Local Clustering Coefficient of Spatial Preferential Attachment Model

Lenar Iskhakov[*]

lenar-iskhakov@yandex.ru

Bogumił Kamiński[†]

bogumil.kaminski@gmail.com

Maksim Mironov[*]

maxim-m94@mail.ru

Paweł Prałat[§]

pralat@ryerson.ca

Liudmila Prokhorenkova[*][¶]

ostroumova-la@yandex.ru

Abstract

In this paper, we study the clustering properties of the Spatial Preferential Attachment (SPA) model introduced by Aiello et al. in 2009. This model naturally combines geometry and preferential attachment using the notion of spheres of influence. It was previously shown in several research papers that graphs generated by the SPA model are similar to real-world networks in many aspects. Also, this model was successfully used for several practical applications. However, the clustering properties of the SPA model were not fully analyzed. The clustering coefficient is an important characteristic of complex networks which is tightly connected with its community structure. In the current paper, we study the behaviour of $C(d)$, which is the average local clustering coefficient for the vertices of degree d. It was empirically shown that in real-world networks $C(d)$ usually decreases as d^{-a} for some $a > 0$ and it was often observed that $a = 1$. We prove that in the SPA model $C(d)$ decreases as $1/d$. Furthermore, we are also able to prove that not only the average but the individual local clustering coefficient of a vertex v of degree d behaves as $1/d$ if d is large enough. The obtained results further confirm the suitability of the SPA model for fitting various real-world complex networks.

Keywords: graph theory; complex networks; spatial preferential attachment; local clustering coefficient

1 Introduction

The evolution of complex networks attracted a lot of attention in recent years. Empirical studies of different real-world networks have shown that such networks have some typical properties: small diameter, power-law degree distribution, clustering structure, and others (Costa et al. 2007; Newman 2003b). Therefore, numerous random graph models have been proposed to reflect and predict such

---

[*]Advanced Combinatorics and Network Applications Lab, Moscow Institute of Physics and Technology, Moscow, Russia

[†]SGH Warsaw School of Economics, Warsaw, Poland

[‡]Department of Mathematics, Ryerson University, Toronto, ON, Canada

[§]Corresponding author

[¶]Machine intelligence and research department, Yandex, Moscow, Russia

quantitative and topological aspects of growing real-world networks (Boccaletti et al. 2006; Bollobás and Riordan 2003).

The most well studied property of complex networks is their vertex degree distribution. For the majority of studied real-world networks, the degree distribution follows a power law with a parameter γ which usually belongs to $(2,3)$ (Barabási and Albert 1999; Faloutsos et al. 1999; Newman 2005).

Another important property of real-world networks is their clustering (or community) structure. The presence of such structure highly affects many processes occurring in networks like promotion of products via viral marketing or the spreading of computer viruses and infection diseases. One way to characterize the presence of clustering structure is to measure the *clustering coefficient*, which is, roughly speaking, the probability that two neighbours of a vertex are connected. There are two well-known formal definitions: the global clustering coefficient and the average local clustering coefficient (see Section 3 for definitions). At some point, it was believed that for many real-world networks both the average local and the global clustering coefficients tend to non-zero limit as the network becomes large; for example, some numerical values can be found in Newman (2003b); however, this statement for the global clustering coefficient is questionable and recently some contradicting theoretical results were presented in Ostroumova Prokhorenkova (2016).

In this paper, we mostly focus on the behaviour of $C(d)$, which is the average local clustering coefficient for the vertices of degree d. The function $C(d)$ gives a better insight into the network structure than just the average clustering coefficient. It was empirically shown that in real-world networks $C(d)$ usually decreases as $d^{-\psi}$ for some $\psi > 0$ (Csányi and Szendrői 2004; Leskovec 2008; Serrano and Boguná 2006a; Vázquez et al. 2002). In particular, for many studied networks, $C(d)$ scales as d^{-1} (Ravasz and Barabási 2003). Moreover, it was shown in Serrano and Boguná (2006a,b) that the behaviour of $C(d)$ is tightly connected to the notion of weak and strong transitivity, which, in turn, affects percolation properties of a network.

We study the clustering properties of the *Spatial Preferential Attachment* (SPA) model introduced in Aiello et al. (2009). This model combines geometry and preferential attachment; the formal definition is given in Section 2.1. It was previously shown that graphs generated by the SPA model are similar to real-world networks in many aspects. For example, it was proven in Aiello et al. (2009) that the vertex degree distribution follows a power law. Also, this model was successfully used for several practical applications like the analysis of a duopoly market (Kamiński et al. 2017). More details on the properties and applications of the SPA model are given in Section 2.2. However, the clustering coefficient $C(d)$, which is an extremely important characteristic, was not previously analyzed for this model. Some basic clustering properties of the closely related model were analyzed: it is proved in Jacob and Mörters (2013) and Jacob et al. (2015) that the average local clustering coefficient converges in probability to a strictly positive limit; also, the global clustering coefficient converges to a nonnegative limit, which is nonzero if and only if the power-law degree distribution has a finite variance.

A short version of this article was published as a conference proceeding (Iskhakov et al. 2018) and was mostly focused on the behaviour of the clustering coefficient on simulated graphs. In the current

paper we provide a thorough theoretical analysis of the asymptotic properties of the SPA model. However, we also include some empirical results from (Iskhakov et al. 2018) to illustrate the obtained theoretical counterparts.

The rest of the paper is organized as follows. In the next section we define the SPA model and discuss its properties. Then, in Section 3, we define the local clustering coefficient. The obtained theoretical results are discussed in Section 4 and illustrated on simulated graphs in Section 5. All proofs are given in Section 6. Section 7 concludes the paper.

2 Spatial Preferential Attachment model

2.1 Definition

This paper focuses on the *Spatial Preferential Attachment* (SPA) model, which was first introduced by Aiello et al. (2009). This model combines preferential attachment with geometry by introducing "spheres of influence" whose volume grows with the degree of a vertex. The parameters of the model are the *link probability* $p \in [0,1]$ and two constants A_1, A_2 such that $0 < A_1 < \frac{1}{p}$, $A_2 > 0$. All vertices are placed in the m-dimensional unit hypercube $S = [0,1]^m$ equipped with the torus metric derived from any of the L_k norms, i.e.,

$$d(x,y) = \min \left\{ ||x - y + u||_k \, : \, u \in \{-1,0,1\}^m \right\} \quad \forall x, y \in S \,.$$

The SPA model generates a sequences random directed graphs $\{G_t\}$, where $G_t = (V_t, E_t)$, $V_t \subseteq S$. Let $\deg^-(v, t)$ be the in-degree of the vertex v in G_t, and $\deg^+(v, t)$ its out-degree. Then, the *sphere of influence* $S(v, t)$ of the vertex v at time $t \geq 1$ is the ball centered at v with the following volume:

$$|S(v,t)| = \min \left\{ \frac{A_1 \deg^-(v,t) + A_2}{t}, 1 \right\} \,.$$

In order to construct a sequence of graphs we start at $t = 0$ with G_0 being the null graph. At each time step t we construct G_t from G_{t-1} by, first, choosing a new vertex v_t *uniformly at random* from S and adding it to V_{t-1} to create V_t. Then, independently, for each vertex $u \in V_{t-1}$ such that $v_t \in S(u, t-1)$, a directed link (v_t, u) is created with probability p. Thus, the probability that a link (v_t, u) is added in time-step t equals $p|S(u, t-1)|$. See Figure 2.1 for a drawing of a simulation of the SPA model (Aiello et al. 2009).

2.2 Properties and applications of the model

In this section, we briefly discuss previous studies on properties and applications of the SPA model. This model is known to produce scale-free networks, which exhibit many of the characteristics of real-life networks (see Aiello et al. 2009; Cooper et al. 2014). Specifically, Aiello et al. (2009) (Theorem 1.1) proved that the SPA model generates graphs with a power law in-degree distribution with coefficient $1 + 1/(pA_1)$. On the other hand, the average out-degree is asymptotic to $pA_2/(1 - pA_1)$ (see Theorem 1.3 in Aiello et al. 2009). In Janssen et al. (2013b), some properties of common neighbours were used
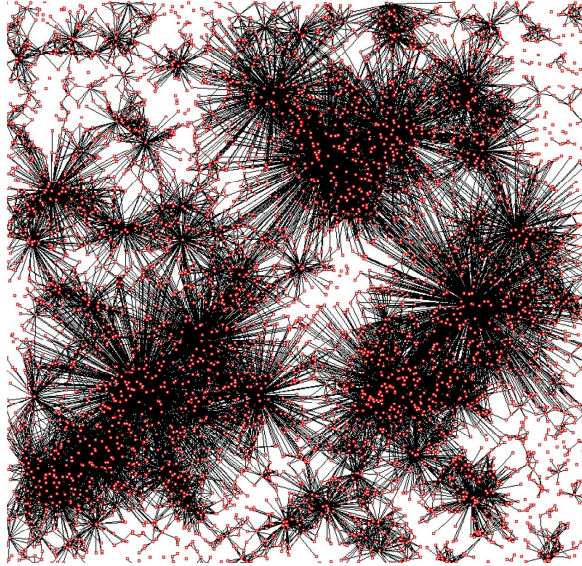
Figure 1: A simulation on the unit square with $t = 5000$, $p = 1$, and $A_1 = A_2 = 1$.

to explore the underlying geometry of the SPA model and quantify vertex similarity based on the distance in the space. Usually, the distribution of vertices in S is assumed to be uniform (Janssen et al. 2013b), but Janssen et al. (2016) also investigated non-uniform distributions, which is clearly a more realistic setting.

Let us briefly discuss the parameters of the model: p, A_1, and A_2. The parameter p is usually highly influenced by the application; for example, if one wants to model the citation network, p would correspond to the ratio of the average number of papers cited and the number of papers a typical author is aware of (presumably larger for, say, computer science papers; lower for mathematics, etc.). Then, A_1 controls the degree distribution and A_2 can be used to tune the average degree.

Importantly, in Janssen et al. (2013a), it was shown that the SPA model gave the best fit, in terms of graph structure, for a series of social networks derived from Facebook. This means that this model is a good synthetic approximation for some real-world networks and can be suitable for various practical applications. For instance, the SPA model was used to analyze a duopoly market on which there is uncertainty of a product quality (Kamiński et al. 2017) and to model the interpersonal network of top managers (Morgan et al. 2018). Fitting the SPA model to real-world networks can also potentially be used for link prediction problems. The model is especially useful when some underlying structure that affects the network is not easily measurable but is important for the application. Consider, for example, a citation network in which vertices correspond to papers and directed edges correspond to citations between them. Clearly, the content of the paper (that is, location of the corresponding vertex in S) affects edges between vertices but very often the content is not provided. Assuming that the citation network is similar to the SPA model (and after a careful tuning of parameters), one can use the model and results from Janssen et al. (2013b, 2016) to predict the similarity between papers and then use some clustering algorithm to extract papers on a similar topic. The same methodology

applies to a more sophisticated scenarios such as predicting people's taste/hobbies/believes based on social networks they are part of.

As we already mentioned, one of the most important properties of real-world networks is their clustering structure: it highly affects many processes occurring in networks. Therefore, to understand the suitability of the SPA model for various applications, it is crucial to analyze its clustering structure. The first step in this direction was made in Ostroumova Prokhorenkova et al. (2017), where modularity of the SPA model was investigated, which is a global criterion to define communities and a way to measure the presence of community structure in a network. In the current paper we analyze the clustering structure using another characteristic — the average local clustering coefficient $C(d)$.

3   Clustering coefficient

Clustering coefficient measures how likely two neighbours of a vertex are connected by an edge. There are several definitions of clustering coefficient proposed in the literature (see, e.g., Bollobás and Riordan (2003)).

