary role in the generating process.

On the other hand, we have the following theorem.

Theorem 2.3. *Let $p \in [0,1)$, $\rho(t) = \Theta(t^s)$ for some $s \in [0,1)$, and H be a finite graph. Let $G = \lim_{t\to\infty} G_t$ be generated by* Copy(p, ρ, H).

1. *If $\rho(t) = 0$, then, with positive probability, G is not* 1*-e.c.*
2. *If $s = 0$, where m is a positive integer, then, with positive probability, G is not* 2*-e.c.*
3. *If $s \in (0,1)$, then, with positive probability, G is not $(n_{p,s} + 1)$-e.c.*

To prove these theorems, we need strict bounds on the number of nodes joined to a set of a specified cardinality. Namely, the probability that a set X of size n will obtain its first common neighbor in the next time-step is largely determined by the number of common neighbors of each of its subsets. If any subset $A \subseteq X$ of size $n-1$ has a significant number of common neighbors, then this probability is relatively high: X will receive a common neighbor if any of the common neighbors of A is chosen as the copy vertex, all links to A are copied, and the sole vertex in $X \setminus A$ receives one of the extra edges. On the other hand, if none of the subsets of X have a common neighbor, then a common neighbor of X can only come about if all vertices in X are chosen as the endpoints of extra edges, an event of vanishingly low probability if $\rho(t)$ is sublinear.

Assume that we are given a graph sequence $(G_t : t \in \mathbb{N})$ generated by Copy$(p, \rho(t), H)$ with infinite limit G, where $\rho(t) = \Theta(t^s)$, $s \in (0,1)$, and $p \in [0,1)$. For a fixed finite set X, let $B_X(t)$ be the event that v_t is joined to all of X:

$$\mathbb{P}(B_X(t)) = \sum_{A\subseteq X} \mathbb{P}(B_X(t) \,|\, N_{t-1}(u_t) \cap X = A)\; \mathbb{P}(N_{t-1}(u_t) \cap X = A) \qquad (2.4)$$

Since u_t is chosen u.a.r., $\mathbb{P}(N_{t-1}(u_t) \cap X = A)$ is proportional to the number of vertices in G_{t-1} that are joined to all vertices in A and none in $X \backslash A$. In particular, to obtain a good estimate of $\mathbb{P}(B_X(t))$, we need good bounds on the number of common neighbors of all subsets of X.

We now introduce two important parameters. For a fixed finite vertex set $X \subseteq V(G)$, let

$$\delta(X, t) \tag{2.5}$$

be the number of vertices in G_t that are joined to all vertices in X. Lemmas 2.4 and 2.5 below will show that, with high probability as t increases, $\delta(X, t)$ is of order $t^{a_k \pm \varepsilon}$ for any $\varepsilon > 0$, where a_k is a constant that depends only on the cardinality k of X. In Lemma 2.4, we obtain the lower-order bound for the number of vertices correctly joined to X and Y, for any finite set Y disjoint from X. We denote this number by

$$\delta(X, Y, t). \tag{2.6}$$

The exponents a_k are defined as follows. For $p, s \in (0, 1)$, and for a non-negative integer k, define

$$a_k = \max\{p^k, 1 - (1 - s)k\}.$$

Note that the a_k are decreasing in k. The constant $n_{p,s}$ from Theorem 2.3 is defined so that it is the least positive integer k with the property that $a_k < 1 - s$. It will become apparent in the proof of Theorem 2.3 why this condition is needed.

Consider the real-valued function $f(x) = p^x - (1 - s)(k - x)$. This function has one extreme value, which is a minimum. Hence, f achieves its maximum over the interval $[0, k]$ at $x = 0$ or $x = k$. It follows that

$$a_k = \max\{p^x - (1 - s)(k - x) \,:\, 0 \le x \le k\}.$$

As a consequence, for any two integers ℓ and k so that $0 \le \ell < k \le n_{p,s}$,

$$\begin{aligned} a_k &\ge \max\{p^x - (1 - s)(k - x) \,:\, 0 \le x \le \ell\} \\ &= \max\{p^x - (1 - s)(\ell - x) : 0 \le x \le \ell\} - (1 - s)(k - \ell) \\ &= a_\ell - (1 - s)(k - \ell). \end{aligned} \tag{2.7}$$

Observe that equality holds in the first displayed line precisely when the maximum of $f(x)$ over the interval $[0, k]$ is achieved at $x = 0$; that is, if $p^k \le 1 - (1 - s)k$.

Let $(A_t : t \in \mathbb{N})$ be a sequence of events. We say that A_t holds *with extreme probability* (*wep*) in t if, for all $t \in \mathbb{N}$, $\mathbb{P}(\overline{A_t}) \le h(t)$, where $h(t)$ is a function that exponentially decreases to zero as t goes to infinity. The proof of the following technical lemma is postponed until Section 4.

Lemma 2.4. *Let $(G_t : t \geq t_0)$ be a sequence of graphs generated by* $\text{Copy}(p, \rho, H)$ *with copy probability $p \in (0, 1)$, extra-edge function $\rho(t) = \Theta(t^s)$ for some $s \in [0, 1]$, and H a finite graph. Let X and Y be disjoint vertex sets born at time $t_1 > t_0$, of cardinality $|X| = k \leq n_{p,s}$ and $|Y| = \ell$. Then for each $\varepsilon > 0$ and for each $t_2 \geq t_1$, there exists a constant $c > 0$ so that wep in t_2, for all $t \geq t_2$*

$$\delta(X, Y, t) \geq ct^{a_k - \varepsilon}. \tag{2.8}$$

Theorem 2.2 follows directly from this lemma.

Proof of Theorem 2.2. Let X and Y be a pair of disjoint, finite vertex sets in G, born at time t_1, and such that $|X| \leq n_{p,s}$. For $t > t_1$, let $A_{X,Y}(t)$ be the event that for some $\tau \geq t$,

$$\delta(X, Y, \tau) < c\tau^{a_k - \varepsilon}.$$

By Lemma 2.4, $\mathbb{P}(A_{X,Y}(t)) = h(t)$, where $h(t)$ is an exponentially decreasing function. Then, for all $t > t_1$,

$$\begin{aligned} \mathbb{P}(G \text{ contains no vertex c.j. to } X \text{ and } Y) &= \mathbb{P}(\text{For all } \tau \geq t,\ \delta(X, Y, \tau) = 0) \\ &\leq \mathbb{P}(A_{X,Y}(t)) \\ &= h(t). \end{aligned}$$

Since the above holds for all t, we have that $h(t) = o(1)$. As in the proof of Theorem 2.1, the proof now follows since there are only countably many choices for X and Y. □

For the proof of Theorem 2.3, we need the following technical lemma whose proof is also given in Section 4.

Lemma 2.5. *Let $(G_t : t \geq t_0)$ be a sequence of graphs generated by* $\text{Copy}(p, \rho, H)$ *with copy probability $p \in (0, 1)$, extra-edge function $\rho(t) = \Theta(t^s)$ for some $s \in [0, 1)$, and H a finite graph. Let X be a set born at time $t_1 > t_0$, of cardinality $|X| = k \leq n_{p,s}$.*

1. *Then, for each $\varepsilon > 0$ and for each $t_2 \geq t_1$, there exists a constant $c > 0$ such that wep in t_2 for all $t \geq t_2$*

$$\delta(X, t) \leq ct^{a_k + \varepsilon}. \tag{2.9}$$

2. *Let $B_X(t)$ denote the event that the new vertex v_t at time t joins to all vertices of X. Then, for any set X of size $|X| = k = n_{p,s} + 1$ born at time*

t_1, there exists $t_2 \geq t_1$, such that

$$\mathbb{P}\left(\bigcap_{t \geq t_2+1} \overline{B_X(t)} | \delta(X, t_2) = 0\right) > 0. \tag{2.10}$$

We now prove our second main result.

Proof of Theorem 2.3. We show that there exists a set X that, with positive probability, has no common neighbor in G. In the following, we always take X to be a set of cardinality $k = n_{p,s} + 1$, born at time t_1, so that the probability that X contains no common neighbor in G_{t_1}, written p_X, is positive. Note that a set X for which $p_X > 0$ always exists: Let t_1 be such that $\rho(t_1 - 1) + \rho(t_1) < t_1$, and choose X so that it contains v_{t_1-1} and v_{t_1}. Then, with positive probability, in time-steps $t_1 - 1$ and t_1 no links are copied, and no neighbors of v_{t_1-1} are chosen as endpoints of the extra edges in step t_1. Therefore, v_{t_1-1} and v_{t_1} have no common neighbor, so neither has X.

For $t \geq t_1$, $B_X(t)$ is the event that v_t is a common neighbor of X. We now consider the case when $p \in (0, 1)$ (the proof in the case $p = 0$ is analogous and so is omitted). Suppose first that $\rho(t) = 0$. If a vertex v_t does not receive any edges when it is born, then it is isolated in G. This occurs when no copying occurs in time-step t, which happens with positive probability since $p < 1$. This completes the proof of item (1).

We now turn to the proof of items (2) and (3). Assume that $s \in [0, 1)$, and

$$k = n_{p,s} + 1.$$

Let $\overline{s} = 1 - s$. By (2.2), for all positive integers ℓ , the function $p_\rho(\ell, 0, t)$ is a decreasing function of order $t^{-\overline{s}\ell}$. Let $c > 0$ be such that, for all $t > 0$ and all integers $1 \leq \ell \leq k$,

$$p_\rho(\ell, 0, t) \leq ct^{-\overline{s}\ell}. \tag{2.11}$$

Note that c does not depend on t_1. Assume X is chosen so that it is born a time t_1 that is great enough so that

$$c(t_1^{-\overline{s}} + 1)^k t_1^{-\overline{s}} < 1. \tag{2.12}$$

By Lemma 2.5, there is a $t_2 > t_1$ such that

$$\mathbb{P}\left(\bigcap_{t \geq t_2+1} \overline{B_X(t)} | \delta(X, t_2) = 0\right) > 0. \tag{2.13}$$

For $t \le t_2$, we estimate $\mathbb{P}(B_X(t+1)|\delta(X,t)=0)$ by observing that $\frac{\delta(A,t)}{t} \le 1$ for all $A \subset X$, and thus, for all $t \ge t_1$,

$$\begin{aligned}\mathbb{P}(B_X(t+1)|\delta(X,t)=0) &\le c\sum_{A\subset X} t^{-\overline{s}(|X-A|)}\\ &= ct^{-\overline{s}}\sum_{\ell=0}^{k}\binom{k}{\ell}t^{-\overline{s}(k-\ell)}\\ &\le c(t^{-\overline{s}}+1)^k t^{-\overline{s}},\end{aligned}$$

where the first inequality follows from (2.11), and the second inequality follows from the binomial theorem.

Therefore,

$$\begin{aligned}\mathbb{P}(\delta(X,t_2)=0) &= p_X\prod_{t=t_1}^{t_2-1}\mathbb{P}(\overline{B_X(t+1)}|\,\delta(X,t)=0)\\ &\ge p_X\prod_{t=t_1}^{t_2-1}(1-c(t^{-\overline{s}}+1)^k t^{-\overline{s}}) \qquad (2.14)\\ &> 0,\end{aligned}$$

where the last inequality follows from (2.12) and because $p_X > 0$.

Then

$$\begin{aligned}&\mathbb{P}(X \text{ has no common neighbor in } G)\\ &\quad = \mathbb{P}(\delta(X,t_2)=0)\mathbb{P}\left(\bigcap_{t\ge t_2+1}\overline{B_X(t)}|\delta(X,t_2)=0\right)\\ &\quad > 0,\end{aligned}$$

where the inequality follows from (2.13) and (2.14). Therefore, with positive probability X has no common neighbor in G, which concludes the proof. □

The strength of the n-e.c. properties satisfied by a limit can be used to distinguish its isomorphism type when the parameter p varies. As a corollary of Theorem 2.3, we obtain the following non-isomorphism result for limits generated by our model.

Corollary 2.6. *Let H be a finite graph, and let $p, p', s, s' \in (0,1)$ such that $\rho(t) = \Theta(t^s)$ and $\rho'(t) = \Theta(t^{s'})$. If $n_{p,s} < n_{p',s'}$, then, with positive probability, a limit generated by* $\mathrm{Copy}(p,\rho,H)$ *is not isomorphic to a limit generated by* $\mathrm{Copy}(p',\rho',H)$.

We do not know whether the hypothesis of Corollary 2.6 may be relaxed to simply $s < s'$.

3. Minimal Graphs for the Model

By Corollary 2.6, limits generated by the model Copy(p, ρ, H) for different choices of ρ and p need not be isomorphic. Contrary to the situation for countable e.c. graphs (which are isomorphic to the infinite random graph R), not all countable graphs that are either locally e.c. or strongly n-e.c. for some n need to be isomorphic. In fact, there are $2^{\aleph_0}$ (that is, cardinality of $\mathbb{R}$) many non-isomorphic countable locally e.c. graphs; see [Bonato and Janssen 04a].