The *global clustering coefficient* $C_{glob}(G)$ of a graph G is the ratio of three times the number of triangles to the number of pairs of adjacent edges in G. In other worlds, if we sample a random pair of adjacent vertices in G, then $C_{glob}(G)$ is the probability that these three vertices form a triangle. The global clustering coefficient in the SPA model was previously studied in (Jacob and Mörters 2013; Jacob et al. 2015) and it was proven that $C_{glob}(G_n)$ converges to a limit, which is positive if and only if the power-law degree distribution has a finite variance.

In this paper, we focus on the *local clustering coefficient*. Let us first define it for an undirected graph $G = (V, E)$. Let $N(v)$ be the set of neighbours of a vertex v, $|N(v)| = \deg(v)$. For any $B \subseteq V$, let $E(B)$ be the set of edges in the graph induced by the vertex set B; that is,

$$E(B) = \{\{u, w\} \in E : u, w \in B\}.$$

Finally, *clustering coefficient* of a vertex v is defined as follows:

$$c(v) = |E(N(v))| \Big/ \binom{\deg(v)}{2}.$$

Clearly, $0 \leq c(v) \leq 1$.

Note that the local clustering $c(v)$ is defined individually for each vertex and it can be noisy, especially for the vertices of not too large degrees. Therefore, the following characteristic was extensively studied in the literature. Let $C(d)$ be the local clustering coefficient averaged over the vertices of degree d; that is,

$$C(d) = \frac{\sum_{v:\deg(v)=d} c(v)}{|\{v : \deg(v) = d\}|}.$$

Further in the paper we will also use the notation $c(v, t)$ and $C(d, t)$ referring to graphs on t vertices.

The local clustering $C(d)$ was extensively studied both theoretically and empirically. For example, it was observed in a series of papers that in real-world networks $C(d) \propto d^{-\varphi}$ for some $\varphi > 0$. In

particular, Ravasz and Barabási (2003) shows that $C(d)$ can be well approximated by d^{-1} for four large networks, Vázquez et al. (2002) obtains power-law in a real network with parameter 0.75, while Csányi and Szendrői (2004) obtain $\varphi = 0.33$. The local clustering coefficient was also studied in several random graph models of complex networks. For instance, it was shown in Dorogovtsev et al. (2002); Krot and Ostroumova Prokhorenkova (2015); Newman (2003a) that some models have $C(d) \propto d^{-1}$. As we prove in this paper, similar behaviour is observed in the SPA model.

Recall that the graph G_t constructed according to the SPA model is directed. Therefore, we first analyze the directed version of the local clustering coefficient and then, as a corollary, we obtain the corresponding results for the undirected version. Let us now define the directed clustering coefficient. By $N^-(v, t) \subseteq V_t$ we denote the set of in-neighbours of a vertex v at time t; $\deg^-(v, t) = |N^-(v, t)|$. So, the directed clustering coefficient of vertex v at time t is defined as

$$c^-(v, t) = |E(N^-(v, t))| \Big/ \binom{\deg^-(v, t)}{2},$$

where this time $E(B) = \{(u, w) \in E : u, w \in B\}$ for any $B \subseteq V_t$. Similarly to the undirected case, we define

$$C^-(d, t) = \frac{\sum_{v : \deg^-(v, t) = d} c^-(v, t)}{|\{v : \deg^-(v, t) = d\}|}.$$

4   Results

Let us start with introducing some notation. As typical in random graph theory, all results in this paper are asymptotic in nature; that is, we aim to investigate properties of G_n for n tending to infinity. We say that an event holds *asymptotically almost surely* (a.a.s.) if it holds with probability tending to one as $n \to \infty$. Also, given a set S we say that *almost all* elements of S have some property P if the number of elements of S that do not have P is $o(|S|)$. Finally, we emphasize that the notations $o(\cdot)$ and $O(\cdot)$ refer to functions of n, not necessarily positive, whose growth is bounded. We use the notations $f \ll g$ for $f = o(g)$ and $f \gg g$ for $g = o(f)$. We also write $f(n) \sim g(n)$ if $f(n)/g(n) \to 1$ as $n \to \infty$ (that is, when $f(n) = (1 + o(1))g(n)$).

Let us first consider the directed clustering coefficient. It turns out that for the SPA model we are able not only to prove the asymptotics for $C^-(d, n)$, which is the average clustering over all vertices of in-degree d, but also analyze the individual clustering coefficients $c^-(v, n)$, which is a much stronger result. However, in order to do this, we need to assume that $\deg^-(v, n)$ is large enough.

From technical point of view, it will be convenient to partition the set of contributing edges, $E(N^-(v, n))$, and independently consider edges to "old" and to "young" neighbours of v. Formally, let us take any function $\omega(n)$ that tends to infinity (arbitrarily slowly) as $n \to \infty$; for example, $\omega(n) = \log \log \log n$ or even $\omega(n)$ could be the inverse Ackermann function that is less than 5 for any practical input size n. The function $\omega(n)$ will remain fixed throughout the rest of the paper. Let $\hat{T}_v$ be the smallest integer t such that $\deg^-(v, t)$ exceeds $\omega \log n$ (or $\hat{T}_v = n$ if $\deg^-(v, n) < \omega \log n$). The reason for introducing $\hat{T}_v$ is that the behaviour of vertices is chaotic and unpredictable at first but it stabilizes once they accumulate enough neighbours; the threshold happens to be around $\log n$.

Vertices in $N^-(v, \hat{T}_v)$ are called *old neighbours of* v; $N^-(v, n) \setminus N^-(v, \hat{T}_v)$ are *new neighbours of* v. So, we can partition $E(N^-(v, n))$ into $E_{old}(N^-(v, n))$ and $E_{new}(N^-(v, n))$, the first contains the edges going from neighbours of v to its old neighbours, the second contains the remaining edges, i.e., ones connecting only new neighbours. Again, the reason for partitioning $E(N^-(v, n))$ is that old neighbours of v are unpredictable (but, fortunately, there are few of them); on the other hand, the behaviour of young neighbours can be well understood. Formally,

$$E_{old}(N^-(v, n)) = \{(u, w) \in E_n : u \in N^-(v, n), w \in N^-(v, \hat{T}_v)\},$$

$$E_{new}(N^-(v, n)) = E(N^-(v, n)) \setminus E_{old}(N^-(v, n));$$

and

$$c^-(v, n) = c_{old}(v, n) + c_{new}(v, n), \tag{1}$$

where

$$c_{old}(v, n) = |E_{old}(N^-(v, n))| \Big/ \binom{\deg^-(v, n)}{2},$$

$$c_{new}(v, n) = |E_{new}(N^-(v, n))| \Big/ \binom{\deg^-(v, n)}{2}.$$

Let us start with the following theorem which is extensively used in our reasonings and is interesting and important on its own. Variants of this results were proved in (Janssen et al. 2013b, 2016); here, we present a slightly modified statement from (Janssen et al. 2016), adjusted to our current needs. We provide the proof in Section 6.1 for completeness.

Theorem 4.1. *Let* $\omega = \omega(n)$ *be any function tending to infinity together with* n. *The following holds with probability* $1 - o(n^{-4})$. *For any vertex* v *with*

$$\deg^-(v, n) = k = k(n) \geq \omega \log n$$

and for all values of t *such that*

$$n \left(\frac{\omega \log n}{k} \right)^{\frac{1}{pA_1}} =: T_v \leq t \leq n,$$

we have

$$\deg^-(v, t) \sim k \left(\frac{t}{n} \right)^{pA_1}.$$

The expression for $T_v = T_v(n)$ is chosen so that at time T_v vertex v has $(1 + o(1))\omega \log n$ neighbours a.a.s. The implication of this theorem is that once a vertex accumulates $\omega \log n$ neighbours, its behaviour can be predicted with high probability until the end of the process (that is, till time n) when its degree reaches k. In other words the following property holds. For a fixed value of n, let $M = M(n)$ and $m = m(n)$ be the maximum and, respectively, the minimum ratio between $\deg^-(v, t)$ and the deterministic function $k(t/n)^{pA_1}$ (taken over the interval $T_v \leq t \leq n$). Then, a.a.s. both M and m tend to one as $n \to \infty$.

This property can be used to show that the contribution to $c^-(v, n)$ coming from edges to new neighbours of v is well concentrated.

Theorem 4.2. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. Then, with probability $1 - o(n^{-1})$ for any vertex v with*

$$\deg^-(v, n) = k = k(n) \geq (\omega \log n)^{4 + (4pA_1 + 2)/(pA_1(1 - pA_1))}$$

we have

$$c_{new}(v, n) = \Theta(1/k).$$

Unfortunately, if a vertex v lands in a densely populated region of S, it might happen that $c_{old}(v, b)$ is much larger than $1/k$. We show the following 'negative' result (without trying to aim for the strongest statement) that shows that there is no hope for extending Theorem 4.2 to $c^-(v, n)$.

Theorem 4.3. *Let $C = 5 \log(1/p)$ and $\xi = \xi(n) = 1/(\omega(\log \log n)^2(\log \log \log n)) = o(1)$ for some $\omega = \omega(n)$ tending to infinity as $n \to \infty$. Suppose that $k = k(n)$ is such that $2 \leq k \leq n^\xi$. Then, a.a.s., there exists a vertex v such that $\deg^-(v, n) \sim k$ and*

(i) $c^-(v, n) = 1$, provided that $2 \leq k \leq \sqrt{\log n/C}$,

(ii) $c^-(v, n) = \Omega(1) \gg 1/k$, provided that $\sqrt{\log n/C} \leq k \leq \log n/\log \log n$,

(iii) $c^-(v, n) \gg (\log \log n)^2(\log \log \log n)/k \gg 1/k$, provided that $\log n/\log \log n \leq k \leq n^\xi$.

On the other hand, Theorem 4.2 implies immediately the following corollary.

Corollary 4.4. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-1})$. For any vertex v for which*

$$\deg^-(v, n) = k = k(n) \geq (\omega \log n)^{4 + (4pA_1 + 2)/(pA_1(1 - pA_1))}$$

it holds that

$$\begin{aligned}
c^-(v, n) &\geq c_{new}(v, n) = \Omega(1/k) \\
c^-(v, n) &= c_{old}(v, n) + c_{new}(v, n) = O(\omega \log n/k) + O(1/k) = O(\omega \log n/k).
\end{aligned}$$

This corollary states that for *all* vertices with large enough degree k we have the desired lower bound $\Omega(1/k)$ for the clustering coefficient $c^-(v, n)$. On the other hand, the upper bound, which is $O(\omega \log n/k)$, grows faster than desired due to the presence of $c_{old}(v, n)$. However, despite the 'negative' result (stated in Theorem 4.3), almost all vertices (of large enough degrees) have clustering coefficients of order $1/k$. Below is a precise statement. The conclusions in cases (i)' and (ii)' follow immediately from Theorem 4.2.

Theorem 4.5. *Let $\varepsilon, \delta \in (0, 1/2)$ be any two constants, and let $k = k(n) \leq n^{pA_1 - \varepsilon}$ be any function of n. Let X_k be the set of vertices of G_n of in-degree between $(1 - \delta)k$ and $(1 + \delta)k$. Then, a.a.s., the following holds.*

(i) *Almost all vertices in X_k have $c_{old}(v, n) = O(1/k)$, provided that $k \gg \log^{C_1} n$, where $C_1 = (1 + (2 + \varepsilon)pA_1)/(1 - pA_1)$.*

(i)' *As a result, almost all vertices in X_k have $c^-(v, n) = \Theta(1/k)$, provided that $k \gg \log^C n$, where $C = 4 + (4pA_1 + 2)/(pA_1(1 - pA_1))$.*

(ii) *The average clustering coefficient $c_{old}(v, n)$ of vertices in X_k is $O(1/k)$; that is,*

$$\frac{1}{|X_k|} \sum_{v \in X_k} c_{old}(v, n) = O(1/k),$$

provided that $k \gg \log^{C_2} n$, where $C_2 = (1 + (2 + pA_1 + \varepsilon)pA_1)/(1 - pA_1)$.