We may define certain graphs that are *minimal* graphs with the given adjacency property, in the sense that they embed in any graph with the property. We devote this section to the study of deterministic graphs that are minimal for our models. We introduce graphs R_H and $R_H^{(n)}$ that are minimal for both the locally e.c. and strongly n-e.c. properties, respectively. Aside from their relationship to our model, these graphs may be of interest in their own right.

We introduce the following infinite graphs.

R_H: Fix a finite graph H. Let $R_{H,0} \cong H$. For a fixed $t \geq 0$, assume that $R_{H,t}$ is defined and is finite. For each vertex $y \in V(R_{H,t})$, and each subset $X \subseteq N(y)$, add a new vertex $z_{y,X}$ joined only to X. This gives a graph $R_{H,t+1}$ that contains $R_{H,t}$ as an induced subgraph. Define $R_H = \lim_{t\to\infty} R_{H,t}$.

$R_H^{(n)}$: Fix a finite graph H. Let $R_{H,0}^{(n)} \cong H$. For a fixed $t \geq 0$, assume that $R_{H,t}^{(n)}$ is defined and is finite. For each vertex $y \in V(R_{H,t}^{(n)})$, and each subset $X \subseteq N(y)$, add a new vertex $z_{y,X}$ joined only to X. For each subset of vertices of Y with cardinality at most n add a new vertex z_Y joined only to Y. This gives a graph $R_{H,t+1}^{(n)}$. Define $R_H^{(n)} = \lim_{t\to\infty} R_{H,t}^{(n)}$.

Observe that R_H is locally e.c., while $R_H^{(n)}$ is both locally e.c. and strongly n-e.c. In addition, for all n, the graph R_H is an induced subgraph of $R_H^{(n)}$. The graphs R_H and $R_H^{(n)}$ play the role of minimal graphs for certain adjacency properties.

Theorem 3.1. *Fix $n \in \mathbb{N}$ and a finite graph H.*

1. *If G is a locally e.c. graph, then $R_H \leq G$ if and only if $H \leq G$.*
2. *If G is a locally e.c., strongly n-e.c graph, then $R_H^{(n)} \leq G$ if and only if $H \leq G$.*
3. *The graphs R_H and $R_H^{(n)}$ are not isomorphic.*

Proof. As the proofs of (1) and (2) are similar, we therefore prove only items (2) and (3). The forward direction of (2) is immediate, since $H \leq R_H^{(n)}$ by the definition of the limit. For the reverse direction of (2), suppose that H is an induced subgraph of G. Let $R_t = R_{H,t}^{(n)}$, where $t \in \mathbb{N}$, be the finite graphs used to define $R_H^{(n)}$. We proceed by induction on t to show that each of the graphs R_t is an induced subgraph of G extending the embedding of H in G. We take R_0 to be this copy of H.

Fix a vertex $y \in V(R_t)$ and a subset $X \subseteq N(y)$. Let $Y = V(R_t)\backslash X$. As G is locally e.c., there is a vertex z correctly joined to X, Y. The vertex z is joined only to X in R_t and plays the role of the vertex $z_{y,X}$ in the definition of $R_H^{(n)}$. Fix a subset Y of vertices of cardinality at most n. Let $Z = V(R_t)\backslash Y$. As G is strongly n-e.c. there is a vertex z correctly joined to Y, Z. Vertex z is joined only to Y in R_t, so plays the role of z_Y in the definition of $R_H^{(n)}$. Therefore,

$$R_H^{(n)} \cong \lim_{t \to \infty} R_t \leq G.$$

For item (3), note that the graph R_H contains isolated vertices. To see this, fix $y \in V(R_0)$ and $X = \emptyset$. Then $z_{y,X}$ is isolated in R_1. By the definition of R_H, the vertex $z_{y,X}$ acquires no new neighbors as t tends to infinity. However, as $n \geq 1$, no vertex of $R_H^{(n)}$ is isolated. To see this, fix $v \in V(R_H^{(n)})$. Then $v \in V(R_t)$ for some $t \geq 0$. The vertex $z_{\{v\}}$ is joined to v in R_{t+1}. □

Corollary 3.2. *Fix $p \in (0,1)$ and H.*

1. *If $\rho(t) = \Theta(t^s)$, with $s \in [0,1]$, then, with probability 1, a limit graph G generated by the model $\mathrm{Copy}(p, \rho, H)$ embeds $R_H^{(n)}$, where $n = n_{p,s}$.*
2. *If $\rho(t) = 0$, then, with probability 1, a limit graph G generated by the model $\mathrm{Copy}(p, \rho, H)$ embeds R_H.*

Proof. By Theorem 2.1 and Theorem 2.2 (2), with probability 1, G is locally e.c. and, for case (1), strongly $n_{p,s}$-e.c. Now apply Theorem 3.1. □

If G is a countable graph, then define the *clique number* of G, written $\omega(G)$, to be the supremum of the set $\{|K| : K \text{ is a clique in } G\}$. We omit the straightforward proof of the following theorem.

Theorem 3.3. *Fix $n \in \mathbb{N}$ and H a finite graph. Then the following equalities hold.*

1. $\omega(R_H) = \omega(H)$.
2. $\omega(R_H^{(n)}) = \max\{n + 1, \omega(H)\}$.

As an immediate corollary of Theorem 3.3, we have the following.

Corollary 3.4.

1. *There are infinitely many non-isomorphic graphs of the form* R_H.
2. *There are infinitely many non-isomorphic graphs of the form* $R_H^{(n)}$.

A mapping $f : V(G) \to V(H)$ is a *homomorphism* if it has the property that if $xy \in E(G)$, then $f(x)f(y) \in E(H)$. The map f is sometimes called an *H-coloring*, and G is referred to as *H-colorable*. We write $G \to H$ to denote that G admits a homomorphism to H without reference to a specific mapping. See [Hell and Nešetřil 04] for more on graph homomorphisms. The following theorem establishes a surprising connection between graph homomorphisms and the induced subgraphs of R_H.

Theorem 3.5. *Fix n a positive integer, with H a finite graph and G a countable graph. Then $G \leq R_H$ if and only if $G \to H$.*

Proof. For the forward direction, suppose that $G \leq R_H$. Then $G \to R_H$. It is not hard to see that each R_t in the definition of R_H admits a homomorphism f_t to H: Each new vertex is assigned the same image as the node it copies from. The union f of the chain $(f_t : t \in \mathbb{N})$ of homomorphisms is a homomorphism from R_H to H. As $G \leq R_H$, we have that $G \to H$ by the transitivity of the homomorphism relation.

For the converse, assume first that G is finite. We introduce an auxiliary graph construction. Fix a homomorphism $f : G \to H$. Assume that $V(G)$ and $V(H)$ are disjoint, and define a graph $H(G, f)$ to have vertices $V(G) \cup V(H)$ and edges

$$E(G) \cup E(H) \cup \{xy : x \in V(G), y \in V(H), \text{and } f(x)y \in E(H)\}.$$

We refer to the induced copy of H in $H(G, f)$ as H'. The graph $H(G, f)$ is the union of G and H, so that, for each vertex x of G, x is joined to all the neighbors of $f(x)$ in H. We proceed by induction on $|V(G)|$ to show that $H(G, f) \leq R_H$. (Since $G \leq H(G, f)$, this proves the statement.) Note that, if $|V(G)| = 1$, then $H(G, f)$ is isomorphic to the disjoint union of H and K_1. The base case follows. The induction hypothesis is that if $|V(G)| = n$, where $n \geq 1$ is fixed, then $H(G, f)$ is an induced subgraph of R_H with H' the copy of H at $t = 0$.

Let $|V(G)| = n + 1$, and fix $x \in V(G)$. By the induction hypothesis, $H(G - x, f \restriction (G - x))$ is a subgraph of R_H with H' the copy of H at $t = 0$. By the definition of $H(G, f)$, all the neighbors of x in $H(G, f)$ are also neighbors of the vertex $f(x)$ in H'. By the locally e.c. property of R_H, there is a vertex z of R_H joined exactly to the neighbors of x in $H(G, f)$. Adding z to the copy of $H(G - x, f \restriction G - x)$ in R_H will give an induced subgraph of R_H that is isomorphic to $H(G, f)$, while H' is unchanged. This completes the induction.

In the case when G is infinite, express G as a limit of some chain of finite induced subgraphs $(G_t : t \in \mathbb{N})$. An easy, therefore omitted, argument using the technique from the finite case shows that we may embed the graphs G_t as a $G'_t \leq R_H$ so that each G'_{t+1} contains G'_t. Hence,

$$G \cong \lim_{t\to\infty} G'_t \leq R_H. \qquad \square$$

One interpretation of Theorem 3.5 is that the graph R_H carries a certain memory of H, made explicit by the homomorphism to H. Hence, we have the following corollary.

Corollary 3.6. *For a fixed finite graph H, all countable H-colorable graphs embed in R_H; that is, R_H is a universal H-colorable graph.*

Corollary 3.6 along with Theorem 3.1 expresses an interesting duality property for R_H: R_H is at once the minimal (with respect to the embedding relation) locally e.c. graph containing H and the maximal H-colorable graph. This form of duality is not unique to R_H, and emerges in other limit graphs arising from network models (see [Bonato et al. 09]).

4. Proofs of Lemmas 2.4 and 2.5

In this final section, we give proofs of Lemmas 2.4 and 2.5. Our main tools are the following versions of the Chernoff bounds, which we state here for completeness (see also Section 2.1 of [Janson et al. 00]).

Theorem 4.1. *Let $Z_t = \sum_{i=1}^{t} \zeta_i$ be the sum of random variables ζ_i, where $\zeta_i \in Be(p_i)$ for $1 \leq i \leq t$; that is, each ζ_i is a Bernoulli random variable where $\mathbb{P}(\zeta_i = 1) = p_i$. (Note that $\mathbb{E}(Z_t) = \sum_{i=1}^{t} p_i$.)*

Then for all $\gamma \geq 0$, we have the following inequalities:

$$\mathbb{P}(Z_t \leq \mathbb{E}(Z_t) - \gamma) \leq exp\left(-\frac{\gamma^2}{2\mathbb{E}(Z_t)}\right), \tag{4.1}$$

$$\mathbb{P}(Z_t \geq \mathbb{E}(Z_t) + \gamma) \leq exp\left(-\frac{\gamma^2}{2(\mathbb{E}(Z_t) + \gamma/3)}\right). \tag{4.2}$$

Proof of Lemma 2.4. Let X, Y of size k and ℓ, respectively, born at t_1, be as in the statement of the lemma. Let $B_{X,Y}(t)$ denote the event that the new vertex v_t at time t joins to all vertices of X and no vertex in Y.

Case 1. $p^k < 1 - \bar{s}k$, so $a_k = 1 - \bar{s}k$.

In this case, the extra edges play the crucial role; a lower bound on $\mathbb{P}(B_{X,Y}(t))$ can be derived from the extra edges. Namely, $B_{X,Y}(t)$ occurs if the extra edges from v_t are joined to all vertices in X and none in Y, and any edge of the copy vertex to a vertex in Y is not copied. Therefore,

$$\mathbb{P}(B_{X,Y}(t)) \geq p_\rho(k, \ell, t)(1-p)^\ell.$$

Recall from (2.6) the definition of $\delta(X, Y, t)$. Now, for $t \geq t_1$, we have that $\delta(X, Y, t) - \delta(X, Y, t_1)$ is bounded below by the sum of independent Bernoulli variables $Z_t = \sum_{i=t_1+1}^{t} \zeta_i$, with $\zeta_i \in Be(p_i)$, where $p_i = p_\rho(k, \ell, i)(1-p)^\ell$. So

$$\mathbb{E}(Z_t) = (1-p)^\ell \sum_{i=t_1+1}^{t} p_\rho(k, \ell, i).$$

By (2.2), $p_\rho(k, \ell, t)$ is an increasing function of order $\Theta(t^{-\bar{s}k})$ so $\mathbb{E}(Z_t) \geq ct^{1-\bar{s}k} = ct^{a_k}$ for some constant $c > 0$. By the Chernoff bounds (4.1) for each $\varepsilon > 0$ and $t > t_1$,

$$\begin{aligned} \mathbb{P}(\delta(X, Y, t) \leq t^{a_k - \varepsilon}) &\leq \mathbb{P}(Z_t \leq t^{a_k - \varepsilon}) \\ &\leq exp(-(c/2)t^{a_k}(1 + O(t^{-\varepsilon}))). \end{aligned}$$

Fix $t_2 > t_1$. Then

$$\mathbb{P}(\text{For some } t \geq t_2,\ \delta(X, Y, t) \leq t^{a_k - \varepsilon}) \leq \sum_{t=t_2}^{\infty} exp(-(c/2)t^{a_k}(1 + O(t^{-\varepsilon}))).$$

The sum decreases exponentially in t_2, so wep in t_2, $\delta(X, Y, t) \geq t^{a_k - \varepsilon}$ for all $t \geq t_2$. Hence, (2.8) follows for Case 1.

Case 2. $p^k > 1 - \bar{s}k$.