(ii)' *As a result, the average clustering coefficient $c^-(v, n)$ of vertices in X_k is $\Theta(1/k)$; that is,*

$$\frac{1}{|X_k|} \sum_{v \in X_k} c^-(v, n) = \Theta(1/k),$$

provided that $k \gg \log^C n$, where $C = 4 + (4pA_1 + 2)/(pA_1(1 - pA_1))$.

Finally, let us discuss the undirected case. The following corollary holds.

Corollary 4.6. *Let $c(v, n)$ be the clustering coefficient defined for the undirected graph $\hat{G}_n$ obtained from G_n by considering all edges as undirected. Then Corollary 4.4 and Theorem 4.5 hold with replacing $c^-(v, n)$ by $c(v, n)$.*

Indeed, according to Lemma 6.6 (see Section 6.2) a.a.s. the out-degrees of all vertices do not exceed $\omega \log n$. Therefore, even if out-neighbours of a vertex form a complete graph, the contribution from them is at most $\binom{\omega \log n}{2}$, which is much smaller than the required lower bound for k.

Finally, let us discuss the connection between the obtained results and the empirical observations discussed in Section 3. Recall that it was observed that in real-world networks $C(d) \propto d^{-\varphi}$ for some $\varphi > 0$ and it is often the case that $\varphi = 1$. Basically, the results discussed in this section mean that in SPA model we have $\varphi = 1$ and other constants cannot be modeled. Indeed, Theorem 4.5 (ii)' (and the corresponding Corollary 4.6) state slightly weaker results than the asymptotics for $C(d)$, since instead of averaging over the vertices of degree d we average over the set X_d of vertices with degree close to d. On the other hand, Theorem 4.5 (i)' states a stronger result for individual vertices.

5 Simulations

In this section, we illustrate the theoretical, asymptotic, results presented in the previous section by analyzing the local clustering coefficient for graphs of various orders generated according to the SPA model. An efficient algorithm used to generate the SPA graphs is described in the proceeding version of this paper (Iskhakov et al. 2018).

It is proven in Theorem 4.5 that $\frac{1}{|X_d|} \sum_{v \in X_d} c^-(v, n) = \Theta(1/d)$ for $d \gg \log^C n$, where $C = 4 + (4pA_1 + 2)/(pA_1(1 - pA_1))$. In order to illustrate this result, we generated 10 graphs for each
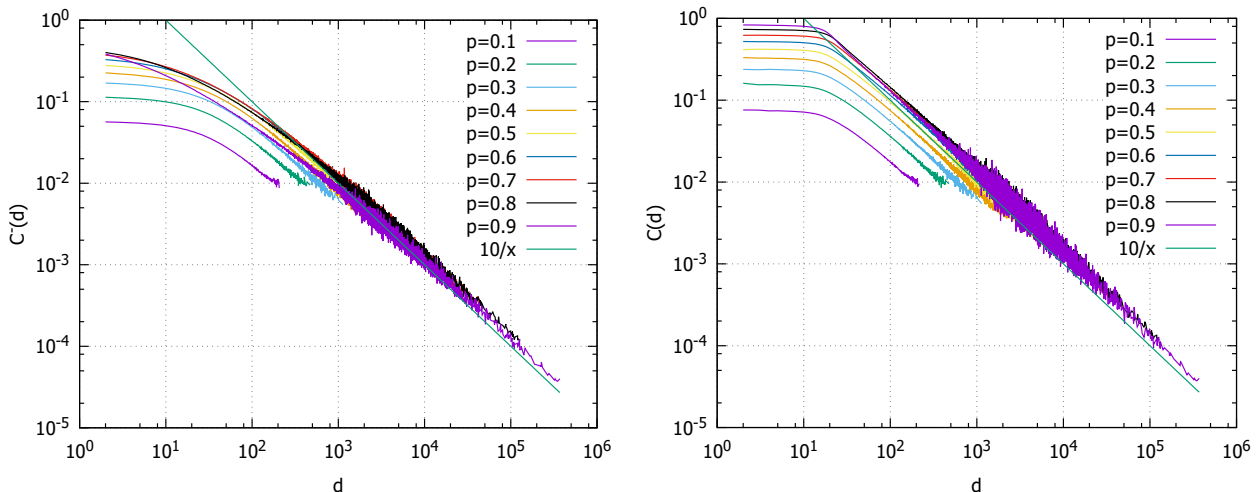
Figure 2: Average local clustering coefficient for directed (left) and undirected (right) graphs.

$p \in \{0.1, 0.2, \ldots, 0.9\}$, $A_1 = 1$, $A_2 = 10(1-p)/p$ (A_2 is chosen to fix the expected asymptotic degree equal 10) and computed the average value of $C^-(d, n)$ for $n = 10^6$, see Figure 2 (left). Similarly, Figure 2 (right) presents the same measurements for the undirected average local clustering $C(d, n)$. Note that in both cases figures agree with our theoretical results: both $C^-(d, n)$ and $C(d, n)$ decrease as c/d with some c for large enough d (we added a function $10/d$ for comparison). Note that for small p the maximum degree is small, therefore the sizes of the generated graphs are not large enough to observe a straight line in log-log scale.

Note that for all $p \in (0, 1)$ we have $C = 4 + \frac{4p+2}{p(1-p)} > 18$, so, our theoretical results are expected to hold for $d \gg \log^C n > 10^{20}$ which is irrelevant as the order of the graph is only 10^6. However, we observe the desired behaviour for much smaller values of d; that is, in some sense, our bound is too pessimistic.

Also, note that the statement $C^-(d, n) = \Theta(1/d)$ is stronger than the statement of Theorem 4.5, since in the theorem we averaged $c^-(v, n)$ over the set X_d of vertices of in-degree between $(1-\delta)d$ and $(1+\delta)d$. In order to illustrate the difference, on Figure 3 we present the smoothed curves for the directed (left) and undirected (right) local clustering coefficients averaged over X_d for $\delta = 0.1$. Note that this smoothing substantially reduce the noise observed on Figure 2.

Next, let us illustrate the fact that the number of edges between "new" neighbours of a vertex is more predictable than the number of edges going from some neighbours to "old" ones. We extensively used this difference in Section 4, where we analyzed new and old edges separately. In our experiments, we split $c^-(v, n)$ into "old" and "new" parts as in Equation (1), but now we take $\hat{T}_v$ be the smallest integer t such that $\deg^-(v, t)$ exceeds $\deg^-(v, n)/2$. As a result, we compute the average local clustering coefficients $C^-_{old}(d)$ and $C^-_{new}(d)$. Figure 4 shows that $C^-_{new}(d)$ can almost perfectly be fitted by c/d with some c, while most of the noise comes from $C^-_{old}(d)$.

Finally, Figure 5 shows the distribution of the individual local clustering coefficients for one graph generated with $p = 0.7$. Theorem 4.3 states that a.a.s. there exist a vertex v of degree d with $c^-(v, n) \gg 1/d$. Also, according to this theorem, the situation is much worse for smaller values of d.
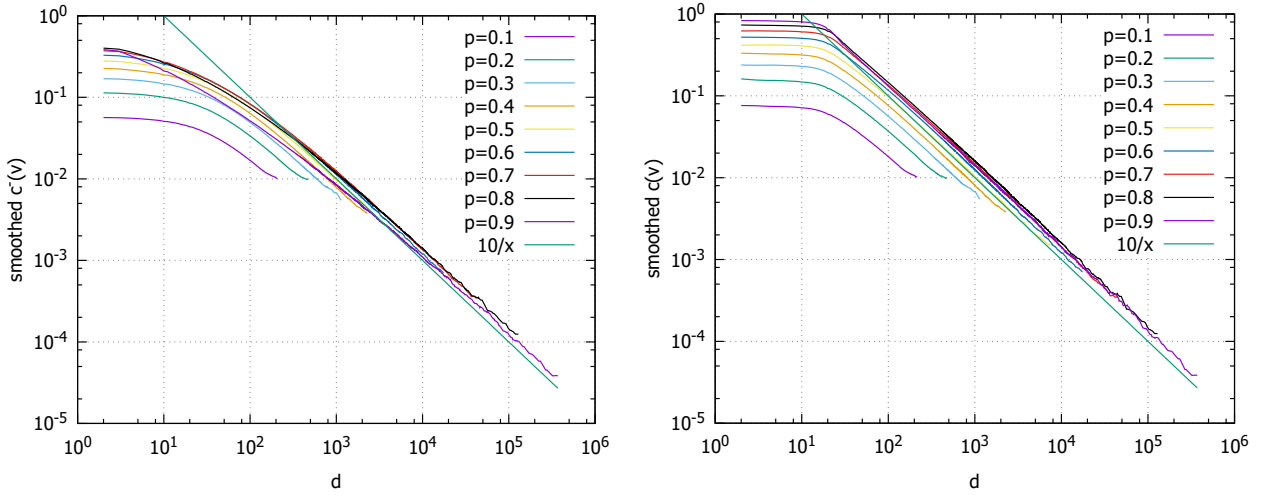
10

Figure 3: Local clustering coefficient for directed (left) and undirected (right) graphs averaged over X_d.
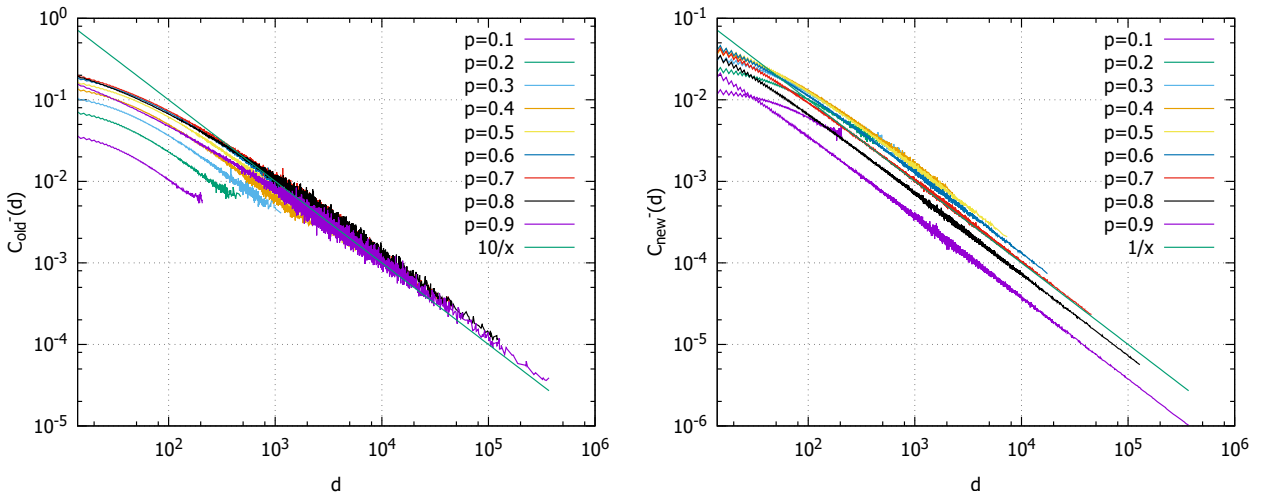

Figure 4: Comparison of "new" and "old" parts of the average local clustering coefficient.
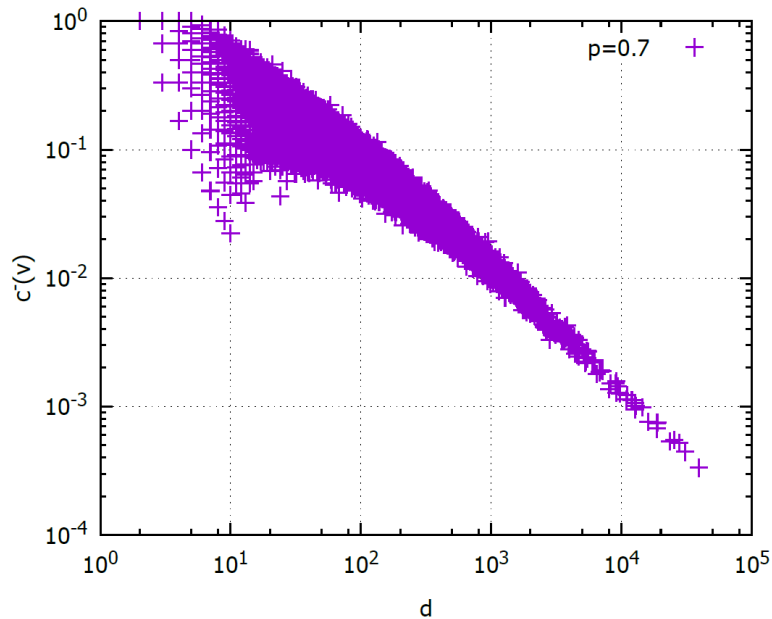
Figure 5: The distribution of individual local clustering coefficients.