In this case, copying plays the central role. We use induction on k. Since $p^0 = 1 - 0\bar{s}$, the base case $k = 0$ follows by Case 1.

For the induction step, fix $k \leq n_{p,s}$, and let X and Y be as stated in the lemma. Let A be a fixed subset of X of size $k-1$, and let w be the unique vertex in $X - A$. For any $t \geq t_1$, the new vertex v_{t+1} is correctly joined to X

and Y if the copy vertex u_{t+1} is correctly joined to A and Y (with probability $\delta(A,Y,t)/t$), and all edges from u_{t+1} to vertices in A are copied (with probability p^{k-1}), and the vertex w receives an extra edge, but no vertex in Y does (with probability $p_\rho(1,\ell,t+1)$). Therefore,

$$\mathbb{P}(B_{X,Y}(t+1)) \geq \frac{\delta(A,Y,t)}{t} p^{k-1} p_\rho(1,\ell,t+1). \tag{4.3}$$

Fix $\varepsilon > 0$ such that $\varepsilon < a_{k-1} - \bar{s}$. Note that $a_{k-1} > \bar{s}$ by the fact that $k \leq n_{p,s}$ and the definition of $n_{p,s}$.

By induction, there exists an exponentially decreasing function $h(t)$ such that for every $t_2 \geq t_1$, there exists a constant c such that

$$\mathbb{P}(\text{For all } t \geq t_2,\ \delta(A,Y,t) \geq c_A t^{a_{k-1}-\epsilon}) \geq 1 - h(t_2).$$

Fix $t_2 > t_1$. Since the statement of the lemma is asymptotic for t_2 going to infinity, we can assume without loss of generality that $t_2 > 2t_1$. Let c_A be such that, with probability $1 - h(t_2/2)$ for all $t \geq t_2/2$,

$$\delta(A,Y,t) \geq c_A t^{a_{k-1}-\varepsilon}. \tag{4.4}$$

Define

$$T = \min\{t > t_2/2 : \delta(A,Y,t) < c_A t^{a_{k-1}-\varepsilon} \text{ or } t = t_2\}.$$

Fix t so that $t_2/2 \leq t < T$. Since $p_\rho(1,\ell,t+1)$ is non-increasing and of order $\Theta(t^{-\bar{s}})$, by (2.2), (4.3), and (4.4) we obtain that

$$\mathbb{P}(B_{X,Y}(t+1)) \geq c t^{a_{k-1}-\varepsilon-1-\bar{s}}$$

for some constant c. Thus, $\delta(X,Y,t) - \delta(X,Y,t_2/2)$ can be bounded below by Z_t, a sum of independent Bernoulli variables, with

$$\begin{aligned}\mathbb{E}(Z_t) &= c \sum_{\tau=t_2/2+1}^{t} \tau^{a_{k-1}-\varepsilon-1-\bar{s}} \\ &= \frac{c}{a_{k-1}-\bar{s}-\varepsilon}(t^{a_{k-1}-\bar{s}-\varepsilon} - (t_2/2)^{a_{k-1}-\bar{s}-\varepsilon}) + o(1).\end{aligned}$$

By (4.1) we derive that

$$\begin{aligned}\mathbb{P}(\delta(X,Y,T) = 0) &\leq \mathbb{P}(Z_T = 0) \\ &\leq \exp(-(1/2)\mathbb{E}(Z_T)).\end{aligned}$$

From the induction hypothesis, we have that, with probability $1-h(t_2/2)$, $T = t_2$. Since $\mathbb{E}(Z_t)$ is increasing in t, we have that the probability that $\delta(X,Y,t_2) = 0$

is bounded above by an exponentially decreasing function in t_2. In other words, wep in t_2, G_{t_2} contains *at least* one vertex correctly joined to X and Y.

In the following, assume that $\delta(X,Y,t_2) > 0$. For any $t \geq t_2$, v_{t+1} is correctly joined to X and Y if the copy vertex u_{t+1} is correctly joined to X and Y, each link from u_{t+1} to X is copied, and no vertex in Y receives an extra edge in step $t+1$. Therefore,

$$\mathbb{P}(B_{X,Y}(t+1)) \geq p^k \frac{\delta(X,Y,t)}{t} p_\rho(0,\ell,t+1).$$

By (2.2), $p_\rho(0,\ell,t)$ converges to 1 as $t \to \infty$. Without loss of generality, assume that t_2 is great enough so that $p_\rho(0,\ell,t) > 1-\varepsilon/(2p^k)$ for all $t \geq t_2$. Let t_3 be such that

$$(t_2/t_3)^{a_k-\varepsilon} < \frac{\varepsilon/2}{a_k-\varepsilon/2}. \tag{4.5}$$

Choose

$$c^* = (t_3)^{-a_k+\varepsilon}.$$

Define

$$S = \{t \geq t_2 : \delta(X,Y,t) > c^* t^{a_k-\varepsilon}\},$$

and

$$T_1 = \begin{cases} \min S & \text{if } S \neq \varnothing, \\ \infty & \text{else.} \end{cases}$$

The proof in Case 2 will follow if we show that, wep in t_2, $T_1 = \infty$. Note that $T_1 \geq t_3$, since $c^* t^{a_k-\varepsilon} < 1$ for all $t < t_3$, and by assumption $\delta(X,Y,t_2) > 0$.

For all t such that $t_2 \leq t < T_1$,

$$\mathbb{P}(B_{X,Y}(t+1)) \geq (p^k - \varepsilon/2) c^* t^{a_k-\varepsilon-1}.$$

(Recall that, for this case, $a_k = p^k$.) For all t such that $t_2 < t \leq T_1$, $\delta(X,Y,t) - \delta(X,Y,t_2)$ can be bounded below by the sum Z_t of independent Bernoulli variables, where

$$\mathbb{E}(Z_t) \geq c^* \left(\frac{a_k-\varepsilon/2}{a_k-\varepsilon}\right)(t^{a_k-\varepsilon} - (t_2)^{a_k-\varepsilon} + 1).$$

By (4.1) we have that

$$\begin{aligned} \mathbb{P}(\delta(X,Y,t) > c^* t^{a_k-\varepsilon}) &\leq \mathbb{P}(Z_t > c^* t^{a_k-\varepsilon}) \\ &\leq \exp\left(-\left(\frac{(\mathbb{E}(Z_t) - c^* t^{a_k-\varepsilon})^2}{2\mathbb{E}(Z_t)}\right)\right). \end{aligned} \tag{4.6}$$

If $t \geq t_3$, then

$$\begin{aligned}\mathbb{E}(Z_t) - c^* t^{a_k-\varepsilon} &\geq c^* \left(\frac{\varepsilon/2}{a_k - \varepsilon} t^{a_k-\varepsilon} - \frac{a_k - \varepsilon/2}{a_k - \varepsilon} (t_2)^{a_k-\varepsilon} - 1 \right) \\ &= c^* \left(\left(\frac{\varepsilon/2}{a_k - \varepsilon} - \frac{a_k - \varepsilon/2}{a_k - \varepsilon} (t_2/t)^{a_k-\varepsilon} \right) t^{a_k-\varepsilon} - 1 \right)\end{aligned}$$

By (4.5) the term in front of $t^{a_k-\varepsilon}$ is positive. Since $\mathbb{E}(Z_t) = \Theta(t^{a_k-\varepsilon})$, by (4.6) for all $t \geq t_3$

$$\mathbb{P}(\delta(X,Y,t) < c^* t^{a_k-\varepsilon} | \delta(X,Y,t_2) > 0) \leq \exp(-ct^{a_k-\varepsilon}),$$

for some constant $c > 0$. Thus,

$$\mathbb{P}(\text{For some } t_2 < t \leq T_1, \delta(X,Y,t) > c^* t^{a_k-\varepsilon} | \delta(X,Y,t_2) > 0) \leq \sum_{\tau=t_2}^{t} \exp(-c\tau^{a_k-\varepsilon}).$$

From this inequality, it follows that if $\delta(X,Y,t_2) > 0$ then the probability that T_1 is finite is exponentially small (in t_2).

Therefore,

$$\begin{aligned}&\mathbb{P}(\text{For some } t > t_2, \delta(X,Y,t) > c^* t^{a_k-\varepsilon}) \\ &\qquad \leq \mathbb{P}(\text{For some } T > t_2, \delta(X,Y,t) > c^* t^{a_k-\varepsilon} | \delta(X,Y,t_2) > 0) \\ &\qquad + \mathbb{P}(\delta(X,Y,t_2) = 0).\end{aligned}$$

The last two probabilities are both bounded by functions exponentially decreasing in t_2, so the result follows. □

Proof of Lemma 2.5. For item (1), the proof will proceed by induction on $k = |X|$, for $0 \leq k \leq n_{p,s}$. Recall the definition of $\delta(X,t)$ in (2.5). For the base case of the induction, note that if $k = 0$, $X = \emptyset$, and then $\delta(X,t) = |V(G_t)| = t$ and $a_0 = 1$. Therefore, $\delta(X,t) \leq t^{a_0}$ for all $t \geq t_1$, so (2.9) holds with probability 1 for $c = 1$ and $t_2 \geq t_1$.

For the induction step, fix a positive integer $k \leq n_{p,s}$, and let X be a set of vertices born at time t_1 of cardinality k. As explained before and by (2.4), we will resolve the probability $\mathbb{P}(B_X(t+1))$ into cases, depending on the overlap of the neighborhood of the copy vertex $u = u_{t+1}$ with X.

Let A be a fixed subset of X of cardinality ℓ, and define $\overline{s} = 1 - s$. Let $f_\ell(t)$ be a function so that $f_\ell(t) = \Theta(1)$, and, for $0 \leq i \leq \ell$,

$$p_\rho(k-i, 0, t+1) \leq f_\ell(t) t^{-(k-i)\overline{s}}.$$

Such a function f_ℓ exists since, by (2.2), $p_\rho(j, 0, t) = \Theta(t^{-j\overline{s}})$ for all $j > 0$.

Then

$$\begin{aligned}\mathbb{P}(B_X(t+1)\,|\,N_t(u)\cap X=A) &= \sum_{i=0}^{\ell}\binom{\ell}{i}p^i(1-p)^{\ell-i}p_\rho(k-i,0,t+1)\\ &\le t^{-\overline{s}(k-\ell)}f_\ell(t)\sum_{i=0}^{\ell}\binom{\ell}{i}p^i\left((1-p)t^{-\overline{s}}\right)^{\ell-i}\\ &= t^{-\overline{s}(k-\ell)}f_\ell(t)(p+(1-p)t^{-\overline{s}})^\ell.\end{aligned}$$

Moreover, if $\ell=k$, then

$$\begin{aligned}&\mathbb{P}(B_X(t+1)\,|\,N_t(u)\cap X=A)\\ &\quad= p^k+\sum_{i=0}^{k-1}\binom{k}{i}p^i(1-p)^{k-i}p_\rho(k-i,0,t+1)\\ &\quad\le p^k+f_k(t)t^{-\overline{s}}k(1-p)\sum_{i=0}^{k-1}\binom{k-1}{i}p^i(1-p)^{k-1-i}t^{-(k-1-i)\overline{s}}\\ &\quad= p^k+g(t)t^{-\overline{s}},\end{aligned}$$

where $g(t)=f_k(t)k(1-p)(p+(1-p)t^{-\overline{s}})^{k-1}$. Note that $g(t)=\Theta(1)$.

Define the function

$$f(t)=\max_{0\le\ell<k}f_\ell(t)\left(p+(1-p)t^{-\overline{s}}\right)^\ell.$$

Note that $f(t)=\Theta(1)$.

Since

$$\mathbb{P}(B_X(t+1)\,|\,N_t(u)\cap X=A)\le\begin{cases}f(t)t^{-\overline{s}(k-\ell)} & \text{if } A\subset X,\\ p^k+g(t)t^{-\overline{s}} & \text{if } A=X,\end{cases}$$

and

$$\mathbb{P}(N_t(u)\cap X=A)\le\mathbb{P}(A\subseteq N_t(u))=\frac{\delta(A,t)}{t},$$

we have that

$$\begin{aligned}\mathbb{P}(B_X(t+1)) &= \sum_{A\subseteq X}\mathbb{P}(B_X(t+1)\,|\,N_t(u)\cap X=A)\,\mathbb{P}(N_t(u)\cap X=A)\\ &\le (p^k+g(t)t^{-\overline{s}})\frac{\delta\,(X,t)}{t}+f(t)\sum_{A\subset X}\frac{\delta(A,t)}{t}t^{-\overline{s}|X-A|}.\end{aligned}\tag{4.7}$$

Let

$$\delta^*(t)=\max_{A\subset X}t^{-a_{|A|}}\delta(A,t).$$

By (4.7), we have that

$$\begin{aligned}\mathbb{P}(B_X(t+1)) &\leq (p^k + g(t)t^{-\overline{s}})\frac{\delta\,(X,t)}{t} + f(t)\delta^*(t)\sum_{\ell=0}^{k-1}\binom{k}{\ell}t^{-\overline{s}(k-\ell)}t^{a_\ell-1} \\ &\leq (p^k + g(t)t^{-\overline{s}})\frac{\delta\,(X,t)}{t} + f(t)\delta^*(t)t^{a_k-1}\sum_{\ell=0}^{k-1}\binom{k}{\ell} \\ &\leq \left(\frac{p^k}{t} + g(t)t^{-\overline{s}-1}\right)\delta(X,t) + 2^k f(t)\delta^*(t)t^{a_k-1}.\end{aligned}$$

The second inequality follows from (2.7):

$$a_k \geq a_\ell - \overline{s}(k-\ell)$$

for all $0 \leq \ell < k$.