Indeed, one can see on Figure 5 that for small d the scatter of points is much larger. On the other hand, in Theorem 4.5 we present bounds for $c^-(v, n)$ for almost all vertices, provided that d is large enough. One can see it on the figure too and, similarly to the previously discussed figures, we observe the expected behaviour even for relatively small n despite the bound $\log^C n$ that is bigger than n in our case.

6 Proofs

6.1 Proof of Theorem 4.1

We will use the following version of the Chernoff bound that can be found, for example, in (Janson et al. 2000, p. 27, Corollary 2.3).

Lemma 6.1. *Let X be a random variable that can be expressed as a sum of independent random indicator variables, $X = \sum_{i=1}^n X_i$, where $X_i \in \mathrm{Ber}(p_i)$ with (possibly) different $p_i = \mathbb{P}(X_i = 1) = \mathbb{E}X_i$. If $\varepsilon \leq 3/2$, then*

$$\mathbb{P}(|X - \mathbb{E}X| \geq \varepsilon \mathbb{E}X) \leq 2 \exp\left(-\frac{\varepsilon^2 \mathbb{E}X}{3}\right). \tag{2}$$

Let us start with the following key lemma.

Lemma 6.2. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. For a given vertex v, suppose that $\deg^-(v, T) = d \geq \omega \log n$. Then, with probability $1 - o(n^{-6})$, for every value of t,*

12

$T \leq t \leq 2T$,

$$\left| \deg^-(v,t) - d \cdot \left(\frac{t}{T} \right)^{pA_1} \right| \leq \frac{5}{pA_1} \cdot \frac{t}{T} \sqrt{d \log n}.$$

Of course, we will be applying the lemma to the graph generated by the SPA model. However, the statement is much more general; that is, one can apply it to any (deterministic) graph on T vertices as long as the in-degree of v in this graph is at least $\omega \log n$ and the next T vertices are added according to the rules of the SPA model. After that, the desired bounds for the in-degree of v hold with probability $1 - o(n^{-6})$.

Proof. Our goal is to estimate $\deg^-(v,t) - d \cdot (t/T)^{pA_1}$. We will show that the upper bound holds; the lower bound can be obtained by using an analogous, symmetric, argument.

Let $X_T, X_{T+1}, \ldots, X_{2T-1}$ be a sequence of independent Bernoulli random variables (with various parameters):

$$\mathbb{P}(X_t = 1) = p \, \frac{A_1 \left(d \left(\frac{t}{T} \right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t}{T} \sqrt{d \log n} \right) + A_2}{t}.$$

We will use the following standard coupling. Let $Z_T = \deg^-(v,T) = d$. For each $T < t \leq 2T$ we define $Z_t = \deg^-(v,t)$ if

$$\deg^-(v,t') \leq d \cdot \left(\frac{t'}{T} \right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t'}{T} \sqrt{d \log n}$$

for every $T \leq t' < t$; otherwise, $Z_t = Z_{t-1} + X_{t-1}$. It follows that $Z_t - d$ can be (stochastically) bounded from above by the sum $X_t' = \sum_{i=T}^{t-1} X_i$ of independent indicator random variables. Note that if T_0 is the smallest t for which $\deg^-(v,t) > d \cdot \left(\frac{t}{T} \right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t}{T} \sqrt{d \log n}$, then $\deg^-(v,t) = Z_t \leq d + X_t'$ for all $t \leq T_0$. Hence, the lemma will follow once we show that with the desired probability

$$d + X_t' \leq d \cdot \left(\frac{t}{T} \right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t}{T} \sqrt{d \log n} \tag{3}$$

for all $T \leq t \leq 2T$.

Clearly, since $pA_1 < 1$,

$$
\begin{aligned}
\mathbb{E} X_t' &= \sum_{i=T}^{t-1} \mathbb{E} X_i \\
&= pA_1 d T^{-pA_1} \left(\sum_{i=T}^{t-1} i^{pA_1 - 1} \right) + \frac{t - T}{T} 5 \sqrt{d \log n} + O(1) \\
&= d \left(\frac{t}{T} \right)^{pA_1} - d \left(\frac{T}{T} \right)^{pA_1} + \frac{t - T}{T} 5 \sqrt{d \log n} + O(1) \\
&= d \left(\frac{t}{T} \right)^{pA_1} - d + \frac{t - T}{T} 5 \sqrt{d \log n} + O(1).
\end{aligned}
$$

If (3) fails, then

$$
\begin{aligned}
X'_t &\geq \left(d \cdot \left(\frac{t}{T}\right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t}{T}\sqrt{d \log n} \right) - d \\
&= \mathbb{E}X'_t + \frac{5}{pA_1} \cdot \frac{t}{T}\sqrt{d \log n} - \frac{t-T}{T}5\sqrt{d \log n} + O(1) \\
&\geq \mathbb{E}X'_t + 5\sqrt{d \log n},
\end{aligned}
$$

using again that it is assumed that $pA_1 < 1$. It follows from the Chernoff bound (2) that

$$
\mathbb{P}(|X'_t - \mathbb{E}X'_t| \geq 5\sqrt{d \log n}) \leq 2\exp\left(-(5\varepsilon/3)\sqrt{d \log n} \right),
$$

where $\varepsilon = 5\sqrt{d \log n}/\mathbb{E}X'_t$. The maximum value of $\mathbb{E}X'_t$ corresponds to $t = 2T$ and so

$$
\begin{aligned}
\mathbb{E}X &\leq d\left(\frac{2T}{T}\right)^{pA_1} - d + \frac{2T-T}{T}5\sqrt{d \log n} + O(1) \\
&\sim d(2^{pA_1} - 1) \leq d.
\end{aligned}
$$

So $\varepsilon \geq 5\sqrt{d^{-1}\log n}$. Therefore, the probability that (3) fails for a given $T \leq t \leq 2T$ is at most $2\exp(-(25/3)\log n) = o(n^{-7})$. Hence, (3) fails for some $T \leq t \leq 2T$ with probability $o(n^{-6})$ and the proof is finished. $\qquad\square$

Now, with Lemma 6.2 in hand we can get Theorem 4.1.

Proof of Theorem 4.1. Let $\omega = \omega(n)$ be a function going to infinity with n. Let v be a vertex with final degree $k \geq \omega \log n$. Let T be the first time that the in-degree of v exceeds $(\omega/2)\log n$. Finally, let $d = \deg^-(v, T)$. We obtain from Lemma 6.2 that, with probability $1 - o(n^{-6})$,

$$
d\left(\frac{t}{T}\right)^{pA_1}\left(1 - \frac{5}{pA_1}\sqrt{d^{-1}\log n}\right) \leq \deg^-(v, t) \leq d\left(\frac{t}{T}\right)^{pA_1}\left(1 + \frac{5}{pA_1}\sqrt{d^{-1}\log n}\right)
$$

for $T \leq t \leq 2T$. It follows that the degree tends to grow but the sphere of influence tends to shrink between T and $2T$, and thus that the conditions of Lemma 6.2 again hold at time $2T$. We can now keep applying the same lemma for times $2T$, $4T$, $8T$, $16T, \ldots$, using the final value as the initial one for the next period, to get the statement for all values of t from T up to and including time n. Precisely, for $1 \leq i < i_{\max} = \lfloor \log_2 n \rfloor + 1$, let $d_i = \deg^-(v, 2^iT)$. Then by Lemma 6.2, we have for $i > 1$ that $d_i \leq d_{i-1}2^{pA_1}(1 + \varepsilon_i)$, where $\varepsilon_i = \frac{5}{pA_1}\sqrt{d_{i-1}^{-1}\log n}$. Since we apply the lemma $O(\log n)$ times (for a given vertex v), the following statement holds with probability $1 - o(n^{-5})$ from time T on: for any $2^{i-1}T \leq t < 2^iT$, we have that

$$
\deg^-(v, t) \leq d\left(\frac{t}{T}\right)^{pA_1}\prod_{j=0}^{i}(1 + \varepsilon_i).
$$

14

It remains to make sure that the accumulated multiplicative error term is still only $(1 + o(1))$. For that, let us note that

$$
\begin{aligned}
\prod_{j=0}^{i}(1 + \varepsilon_i) &= \prod_{j=1}^{i}\left(1 + \frac{5}{pA_1}\sqrt{d^{-1}2^{-pA_1 j}\log n}\right) \\
&\sim \exp\left(\frac{5}{pA_1}\sqrt{d^{-1}\log n}\sum_{j=1}^{i}2^{-pA_1 j/2}\right) \\
&= \exp\left(O(\sqrt{d^{-1}\log n})\right) \sim 1,
\end{aligned}
$$

since d grows faster than $\log n$. A symmetric argument can be used to show a lower bound for the error term and so the result holds.

It follows that we have the desired behaviour from time T. Precisely, for times $T \le t \le n$, we have that

$$
\deg^-(v, t) \sim d\left(\frac{t}{T}\right)^{pA_1},
$$

where $d = \deg^-(v, T) \sim (\omega/2)\log n$. Setting $t = n$ and $\deg^-(v, n) = k$, we obtain that

$$
T \sim \left(\frac{d}{k}\right)^{1/pA_1} n \sim \left(\frac{\omega \log n}{2k}\right)^{1/pA_1} n \sim \left(\frac{1}{2}\right)^{1/pA_1} T_v.
$$

Therefore, for large enough n, we have that $T < T_v$. As a result, we obtain that, for $T_v \le t \le n$,

$$
\deg^-(v, t) \sim k\left(\frac{t}{n}\right)^{pA_1}.
$$

Finally, since the statement holds for any vertex v with probability $1 - o(n^{-5})$, with probability $1 - o(n^{-4})$ the statement holds for all vertices. The proof of the theorem is finished. $\qquad\square$

Let us note that Theorem 4.1 immediately implies the following two corollaries.

Corollary 6.3. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-4})$. For every vertex v, and for every time T so that $\deg^-(v, T) \ge \omega \log n$, for all times t, $T \le t \le n$,*

$$
\deg^-(v, t) \sim \deg^-(v, T)\left(\frac{t}{T}\right)^{pA_1}.
$$

Corollary 6.4. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-4})$. For any vertex v_i born at time $i \ge 1$, and $i \le t \le n$ we have that*

$$
\deg^-(v_i, t) \le \omega \log n \left(\frac{t}{i}\right)^{pA_1}. \tag{4}
$$

6.2 Proof of Theorem 4.2

Let B be a ball of volume $b = b(n)$ and $t = t(n) \in \mathbb{N}$ be any function of n such that $bt \to \infty$ as $n \to \infty$. It will be crucial for the argument to understand the behaviour of the random variables $N_{i,t} = N_{i,t}(b)$ counting the number of vertices in B that are of in-degree i at time t; that is,

$$
N_{i,t} = |\{w \in B : \deg^-(w, t) = i\}|.
$$

The arguments presented below are similar to the ones in (Aiello et al. 2009) showing that the degree distribution of G_n follows a power-law.

The equations relating the random variables $N_{i,t}$ are described as follows. As G_0 is the null graph, $N_{i,0} = 0$ for $i \geq 0$. For all $t \in \mathbb{N} \cup \{0\}$, we derive that

$$\mathbb{E}(N_{0,t+1} - N_{0,t} \mid G_t) = b - N_{0,t}p\frac{A_2}{t}, \tag{5}$$

$$\mathbb{E}(N_{i,t+1} - N_{i,t} \mid G_t) = N_{i-1,t}p\frac{A_1(i-1) + A_2}{t} - N_{i,t}p\frac{A_1 i + A_2}{t}. \tag{6}$$

Recurrence relations for the expected values of $N_{i,t}$ can be derived by taking the expectation of the above equations. To solve these relations, we use the following lemma on real sequences, which is Lemma 3.1 from Chung and Lu (2006).

Lemma 6.5. *If (α_t), (β_t) and (γ_t) are real sequences satisfying the relation*

$$\alpha_{t+1} = \left(1 - \frac{\beta_t}{t}\right)\alpha_t + \gamma_t,$$

and $\lim_{t\to\infty} \beta_t = \beta > 0$ and $\lim_{t\to\infty} \gamma_t = \gamma$, then $\lim_{t\to\infty} \frac{\alpha_t}{t}$ exists and equals $\frac{\gamma}{1+\beta}$.

Applying this lemma with $\alpha_t = \mathbb{E}(N_{0,t})/b$, $\beta_t = pA_2$, and $\gamma_t = 1$ gives that $\mathbb{E}(N_{0,t}) \sim c_0 bt$ with

$$c_0 = \frac{1}{1 + pA_2}.$$

For $i \geq 1$, the lemma can be inductively applied with $\alpha_t = \mathbb{E}(N_{i,t})/b$, $\beta_t = p(A_1 i + A_2)$, and $\gamma_t = \mathbb{E}(N_{i-1,t})p(A_1(i-1) + A_2)/(bt)$ to show that $\mathbb{E}(N_{i,t}) \sim c_i bt$, where

$$c_i = c_{i-1}p\frac{A_1(i-1) + A_2}{1 + p(A_1 i + A_2)}.$$

It is straightforward to verify that

$$c_i = \frac{p^i}{1 + pA_2 + ipA_1}\prod_{j=0}^{i-1}\frac{jA_1 + A_2}{1 + pA_2 + jpA_1}.$$

The above formula implies that $c_i = (1 + o(1))ci^{-(1+1/(pA_1))}$ (as $i \to \infty$) for some constant c, so the expected proportion $N_{i,t}/(bt)$ asymptotically follows a power-law with exponent $1 + 1/(pA_1)$.

We prove concentration for $N_{i,t}$ when $i \leq i_f$ (for some function $i_f = i_f(n)$) by using a relaxation of Azuma-Hoeffding martingale techniques. The random variables $N_{i,t}$ do not a priori satisfy the c-Lipschitz condition: indeed, a new vertex may fall into many overlapping regions of influence and so it can potentially change degrees of many vertices. Nevertheless, we will prove that deviations from the c-Lipschitz condition occur with very small probability. The following lemma gives a deterministic bound for $|N_{i,t+1} - N_{i,t}|$ which holds with high probability. Indeed, it is obvious that $|N_{i,t+1} - N_{i,t}| \leq \max\{\deg^+(v_{t+1}, t+1), 1\}$. Note that a weaker bound of $\log^2 n$ was proved in Aiello et al. (2009); with Corollary 6.4 in hand, we can get slightly better bound but the argument remains the same. We present the proof for completeness.

Lemma 6.6. *Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-3})$. For every vertex v_i,*

$$\deg^+(v_i, i) = \deg^+(v_i, n) \le \omega \log n.$$

Proof. Let us focus on any $1 \le i \le n$. Since v_i is chosen uniformly at random from the unit hypercube (note that the history of the process does not affect this distribution) with the torus metric, without loss of generality, we may assume that v_i lies in the centre of the hypercube. For $1 \le j < i$, let X_j denote the indicator random variable of the event that v_j lies in the ball around v_i (or vice versa) with volume

$$\alpha_j = j^{-pA_1} i^{pA_1 - 1} \omega^{2/3} \log n.$$

By Corollary 6.4 (applied with $\omega^{1/3}$ instead of ω), we may assume that

$$\deg^-(v_j, i) \le (i/j)^{pA_1} \omega^{1/3} \log n,$$

for all $j \in [i-1]$. Note that $(A_1 \deg^-(v_j, i) + A_2)/i = o(\alpha_j)$. Hence, for all $j \in [i-1]$, $X_j = 0$ implies that v_i is not in the influence region of v_j and so there is no directed edge from v_i to v_j. Therefore, we have that

$$\deg^+(v_i, i) \le \sum_{j=1}^{i-1} X_j.$$

Since

$$\mathbb{E}\left(\sum_{j=1}^{i-1} X_j \right) = \sum_{j=1}^{i-1} \alpha_j = i^{pA_1 - 1} \omega^{2/3} \log n \sum_{j=1}^{i-1} j^{-pA_1} = O(\omega^{2/3} \log n) = o(\omega \log n),$$

the assertion follows easily from the Chernoff bound. $\qquad\square$

Now, we are ready to prove concentration for the random variables $N_{i,t}$. In order to explain the technique, we investigate $N_{0,t}$, the number of vertices of in-degree zero. The argument easily generalizes to other values of i and we explain it afterwards. We will use the supermartingale method of Pittel et al. Pittel et al. (1996), as described in Wormald (1999).

Lemma 6.7. *Let $G_0, G_1, \ldots, G_n$ be a random graph process and X_t a random variable determined by $G_0, G_1, \ldots, G_t$, $0 \le t \le n$. Suppose that for some real constants β_t and constants γ_t,*

$$\mathbb{E}(X_t - X_{t-1} \mid G_0, G_1, \ldots, G_{t-1}) < \beta_t$$

and

$$|X_t - X_{t-1} - \beta_t| \le \gamma_t$$

for $1 \le t \le n$. Then for all $\alpha > 0$,

$$\mathbb{P}\left(\text{For some } s \text{ with } 0 \le s \le n : X_s - X_0 \ge \sum_{t=1}^{s} \beta_t + \alpha \right) \le \exp\left(-\frac{\alpha^2}{2 \sum \gamma_t^2} \right).$$

Now, we are ready to prove the concentration for $N_{0,t}$.

Theorem 6.8. *Let B be a ball of volume $b = b(n)$ and $t = t(n) \in \mathbb{N}$ be any function of n such that $bt \to \infty$ as $n \to \infty$. Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-3})$.*

$$N_{0,t} = N_{0,t}(B) = \frac{bt}{1 + A_2 p} + O((bt)^{1/2}(\omega \log n)^{3/2}) = c_0 bt + O((bt)^{1/2}(\omega \log n)^{3/2}).$$

In particular, if $bt \gg \log^3 n$, then $N_{0,t} \sim c_0 bt$.

Proof. We first need to transform $N_{0,s}$ $(1 \leq s \leq t)$ into something close to a martingale. It provides some insight if we define real function $f(x)$ to model the behaviour of the scaled random variable $N_{0,xt}/t$. If we presume that the changes in the function correspond to the expected changes of the random variable (see (5)), we obtain the following differential equation

$$f'(x) = b - f(x)\frac{pA_2}{x}$$

with the initial condition $f(0) = 0$. The general solution of this equation can be put in the form

$$f(x)x^{pA_2} - \frac{bx^{1+pA_2}}{1 + pA_2} = C.$$

Consider the following real-valued function

$$H(x, y) = x^{pA_2}y - \frac{bx^{1+pA_2}}{1 + pA_2} \tag{7}$$

(note that we expect $H(s, N_{0,s})$ to be close to zero). Let $\mathbf{w}_s = (s, N_{0,s})$, and consider the sequence of random variables $(H(\mathbf{w}_s) : 1 \leq s \leq t)$. The second-order partial derivatives of H evaluated at $\mathbf{w}_s$ are all $O(s^{pA_2-1})$. Moreover, it follows from Lemma 6.6 that we may assume that

$$|N_{0,s+1} - N_{0,s}| \leq \omega \log n. \tag{8}$$

Therefore, we have

$$H(\mathbf{w}_{s+1}) - H(\mathbf{w}_s) = (\mathbf{w}_{s+1} - \mathbf{w}_s) \cdot \operatorname{grad} H(\mathbf{w}_s) + O(s^{pA_2-1}\omega^2 \log^2 n), \tag{9}$$

where "$\cdot$" denotes the inner product and $\operatorname{grad} H(\mathbf{w}_s) = (H_x(\mathbf{w}_s), H_y(\mathbf{w}_s))$.

Observe that from our choice of H, we have that

$$\mathbb{E}(\mathbf{w}_{s+1} - \mathbf{w}_s \mid G_s) \cdot \operatorname{grad} H(\mathbf{w}_s) = 0.$$

Hence, taking the expectation of (9) conditional on G_s, we obtain that

$$\beta_{s+1} = \mathbb{E}(H(\mathbf{w}_{s+1}) - H(\mathbf{w}_s) \mid G_s) = O(s^{pA_2-1}\omega^2 \log^2 n).$$

From (9) and (8), noting that

$$\operatorname{grad} H(\mathbf{w}_s) = \left(pA_2 s^{pA_2-1}N_{0,s} - bs^{pA_2}, s^{pA_2}\right),$$

18

we have that
$$\gamma_{s+1} = |H(\mathbf{w}_{s+1}) - H(\mathbf{w}_s)| \leq O(s^{pA_2}\omega \log n).$$

Our goal is to apply Lemma 6.7 to the sequence $(H(\mathbf{w}_s) : 1 \leq s \leq t)$ to get an upper bound for $H(\mathbf{w}_s)$. A symmetric argument applied to $(-H(\mathbf{w}_s) : 1 \leq s \leq t)$ will give us the desired lower bound so let us concentrate on the upper bound. The bounds for β_{s+1} and γ_{s+1} we derived above are universal; however, typically vertex v_s lies far away from the ball B so that $N_{0,s}$ is not affected. This certainly happens if the distance from the ball B to v_s is more than the radius of the ball of volume A_2/s, and so this situation occurs with probability $1 - O(b + s^{-1})$. Moreover, if this happens and $H(\mathbf{w}_s) \geq 0$, then $H(\mathbf{w}_s)$ decreases (it can be viewed as some kind of "self-correcting" behaviour); hence, since we aim for an upper bound, we may assume that $H(\mathbf{w}_s)$ does not change. It follows that

$$
\begin{aligned}
\sum_{s=1}^{t} \beta_s &= O\left(\sum_{s=1}^{t} s^{pA_2-1}\omega^2 \log^2 n \cdot (b + s^{-1})\right) \\
&= O\left(b\,\omega^2 \log^2 n \sum_{s=1}^{t} s^{pA_2-1}\right) + O\left(\omega^2 \log^2 n \sum_{s=1}^{t} s^{pA_2-2}\right) \\
&= O\left(b\,t^{pA_2}\omega^2 \log^2 n\right) + O\left(t^{pA_2-1}\omega^2 \log^2 n\right) = O\left(b\,t^{pA_2}\omega^2 \log^2 n\right),
\end{aligned}
\tag{10}
$$

since it is assumed that $bt \to \infty$. Similarly, we get that

$$
\sum_{s=1}^{t} \gamma_s^2 = O\left(\sum_{s=1}^{t}(s^{pA_2}\omega \log n)^2 \cdot (b + s^{-1})\right) = O\left(b\,t^{1+2pA_2}\omega^2 \log^2 n\right).
\tag{11}
$$