It can be deduced from the induction hypothesis that there exists an exponentially decreasing function $h(t)$ such that for each $t_2 \geq t_1$, there exists a constant c' so that, for each $A \subset X$ of size ℓ with probability at least $1 - h(t_2)$, for all $t \geq t_2$

$$\delta(A,t) \leq c' t^{a_\ell+\varepsilon}.$$

Fix $\varepsilon > 0$. As $g(t) = \Theta(1)$ assume that $g(t)t^{-\overline{s}} \leq \varepsilon/4$ for all $t \geq t_2$, where $t_2 > t_1$ is sufficiently large. As $g(t) = \Theta(1)$, let c_f be so that $f(t) \geq c_f$ for all $t \geq t_2$. Choose $t_3 > t_2$ so that

$$t_2(t_3)^{-a_k-\varepsilon} \leq \frac{\varepsilon/2}{a_k+\varepsilon}. \tag{4.8}$$

Without loss of generality, assume that $c' \geq t_3$. Let $c^* > 1$ be such that

$$\begin{aligned}c^* &\geq (t_3)^{1-a_k-\varepsilon}, \\ c^* &\geq c' c_f 2^k (4/\varepsilon).\end{aligned}$$

Define

$$S = \{t \geq t_2 : \delta(X,t) > c^* t^{a_k+\varepsilon} \text{ or } \delta^*(t) > c' t^{\varepsilon}\}$$

and

$$T = \begin{cases} \min S & \text{if } S \neq \varnothing, \\ \infty & \text{else.}\end{cases}$$

Item (1) of the lemma will follow if we prove that, wep in t_2, $T = \infty$. Note that $T \geq t_3$, since $c^* t^{a_k+\varepsilon} \geq t$ and $c' t^{a_\ell+\varepsilon} \geq t$ for all $t \leq t_3$.

Fix t so that $t_2 \le t < T$. Then $\delta(X,t) - \delta(X,t_2)$ is bounded above by the sum Z_t of independent Bernoulli variables with probabilities $p_t = (a_k + \varepsilon/2)c^*(t-1)^{a_k+\varepsilon-1}$, and

$$\mathbb{E}(Z_t) \le c^* \left(\frac{a_k + \varepsilon/2}{a_k + \varepsilon}\right) t^{a_k+\varepsilon}.$$

By (4.2), for all $\gamma > 0$,

$$\mathbb{P}(Z_t \ge \mathbb{E}(Z_t) + \gamma) \le \exp\left(-\frac{\gamma^2}{2(\mathbb{E}(Z_t) + \gamma/3)}\right).$$

Setting $\gamma = c^* \frac{\varepsilon/2}{a_k+\varepsilon} t^{a_k+\varepsilon} - t_2$, and using the facts that $\delta(X,t_2) \le t_2$ and that $c^* t^{a_k+\varepsilon} - \mathbb{E}(Z_t) \ge \frac{\varepsilon/2}{a_k+\varepsilon} t^{a_k+\varepsilon}$, we obtain by (4.2) that

$$\begin{aligned} \mathbb{P}(\delta(X,t) \ge c^* t^{a_k+\varepsilon}) &\le \mathbb{P}(Z_t \ge c^* t^{a_k+\varepsilon} - t_2) \\ &\le \exp\left(-\frac{\gamma^2}{2\mathbb{E}(Z_t) + \gamma/3}\right). \end{aligned}$$

Since

$$\gamma = \left(\frac{\varepsilon/2}{a_k + \varepsilon} c^* - t_2 t^{-a_k-\varepsilon}\right) t^{a_k+\varepsilon},$$

it follows from (4.8) that $\gamma > 0$ for all $t > t_3$, and $\gamma = \Theta(t^{a_k+\varepsilon})$. Since $\mathbb{E}(Z_t) = O(t^{a_k+\varepsilon})$, for all $t_3 \le t \le T$,

$$\mathbb{P}(\delta(X,t) \ge c^* t^{a_k+\varepsilon}) \le \exp(-ct^{a_k+\varepsilon})$$

for some constant $c > 0$. Thus,

$$\begin{aligned} &\mathbb{P}(\text{For some } t_2 < t \le T,\ \delta(X,t) \ge c^* t^{a_k+\varepsilon}) \\ &\qquad = \mathbb{P}(\text{For some } t_3 < t \le T,\ \delta(X,t) \ge c^* t^{a_k+\varepsilon}) \\ &\qquad \le \sum_{\tau=t_3}^{T} \exp(-c\tau^{a_k-\varepsilon}). \end{aligned} \tag{4.9}$$

Using (4.9) and the induction hypothesis, we have that the probability that T is finite is exponentially decreasing in t_2. Item (1) now follows.

We now prove item (2) of the lemma. Let X be a set of cardinality $k = n_{p,s}+1$. The probability of $B_X(t+1)$ can be bounded exactly the same as above. So from (4.7),

$$\begin{aligned} \mathbb{P}(B_X(t+1) | \delta(X,t) = 0) &\le f(t) \sum_{A \subset X} t^{-\overline{s}|X-A|} \left(\frac{\delta(A,t)}{t}\right) \\ &\le f(t)\delta^*(t) \sum_{\ell=0}^{k-1} \binom{k}{\ell} t^{-\overline{s}(k-\ell)} t^{a_\ell - 1} \\ &\le 2^k f(t)\delta^*(t) t^{a_{n_{p,s}} - \overline{s} - 1}, \end{aligned} \tag{4.10}$$

where the last inequality follows from (2.7):

$$a_{n_{p,s}} = a_{k-1} \geq a_\ell + \overline{s}(k-1-\ell) \text{ for all } 0 \leq \ell \leq k-1.$$

Let $\varepsilon = \overline{s} - a_{n_{p,s}}$. It follows from the definition of $n_{p,s}$ that $\varepsilon > 0$. By item (1) of the lemma, for all $t_2 > t_1$ there exists a constant $c > 0$ so that, wep in t_2, $\delta^*(t) \leq ct^{\varepsilon/2}$ for all $t \geq t_2$.