Finally, we are ready to apply Lemma 6.7 with $\alpha = b^{1/2}t^{1/2+pA_2}(\omega \log n)^{3/2}$ to obtain that with probability $1 - o(n^{-3})$,

$$|H(\mathbf{w}_t) - H(\mathbf{w}_0)| = O(\alpha) = O(b^{1/2}t^{1/2+pA_2}(\omega \log n)^{3/2}).$$

As $H(\mathbf{w}_0) = 0$, it follows from the definition (7) of the function H, that with the desired probability

$$N_{0,t} = \frac{bt}{1 + pA_2} + O((bt)^{1/2}(\omega \log n)^{3/2}),$$

which finishes the proof of the theorem. $\qquad\square$

We may repeat (recursively) the argument as in the proof of Theorem 6.8 for $N_{i,t}$ with $i \geq 1$. Since the expected change for $N_{i,t}$ is slightly different now (see (6)), we obtain our result by considering the following function:

$$H(x, y) = x^{p(A_1 i + A_2)}y - c_{i-1}\frac{p(A_1(i-1) + A_2)}{1 + p(A_1 i + A_2)}x^{1+p(A_1 i + A_2)}.$$

Moreover, in bounding $\sum \beta_s$ and $\sum \gamma_s^2$ (see (10) and (11)) we need b to be of order at least $(A_1 i + A_2)/t$; say, $bt \gg i$. Other than these minor adjustments, the argument is similar as in the case $i = 0$, and we get the following result. Note that the conclusion (the last claim) follows as

$$c_i bt = \Theta(i^{-1-1/(pA_1)}bt) = \Theta(i(bt)^{1/2}i^{-2-1/(pA_1)}(bt)^{1/2}) \gg i(bt)^{1/2}(\log n)^{3/2},$$

provided $bt \gg i^{4+2/(pA_1)}\log^3 n$.

Theorem 6.9. *Let B be a ball of volume $b = b(n)$, $t = t(n) \in \mathbb{N}$, and $i_f = i_f(n) \in \mathbb{N}$ be any functions of n such that $bt \gg i_f$. Let $\omega = \omega(n)$ be any function tending to infinity together with n. The following holds with probability $1 - o(n^{-2})$. For any $0 \le i \le i_f$,*

$$N_{i,t} = N_{i,t}(B) = c_i bt + O(i(bt)^{1/2}(\omega \log n)^{3/2}).$$

In particular, if $bt \gg i^{4+2/(pA_1)} \log^3 n$, then $N_{i,t} \sim c_i bt$.

Finally, we can move to the proof of Theorem 4.2.

Proof of Theorem 4.2. Let us fix any vertex v for which

$$\deg^-(v, n) = k = k(n) \ge (\omega \log n)^{4 + (4pA_1 + 2)/(pA_1(1 - pA_1))}.$$

Based on Theorem 4.1, we may assume that for all values of t such that

$$n \left(\frac{\omega \log n}{k} \right)^{\frac{1}{pA_1}} =: T_v \le t \le n,$$

we have

$$\deg^-(v, t) \sim k \left(\frac{t}{n} \right)^{pA_1}.$$

For any $\ell \in \mathbb{N} \cup \{0\}$, let

$$t_\ell = 2^\ell T_v, \qquad b_\ell = A_1 k t_\ell^{pA_1 - 1} n^{-pA_1},$$

B_ℓ be the ball around v of volume b_ℓ, and L be the smallest integer ℓ such that $t_\ell \ge n$. In fact, we will assume that $t_L = n$, as we may adjust ω (that is, multiply by a constant in $(1/2, 1)$), if needed. Let $t_v := n(\omega \log n)^{-1/(pA_1)}$; since $k \ge (\omega \log n)^2$, we have $T_v \le t_v \le n$. Let L' be the smallest integer ℓ such that $t_\ell \ge t_v$.

Times $t_0 = T_v$, $t_{L'} = \Theta(t_v)$, and $t_L = n$ are important stages of the process; vertex v has, respectively, degree $(1 + o(1))\omega \log n$, $\Theta(k/(\omega \log n))$, and k. Note that at time t_ℓ (for any $0 \le \ell \le L$) the sphere of influence of v has volume $(1 + o(1))b_\ell$. Moreover, based on Corollary 6.4 (applied with, say, $\sqrt{\omega}$ instead of ω) we may assume that any vertex v_i born after time T_v satisfies (for any $T_v \le t \le n$)

$$\deg(v_i, t) \le \sqrt{\omega} \log n \left(\frac{t}{i} \right)^{pA_1} = o \left(\omega \log n \left(\frac{t}{i} \right)^{pA_1} \right) = o(\deg(v, t)); \tag{12}$$

as a result, the sphere of influence of w has negligible volume comparing to the one of v.

We will independently prove an upper bound and a lower bound of $c_{new}(v, n)$. In order to do it, we need to estimate $|E_{new}(N^-(v, n))|$, the number of directed edges from u to w, where both u and w are neighbours of v born after time T_v.

Proof of $c_{new}(v, n) = O(1/k)$: Suppose that a neighbour w of v lies in $B_{\ell-1} \setminus B_\ell$ for some ℓ. An easy but an important observation is that at any time $t \ge t_{\ell+1}$, the sphere of influence of v is completely

disjoint from the one of w. Hence, the number of edges to w that contribute to $c_{new}(v, n)$ can be upper bounded by $\deg^-(w, t_{\ell+1})$. It follows that

$$
\begin{aligned}
|E_{new}(N^-(v, n))| &\leq \sum_{\ell=0}^{L-2} \sum_{w \in B_{\ell-1} \backslash B_\ell} \deg^-(w, t_{\ell+1}) + \sum_{w \in B_{L-2}} \deg^-(w, t_L) \\
&\leq \sum_{\ell=0}^{L-1} \sum_{w \in B_{\ell-1}} \deg^-(w, t_{\ell+1}).
\end{aligned}
$$

Let $i_f = (\omega \log n)^{1/(1-pA_1)}$. We will independently deal with the largest balls, namely B_ℓ for $\ell < L'$; for the remaining ones, we will deal with vertices of degree more than i_f before analyzing the contribution from low degree ones. In other words, we are going to show that each of the following three functions is of order at most k:

$$
\begin{aligned}
\alpha &= \sum_{\ell=0}^{L'-1} \sum_{w \in B_{\ell-1}} \deg^-(w, t_{\ell+1}), \\
\beta &= \sum_{\ell=L'}^{L-1} \sum_{\substack{w \in B_{\ell-1} \\ \deg^-(w, t_{\ell+1}) \leq i_f}} \deg^-(w, t_{\ell+1}), \\
\gamma &= \sum_{\ell=L'}^{L-1} \sum_{\substack{w \in B_{\ell-1} \\ \deg^-(w, t_{\ell+1}) > i_f}} \deg^-(w, t_{\ell+1}).
\end{aligned}
$$

The conclusion will follow immediately as $|E_{new}(N^-(v, n))| \leq \alpha + \beta + \gamma$.

In order to bound α, we only need to use (12) to get that

$$
\begin{aligned}
\mathbb{E}(\alpha) &= \sum_{\ell=0}^{L'-1} b_{\ell-1} \sum_{i=1}^{t_{\ell+1}} \deg^-(v_i, t_{\ell+1}) \leq \sum_{\ell=0}^{L'-1} b_{\ell-1} \sum_{i=1}^{t_{\ell+1}} \sqrt{\omega} \log n \left(\frac{t_{\ell+1}}{i}\right)^{pA_1} \\
&= \sum_{\ell=0}^{L'-1} \sqrt{\omega} \log n \, b_{\ell-1} \, t_{\ell+1}^{pA_1} \sum_{i=1}^{t_{\ell+1}} i^{-pA_1} = \sum_{\ell=0}^{L'-1} \Theta\left(\sqrt{\omega} \log n \, b_{\ell-1} \, t_{\ell+1}\right) \\
&= \sum_{\ell=0}^{L'-1} \Theta\left(\sqrt{\omega} \log n \, k \left(\frac{t_{\ell+1}}{n}\right)^{pA_1}\right) \\
&= \Theta\left(\sqrt{\omega} \log n \, k \left(\frac{t_{L'}}{n}\right)^{pA_1}\right) \sum_{\ell=0}^{L'-1} 2^{-\ell pA_1} \\
&= \Theta\left(\sqrt{\omega} \log n \, k \left(\frac{t_{L'}}{n}\right)^{pA_1}\right) = o(k).
\end{aligned}
$$

The fact that, with the desired probability, $\alpha = O(k)$ follows from a standard martingale argument (for example, one could use Lemma 6.7).

Similarly, we can deal with γ. It follows from (12) that no vertex born after time

$$
\left(\frac{\sqrt{\omega} \log n}{i_f}\right)^{1/(pA_1)} t_{\ell+1} \leq (\sqrt{\omega} \log n)^{(1-1/(1-pA_1))/(pA_1)} \, t_{\ell+1} = (\sqrt{\omega} \log n)^{-1/(1-pA_1)} \, t_{\ell+1}
$$

can satisfy $\deg^-(w, t_{\ell+1}) > i_f$. Hence,

$$
\begin{aligned}
\mathbb{E}(\gamma) &= \sum_{\ell=L'}^{L-1} b_{\ell-1} \sum_{i=1}^{(\sqrt{\omega}\log n)^{-1/(1-pA_1)}t_{\ell+1}} \deg^-(v_i, t_{\ell+1}) \leq \sum_{\ell=L'}^{L-1} \Theta\left(b_{\ell-1} \, t_{\ell+1}\right) \\
&= \sum_{\ell=L'}^{L-1} \Theta\left(k \left(\frac{t_{\ell+1}}{n}\right)^{pA_1}\right) = \Theta(k) \sum_{\ell=0}^{L'-1} 2^{-\ell p A_1} = O(k).
\end{aligned}
$$

Finally, we need to deal with β. This time, we need to use Theorem 6.9 to count (independently) the number of vertices in $B_{\ell-1}$ of a certain degree. We may apply this theorem as for any $L' \leq \ell \leq L-1$, we have

$$
\begin{aligned}
b_{\ell-1} t_{\ell+1} &\geq b_{L'-1} t_{L'+1} = \Theta\left(k \left(\frac{t_v}{n}\right)^{pA_1}\right) = \Theta\left(\frac{k}{\omega \log n}\right) \\
&= \Omega\left((\omega \log n)^{3+(4pA_1+2)/(pA_1(1-pA_1))}\right) \\
&= \Omega\left((\omega \log n)^3 \, i_f^{(4pA_1+2)/(pA_1)}\right) \\
&\gg i_f^{4+2/(pA_1)} \, \log^3 n,
\end{aligned}
$$

since $k \geq (\omega \log n)^{4+(4pA_1+2)/(pA_1(1-pA_1))}$. (In fact, this is the main bottleneck that forces us to assume that k is large enough.) We get the following:

$$
\begin{aligned}
\beta &= \sum_{\ell=L'}^{L-1} \sum_{i=1}^{i_f} i N_{i,t_{\ell+1}}(B_{\ell-1}) = (1+o(1)) \sum_{\ell=L'}^{L-1} \sum_{i=1}^{i_f} i c_i b_{\ell-1} t_{\ell+1} \\
&= \Theta\left(\sum_{\ell=L'}^{L-1} b_{\ell-1} t_{\ell+1} \sum_{i=1}^{i_f} i^{-1/(pA_1)}\right) = \Theta\left(\sum_{\ell=L'}^{L-1} b_{\ell-1} t_{\ell+1}\right) = O(k),
\end{aligned}
$$

as argued before.