Let t_2 and c be such that, with positive probability p_1, for all $t \geq t_2$, $\delta^*(t) \leq ct^{\varepsilon/2}$. Define

$$S_1 = \{t \geq t_2 : \delta^*(t) > ct^{\varepsilon/2}\},$$

and

$$T_1 = \begin{cases} \min S_1 & \text{if } S_1 \neq \varnothing, \\ \infty & \text{else.} \end{cases}$$

Observe that, for all $t \geq t_2$,

$$\mathbb{P}(T_1 \geq t) \geq p_1. \tag{4.11}$$

Let $c^* > 0$ be so that $f(t) \leq c^* c^{-1} 2^{-k}$ for all $t \geq t_2$, and let

$$\beta = 1 + \varepsilon/2 > 1.$$

Then by (4.10) for all $t > t_2$,

$$\mathbb{P}(B_X(t+1) | \delta(X,t) = 0) \leq c^* t^{-\beta}. \tag{4.12}$$

Choose t_2 to satisfy the previous requirements and be sufficiently large so that $c^*(t_2)^{-\beta} < 1$. Thus

$$\begin{aligned} \mathbb{P}\left(\bigcap_{t_2 < \tau \leq t} \overline{B_X(\tau)} |\ \delta(X,t_2) = 0 \right) \\ &= \prod_{\tau=t_2}^{t} \mathbb{P}(\overline{B_X(\tau+1)} | \delta(X,\tau) = 0) \\ &\geq \mathbb{P}(T_1 \geq t) \prod_{\tau=t_2}^{t} \mathbb{P}(\overline{B_X(\tau+1)} | \delta(X,\tau) = 0, \tau \leq T_1) \\ &\geq p_1 \prod_{\tau=t_2}^{t} (1 - c^* \tau^{-\beta}), \end{aligned}$$

where the final inequality follows by (4.11) and (4.12). Since $\beta > 1$ and by the choice of c^*, the last product is bounded by a constant in $(0,1)$. By taking the limit as $t \to \infty$, (2.10) follows. □

Acknowledgments. Part of the research for this paper was conducted while the authors were visiting the Banff Research Institute as part of the Research in Teams program, and while the authors were visiting the Fields Institute for Research in Mathematical Sciences. The authors would also like to thank Colin Cooper, Fan Chung Graham, Nauzer Kalyaniwalla, Pawel Pralat and Changping Wang for helpful discussions on the topics of this paper. The authors gratefully acknowledge support from NSERC research grants and from a MITACS grant.

References

[Adler and Mitzenmacher 01] A. Adler and M. Mitzenmacher. "Towards Compressing Web Graphs." In *Proceedings of the Data Compression Conference*, p. 203. Los Alamitos, CA: IEEE Press, 2001.

[Bollobás and Riordan 05] B. Bollobás and O. Riordan. "Slow Emergence of the Giant Component in the Growing m-out Graph." *Random Structures and Algorithms* 27:1 (2005), 1–24.

[Bonato, to appear] A. Bonato. "The Search for n-e.c. Graphs." To appear in *Contributions to Discrete Mathematics*.

[Bonato and Janssen 04a] A. Bonato and J. Janssen. "Infinite Limits of Copying Models of the Web Graph." *Internet Mathematics* 1:2 (2004), 193–213.

[Bonato and Janssen 04b] A. Bonato and J. Janssen. "Limits and Power Laws of Models for the Web Graph and Other Networked Information Spaces." In *Proceedings of Combinatorial and Algorithmic Aspects of Networking*, pp. 42–48. Lecture Notes in Computer Science 3405. Berlin-Heidelberg: Springer-Verlag, 2004.

[Bonato et al. 09] A. Bonato, J. Janssen, and C. Wang. "The n-ordered Graphs—A New Graph Class." *Journal of Graph Theory* 60:3 (2009), 173–256.

[Brightwell and Winkler 00] G. Brightwell and P. Winkler. "Gibbs Measures and Dismantlable Graphs." *Journal of Combinatorial Theory Series B* 78:1 (2000), 141–169.

[Cameron 97] P. J. Cameron. "The Random Graph." In *The Mathematics of Paul Erdős*, edited by (R. L. Graham and J. Nešetřil), pp. 331–351. Berlin-Heidelberg: Springer-Verlag, 1997.

[Cameron 01] P. J. Cameron. "The Random Graph Revisited." In *European Congress of Mathematics, Barcelona, July 10-14, 2000, Volume II*, edited by C. Casacuberta, R. M. Mir-Roig, J. Verdera, and S. Xamb-Descamps, pp. 267–274. Basel: Birkhäuser, 2001.

[Chung et al. 03] F. R. K. Chung, G. Dewey, D. J. Galas, and L. Lu. "Duplication Models for Biological Networks." *Journal of Computational Biology* 10:5 (2003), 677–688.

[Chung and Lu 06] F. R. K. Chung and L. Lu. *Complex Graphs and Networks*. Providence, RI: American Mathematical Society, 2006.

[Erdős and Rényi 63] P. Erdős and A. Rényi. "Asymmetric Graphs." *Acta Mathematica Academiae Scientiarum Hungaricae* 14 (1963), 295–315.

[Grimmett and Stirzaker 01] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford, UK: Oxford University Press, 2001

[Hell and Nešetřil 04] P. Hell and J. Nešetřil. *Graphs and Homomorphisms*. Oxford, UK: Oxford University Press, 2004.

[Janson et al. 00] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. New York: Wiley-Interscience, 2000.

[Kleinberg and Kleinberg 05] J. Kleinberg and R. Kleinberg. "Isomorphism and Embedding Problems for Infinite Limits of Scale-Free Graphs." In *Proceedings of ACM-SIAM Symposium on Discrete Algorithms*, pp. 277–286. Philadelphia, PA: SIAM, 2005.

[Kim et al. 02] J. Kim, P. L. Krapivsky, B. Kahng, and S. Redner. "Infinite-Order Percolation and Giant Fluctuations in a Protein Interaction Network." *Phys. Rev. E* 66:5 (2002), 055101.

[Krapivsky and Redner 05] P. L. Krapivsky and S. Redner. "Network Growth by Copying." *Phys. Rev. E* 71:3 (2005), 036118.

[Kumar et al. 00] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. "Stochastic Models for the Web Graph." In *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science*, p. 57. Los Alamitos, CA: IEEE Press, 2000.

Anthony Bonato, Department of Mathematics, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada (abonato@ryerson.ca)

Jeannette Janssen, Department of Mathematics and Statistics Dalhousie University, Halifax, NS, B3H 3J5, Canada (janssen@mathstat.dal.ca)

Received March 15, 2007; accepted August 7, 2007.