Proof of $c_{new}(v,n) = \Omega(1/k)$: The lower bound is straightforward. Clearly, B_{L+1} is contained in the sphere of influence of vertex v not only at time n but, in fact, at any point of the process. It follows from Theorem 6.9 that the number of vertices of in-degree 1 that lie in B_{L+1} is $\Theta(b_{L+1}n) = \Theta(k)$. Moreover, their in-neighbours are also contained in the sphere of influence of v and, with the desired probability, say, half of them are born after time T_v. In order to avoid complications with events not being independent, we can select a family of $\Theta(k)$ directed edges such that no endpoint belongs to more than one edge. Now, each of these selected edges have both endpoints in the in-neighbourhood of v with probability p^2, independently on the other edges. Hence, the expected number of edges in $|E_{new}(N^-(v,n))|$ is $\Omega(1/k)$ and the conclusion follows easily from the Chernoff bound. $\square$

6.3 Proof of Theorem 4.3

We move immediately to the proof.

Proof of Theorem 4.3. Let $1 \leq \alpha = \alpha(n) = n^{o(1)}$ and $2 \leq \beta = \beta(n) = O(\log n)$ be any functions of n. We will tune these functions at the end of the proof for a specific value of k, depending on the case

(i), (ii), or (iii) we deal with. Pick any point s in S and consider two balls, B_1 and B_2, centered at s; the first one of volume C_1 and the second one of volume C_2, where

$$C_1 = \frac{A_2}{10n}, \quad \text{and} \quad C_2 = \frac{2(A_1 + A_2)\beta}{n/(2\alpha)}.$$

Let v be the first vertex that lands in B_1. We independently consider three phases.

Phase 1: Up to time $T_1 = n/\alpha$ when $\deg^-(v, T_1) = \Theta(\beta)$.

Consider the time interval between $n/(2\alpha)$ and n/α. We are interested in the following event D: during the time interval under consideration, β vertices land in B_1 but no vertex lands in $B_2 \setminus B_1$. Clearly,

$$\mathbb{P}(D) = \binom{n/(2\alpha)}{\beta} C_1^\beta (1 - C_2)^{n/(2\alpha) - \beta} \geq \left(\frac{nC_1}{3\alpha\beta}\right)^\beta \exp\left(-\frac{C_2 n}{\alpha}\right).$$

Straightforward but important observations are that every vertex in B_1 is inside a ball around any other vertex in B_1 (balls have volumes at least $A_2/(n/\alpha) \geq A_2/n$, deterministically); moreover, conditioning on D, during the whole time interval all balls around β vertices in B_1 are contained in B_2 (balls have volumes at most $(A_1\beta + A_2)/(n/(2\alpha))$).

We condition on event D and consider two scenarios that will be applied for two different ranges of k.

Event F_1: vertices in B_1 form a (directed) complete graph on β vertices; in particular, $\deg^-(v, n/\alpha) = \beta - 1$ and $c^-(v, n/\alpha) = 1$. It follows that

$$\mathbb{P}(F_1 | D) = p^{\binom{\beta}{2}},$$

and so

$$
\begin{aligned}
\mathbb{P}(D \wedge F_1) & \geq \left(\frac{nC_1}{3\alpha\beta}\right)^\beta \exp\left(-\frac{C_2 n}{\alpha}\right) p^{\binom{\beta}{2}} \\
& = \exp\left(-\beta \log\left(30\alpha\beta/A_2\right) - 4(A_1 + A_2)\beta - \binom{\beta}{2}\log\left(1/p\right)\right) \\
& \geq \exp\left(-\beta \log \alpha - 2\beta \log\log n - \beta^2 \log\left(1/p\right)\right) \\
& \geq n^{-1/5 - o(1) - 1/5} \geq n^{-1/2},
\end{aligned}
$$

provided that

$$\max\left\{\beta \log \alpha, \beta^2 \log\left(1/p\right)\right\} \leq \frac{1}{5}\log n. \tag{13}$$

Event F_2: the first $\beta p/8 - 1$ vertices that landed in B_1 right after v connected to v but the remaining $\beta(1 - p/8)$ vertices did not do this; moreover, each of $\beta p/8 - 1$ neighbours of v got connected to at least $\beta p/4$ other vertices. In particular, $\deg^-(v, n/\alpha) = \beta p/8 - 1$ and all neighbours w of v satisfy $\deg^-(w, n/\alpha) \geq \beta p/4$. It follows that

$$
\begin{aligned}
\mathbb{P}(F_2 | D) & = p^{\beta p/8 - 1}(1 - p)^{\beta(1 - p/8)} \prod_{i=1}^{\beta p/8} \mathbb{P}\left(\text{Bin}(\beta - i, p) \geq \beta p/4\right) \\
& \geq [p(1 - p)]^\beta \, \mathbb{P}\left(\text{Bin}(\beta(1 - p/8), p) \geq \beta p/4\right)^{\beta p/8} \\
& \geq \left[\frac{p(1 - p)}{2}\right]^\beta,
\end{aligned}
$$

23

since $\mathbb{E}(\text{Bin}(\beta(1-p/8), p)) = \beta(1-p/8)p \geq \beta p/2$. This time we get

$$
\begin{aligned}
\mathbb{P}(D \wedge F_2) &\geq \left(\frac{nC_1}{3\alpha\beta}\right)^\beta \exp\left(-\frac{C_2 n}{\alpha}\right)\left[\frac{p(1-p)}{2}\right]^\beta \\
&= \exp\left(-\beta \log\left(30\alpha\beta/A_2\right) - 4(A_1 + A_2)\beta - \beta \log\left(\frac{2}{p(1-p)}\right)\right) \\
&\geq \exp\left(-\beta \log \alpha - \beta \log \beta - O(\beta)\right) \\
&\geq n^{-1/5-1/5-o(1)} \geq n^{-1/2},
\end{aligned}
$$

provided that

$$
\max\left\{\beta \log \alpha, \beta \log \beta\right\} \leq \frac{1}{5} \log n. \tag{14}
$$

Phase 2: Between time $T_1 = n/\alpha$ and time T_2 when $\deg^-(v, T_2) \geq \omega \log n$ for some $\omega = \omega(n) \leq \log\log n$ tending to infinity as $n \to \infty$.

We assume that events D and F_2 hold. Let W be the set of the first $\beta p/8 - 1$ neighbours of v considered in the previous phase. Using the same argument as in Lemma 6.2, we are going to show that with probability at least $1/2$ for any t in the time interval under consideration and any vertex $w \in W \cup \{v\}$,

$$
\deg^-(w, t) \sim \deg^-(w, n/\alpha)\left(\frac{t}{n/\alpha}\right)^{pA_1}.
$$

Let $\varepsilon = 1/(\omega \log\log n)$ and suppose that

$$
\deg^-(v, T) = d \geq \beta p/8 - 1.
$$

Then, with 'failing' probability $\exp(-\Omega(\varepsilon^2 d))$, for some value of t, $T \leq t \leq 2T$,

$$
\left|\deg^-(v, t) - d \cdot \left(\frac{t}{T}\right)^{pA_1}\right| > \frac{5}{pA_1} \cdot \frac{t}{T} \varepsilon.
$$

We will apply this bound for $T = 2^i n/\alpha$ for $0 \leq i = O(\log\log n)$. Hence, the probability that we fail for some vertex (at some time t between T_1 and T_2) is at most

$$
\frac{\beta p}{8} \, O(\log\log n) \, \exp(-\Omega(\varepsilon^2 d)) = O(\beta \log\log n) \exp\left(-\Omega\left(\frac{\beta}{(\omega \log\log n)^2}\right)\right) \leq \frac{1}{2},
$$

provided that

$$
\beta \geq \omega^3 \, (\log\log n)^2 (\log\log\log n). \tag{15}
$$

The claim holds as the cumulative error term is

$$
(1 + O(\varepsilon))^{O(\log\log n)} = 1 + O(\varepsilon \log\log n) \sim 1.
$$

Phase 3: Between time T_2 and time n.

We assume that events D and F_2 hold, and Phase 2 finished successfully (that is, concentration holds for all vertices in W). It follows immediately from Corollary 6.3 that with probability $1 - o(n^{-1}\beta)$ for any t in the time interval between T_2 and n, and any vertex $w \in W \cup \{v\}$,

$$\deg^-(w, t) \sim \deg^-(w, n/\alpha) \left(\frac{t}{n/\alpha} \right)^{pA_1}.$$

The conclusion is that with probability at least $n^{-1/2}/3$, for a given point s in S, there exists vertex v in B_1 that has $\Theta(\beta)$ in-neighbours in B_1. Moreover, between time n/α and n, the degree of these neighbours of v are larger by a factor of at least $2 + o(1)$ than the degree of v. It follows that in this time interval, the ball around v is contained in all the balls of early neighbours of v. Conditioning on this event and assuming that, say, $\alpha \geq 2$, with probability at least $1 - \beta \exp(-\Omega(\omega \log n)) \geq 1 - n^{-1}$, each early neighbour has a positive fraction of neighbours of v as its neighbours at time n. (Note that this time events are not independent but the failing probability is small enough for the union bound to be applied.) It follows that with probability at least $n^{-1/2}/4$, we have $c^-(v, n) = \Omega(\beta/k)$.

Finally, tessellate S into $n^{1-o(1)}$ squares of volumes, say,

$$C_3 = \frac{\omega^2 \log n}{n/(2\alpha)} = n^{-1+o(1)},$$

as it is assumed that $\alpha = n^{o(1)}$, and take various s to be the centers of the corresponding squares. Note that conditioning of all the phases to end up with success, balls of all vertices under consideration are contained in the square. Moreover, in order to decide if a given square is successful does not require to expose vertices outside of this square. Hence, the events associated with different squares are almost independent. Formally, one would need to use (in a straightforward way) the second moment method to show this claim. It follows that a.a.s. there is at least one square that is successful.

Now, we are ready to tune α and β for a specific function k. For case (i), we take $\alpha = 1$ (that is, no phase 2 and 3) and $\beta = k$. It is straightforward to see that conditions (13) are satisfied. For case (ii), we take

$$\beta = \frac{k}{5} \leq \frac{\log n}{5 \log \log n} \quad \text{and} \quad \alpha = \left(\frac{k}{\beta} \right)^{1/(pA_1)} = 5^{1/(pA_1)} \geq 5.$$

(This time, there is no phase 3.) Again, it is straightforward to see that conditions (14) and (15) are satisfied. For case (iii), we take

$$\beta = \frac{pA_1}{5} \omega (\log \log n)^2 (\log \log \log n) \quad \text{and} \quad \alpha = \left(\frac{k}{\beta} \right)^{1/(pA_1)} \leq k^{1/(pA_1)} \leq n^{\xi/(pA_1)}.$$

(Clearly, $\alpha \gg 1$.) As usual, it is straightforward to see that conditions (14) and (15) are satisfied, and the proof is finished. $\qquad\square$

6.4 Proof of Theorem 4.5

We move immediately to the proof.

Proof of Theorem 4.5. Let $\omega = \omega(n) = \log^{o(1)} n$ be any function tending to infinity as $n \to \infty$ (arbitrarily slowly). First, note that a.a.s.

$$|X_k| = \sum_{i=(1-\delta)k}^{(1+\delta)k} \Theta(i^{-1-1/(pA_1)}n) = \Theta_\delta(nk^{-1/(pA_1)}),$$

as the degree distribution of G_n follows power law with exponent $1 + 1/(pA_1)$ (Aiello et al. 2009). Let

$$rT = T(n) := n\left(\frac{2\omega \log n}{k}\right)^{1/(pA_1)}.$$

Note that $T \geq n^{\varepsilon/(pA_1)}$, as $k \leq n^{pA_1-\varepsilon}$. It follows from Theorem 4.1 that a.a.s., for each $v \in X_k$,

$$(1+o(1))(1-\delta)(2\omega \log n) \leq \deg^-(v,T) \leq (1+o(1))(1+\delta)(2\omega \log n).$$

In particular, for n large enough,

$$\deg^-(v,T) > \omega \log n \tag{16}$$

(as $\delta < 1/2$) and so all old neighbours of v are born before time T.

We start from part (i). As we aim for the statement that holds for almost all vertices in X_k, we may concentrate on any vertex $v \in X_k$ that is born after time $nk^{-1/(pA_1)}/\omega = o(nk^{-1/(pA_1)})$ and simply ignore the remaining ones (as the number of them is negligible comparing to $|X_k|$). Since each in-neighbour v_u of v is also born after time $nk^{-1/(pA_1)}/\omega$, we can use Corollary 6.4 to be able to assume that for any $u \leq t \leq n$,

$$\deg(v_u, t) \leq \omega \log n \left(\frac{t}{u}\right)^{pA_1} \leq \omega \log n \left(\frac{t}{nk^{-1/(pA_1)}/\omega}\right)^{pA_1}.$$

As a result, for any $T \leq t \leq n$,

$$\begin{aligned}
\frac{|S(v_u, t)|}{|S(v, t)|} &\leq (1+o(1))\frac{\omega \log n \left(\frac{t}{nk^{-1/(pA_1)}/\omega}\right)^{pA_1}}{\deg^-(v,T)(t/T)^{pA_1}} \leq (1+o(1))\left(\frac{T}{nk^{-1/(pA_1)}/\omega}\right)^{pA_1} \\
&\sim 2\omega^{pA_1+1} \log n \leq \omega^2 \log n.
\end{aligned}$$

(Here we used Theorem 4.1 and (16).) Moreover, we may ignore all vertices that have too many vertices that are too close to them at time T. Formally, we ignore all vertices v that have at least $C = \lceil 8/(\varepsilon pA_1) \rceil > e$ vertices in the ball of volume $B = 1/(T \log^{\varepsilon/2} n)$ around v at time T. Indeed, suppose that T points are placed independently and uniformly at random in S (without generating the graph). The probability that a given point v has too many points around is at most

$$\binom{T}{C}B^C \leq \left(\frac{eTB}{C}\right)^C \leq (TB)^C = \log^{-\varepsilon C/2} n = \log^{-4/(pA_1)} n.$$

Since the expected number of such points is at most

$$\begin{aligned}
T \log^{-4/(pA_1)} n &= nk^{-1/(pA_1)}(2\omega \log n)^{1/(pA_1)} \log^{-4/(pA_1)} n \\
&\leq nk^{-1/(pA_1)} \log^{-2/(pA_1)} n,
\end{aligned}$$

26

it follows from Markov's inequality that a.a.s. there are at most $nk^{-1/(pA_1)}\log^{-1/(pA_1)}n = o(|X_k|)$ of them, as claimed. (In fact, $nk^{-1/(pA_1)}\log^{-1/(pA_1)}n = O(|X_k|/(\omega^2\log n))$, which will be needed for part (ii).)

Our goal is to show that $c_{old}(v,n) = O(1/k)$. Since there are at most $C = O(1)$ close in-neighbours of v, their contribution to $c_{old}(v,n)$ is only $O(1/k)$ and so we need to concentrate on far in-neighbours of v. Let

$$\hat{T} := T\log^{(2+\varepsilon)/(1-pA_1)}n,$$

and note that

$$\hat{T} = n\left(\frac{2\omega\log n}{k}\right)^{1/(pA_1)}\log^{(2+\varepsilon)/(1-pA_1)}n = o(n),$$

assuming that $k \geq \omega^2\log^{1+(2+\varepsilon)pA_1/(1-pA_1)}n$, which we may by taking ω small enough. Let u be any (far) in-neighbour of v that is outside of the ball of volume B around v at time T. Note that

$$
\begin{aligned}
|S(u,\hat{T})| &\leq (\omega^2\log n)|S(v,\hat{T})|\\
&\sim (\omega^2\log n)(A_1\deg^-(v,\hat{T}))/\hat{T}\\
&\leq (\omega^2\log n)(4A_1\omega\log n)(\hat{T}/T)^{pA_1}/\hat{T}\\
&\leq (\omega^4\log^2 n)\hat{T}^{pA_1-1}T^{-pA_1}\\
&= (\omega^4\log^2 n)(\log^{-(2+\varepsilon)}n)/T\\
&= 1/(T\log^{-\varepsilon+o(1)}n) = o(B)
\end{aligned}
$$

and so also $|S(v,\hat{T})| = o(B)$, which implies that at time $\hat{T}$ spheres of influence of u and v are disjoint and will continue to shrink. As a result, the number of common neighbours of v and u is at most

$$\deg^-(v,\hat{T}) = k(\hat{T}/n)^{pA_1} = o(k),$$

and so the number of common neighbours of v and its far neighbours is negligible. Part (i) holds.

The proof of part (ii) is almost the same so we only point out small adjustments that need to be implemented. It follows from Corollary 4.4 that we may assume that $c^-(v,n) = O(\omega\log n/k)$ for any vertex $v \in X_k$. Hence, it is enough to show that all but at most $O(|X_k|/(\omega^2\log n)) = O(nk^{-1/(pA_1)}/(\omega^2\log n))$ vertices in X_k have $c^-(v,n) = O(1/k)$. This time we can only ignore vertices born before time $nk^{-1/(pA_1)}/(\omega^2\log n)$ which gives slightly weaker bound for the ratio of the volumes of influence of a neighbour of v and v itself:

$$\frac{|S(u,t)|}{|S(v,t)|} \leq \omega^3\log^{1+pA_1}n.$$

As a result, we need to define $\bar{T}$ as a counterpart of $\hat{T}$ as follows:

$$\bar{T} := T\log^{(2+pA_1+\varepsilon)/(1-pA_1)}n,$$

and note that $\bar{T} = o(n)$, assuming the stronger lower bound for k. The rest of the proof is not affected. $\qquad\square$

7 Conclusion and future directions

In this paper, we analyzed the clustering properties of the SPA model. Namely, we proved that the average local clustering coefficient $C(d)$ asymptotically behaves as d^{-1}. Moreover, we were able to prove that not only the average but the individual local clustering coefficient of a vertex of degree d behaves as $1/d$ if d is large enough.

The behaviour $C(d) \propto d^{-1}$ is often observed in real-world networks. However, in some cases $d^{-\varphi}$ with $\varphi \neq 1$ is claimed. In random graph models of complex networks, it is usually the case that $C(d) \propto d^{-1}$. For example, such decay is proved in Krot and Ostroumova Prokhorenkova (2015) for a variety of networks based on preferential attachment.

As a future work, it would be interesting to obtain a precise constant C in the expression $C(d) \sim Cd^{-1}$. We believe that the value C is defined by the parameters p and A_1 of the model, similarly to what was previously observed for the general class of preferential attachment models (Krot and Ostroumova Prokhorenkova 2015). This would provide an insight into percolation properties of the network. Namely, it was shown in Serrano and Boguná (2006b) that percolation properties of a network are defined by the type (weak or strong) of its connectivity and the type of connectivity is defined by the constant C (whether it is greater or smaller than one). Another interesting direction for future research is to analyze the nature of usually obtained value $\varphi = 1$ and try to modify the SPA model in order to make it more flexible and allowing to generate graphs with $\varphi \neq 1$. Similar results with $\varphi \neq 1$ were recently obtained for affiliation networks (Bloznelis and Petuchovas 2017).

Acknowledgements

References

Aiello, W., Bonato, A., Cooper, C., Janssen, J., and Prałat, P. (2009). A spatial web graph model with local influence regions. *Internet Mathematics*, 5:175–196.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Bloznelis, M. and Petuchovas, J. (2017). Correlation between clustering and degree in affiliation networks. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 90–104. Springer.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308.

Bollobás, B. and Riordan, O. M. (2003). Mathematical results on scale-free random graphs. *Handbook of graphs and networks: from the genome to the internet*, pages 1–34.

Chung, F. and Lu, L. (2006). *Complex Graphs and Networks*. American Mathematical Society, Providence, Rhode Island, USA.

Cooper, C., Frieze, A., and Prałat, P. (2014). Some typical properties of the spatial preferred attachment model. *Internet Mathematics*, 10:27–47.

Costa, L. d. F., Rodrigues, F. A., Travieso, G., and Villas Boas, P. R. (2007). Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1):167–242.

Csányi, G. and Szendrői, B. (2004). Structure of a large social network. *Physical Review E*, 69(3):036131.

Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F. (2002). Pseudofractal scale-free web. *Physical review E*, 65(6):066122.

Faloutsos, M., Faloutsos, P., and Faloutsos, C. (1999). On power-law relationships of the internet topology. In *ACM SIGCOMM computer communication review*, volume 29, pages 251–262. ACM.

Iskhakov, L., Kamiński, B., Mironov, M., Prałat, P., and Prokhorenkova, L. (2018). Clustering properties of spatial preferential attachment model. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 30–43. Springer.

Jacob, E. and Mörters, P. (2013). A spatial preferential attachment model with local clustering. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 14–25. Springer.

Jacob, E., Mörters, P., et al. (2015). Spatial preferential attachment networks: Power laws and clustering coefficients. *The Annals of Applied Probability*, 25(2):632–662.

Janson, S., Łuczak, T., and Ruciński, A. (2000). *Random Graphs*. Wiley, New York, USA.

Janssen, J., Hurshman, M., and Kalyaniwalla, N. (2013a). Model selection for social networks using graphlets. *Internet Mathematics*, 8(4):338–363.

Janssen, J., Prałat, P., and Wilson, R. (2013b). Geometric graph properties of the spatial preferred attachment model. *Advances in Applied Mathematics*, 50:243–267.

Janssen, J., Prałat, P., and Wilson, R. (2016). Non-uniform distribution of nodes in the spatial preferential attachment model. *Internet Mathematics*, 12(1–2):121–144.

Kamiński, B., Olczak, T., and Prałat, P. (2017). Endogenous differentiation of consumer preferences under quality uncertainty in a spa network. In *Algorithms and Models for the Web Graph: 14th International Workshop, WAW 2017, Toronto, ON, Canada, June 15–16, 2017, Revised Selected Papers*, pages 42–59. Springer.

Krot, A. and Ostroumova Prokhorenkova, L. (2015). Local clustering coefficient in generalized preferential attachment models. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 15–28. Springer.

Leskovec, J. (2008). Dynamics of large networks (proquest). *Ann Arbor*.

Morgan, H., Kamiński, B., and Prałat, P. (2018). Cognition, network structure, and learning in top managers' interpersonal networks. *preprint*.

Newman, M. E. (2003a). Properties of highly clustered networks. *Physical Review E*, 68(2):026121.

Newman, M. E. (2003b). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

Newman, M. E. (2005). Power laws, pareto distributions and zipf's law. *Contemporary physics*, 46(5):323–351.

Ostroumova Prokhorenkova, L. (2016). Global clustering coefficient in scale-free weighted and unweighted networks. *Internet Mathematics*, 12(1-2):54–67.

Ostroumova Prokhorenkova, L., Prałat, P., and Raigorodskii, A. (2017). Modularity of complex networks models. *Internet Mathematics, 23pp*.

Pittel, B., Spencer, J., and Wormald, N. (1996). Sudden emergence of a giant k-core in a random graph. *Journal of Combinatorial Theory, Series B*, 67:111–151.

Ravasz, E. and Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, 67(2):026112.

Serrano, M. Á. and Boguná, M. (2006a). Clustering in complex networks. i. general formalism. *Physical Review E*, 74(5):056114.

Serrano, M. Á. and Boguná, M. (2006b). Clustering in complex networks. ii. percolation properties. *Physical Review E*, 74(5):056115.

Vázquez, A., Pastor-Satorras, R., and Vespignani, A. (2002). Large-scale topological and dynamical properties of the internet. *Physical Review E*, 65(6):066130.

Wormald, N. (1999). The differential equation method for random graph processes and greedy algorithms. *Lectures on Approximation and Randomized Algorithms*, pages 73–155.