Clustering Properties of Spatial Preferential Attachment Model

Lenar Iskhakov[1], Bogumił Kamiński[2], Maksim Mironov[1], Liudmila Ostroumova Prokhorenkova[1,3], and Paweł Prałat[4]

[1] Advanced Combinatorics and Network Applications Lab, Moscow Institute of Physics and Technology, Moscow, Russia
[2] Warsaw School of Economics, Warsaw, Poland
[3] Machine intelligence and research department, Yandex, Moscow, Russia
[4] Department of Mathematics, Ryerson University, Toronto, Canada

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. For example, the vertex degree distribution was shown to follow a power law. In the current paper, we study the behaviour of $C(d)$, which is the average local clustering coefficient for the vertices of degree d. This characteristic was not previously analyzed in the SPA model. However, 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 are illustrated by numerous experiments with simulated graphs.

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 [9,24]. Therefore, numerous random graph models have been proposed to reflect and predict such quantitative and topological aspects of growing real-world networks [4,5].

The most well studied property of complex networks is their vertex degree distribution. For the majority of studied real-world networks, the degree distribution was shown to follow a heavy-tailed distribution [2,12,25]. Another important property of real-world networks is their clustering structure. 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 details). It is widely 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; some numerical values can be found in [24]; however, some contradicting theoretical results are presented in [26].

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. It was empirically shown that in real-world networks $C(d)$ usually decreases as $d^{-\psi}$ for some $\psi > 0$ [10,22,30,31]. In particular, for many studied networks, $C(d)$ scales as d^{-1} [29].

We study the clustering properties of the *Spatial Preferential Attachment* (SPA) model introduced in [1]. 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 [1] that the vertex degree distribution follows a power law. More details on the properties of the SPA model are given in Section 2.2. However, the clustering coefficient $C(d)$ was not previously analyzed for this model, although some clustering properties were analyzed for the generalized SPA model proposed in [14]. It is proved in [14] and [15] 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.

In this paper, we prove that the local clustering coefficient $C(d)$ decreases as $1/d$ in the SPA model. We also obtain some bounds for the individual local clustering coefficients of vertices. The obtained theoretical results are compared with and illustrated by numerous experiments on simulated graphs. Our theoretical results are asymptotic in nature, so we empirically investigate how the model behaves for finite size graphs and see that the asymptotic predictions are still close to empirical observations even for small graph sizes. Additionally, we demonstrate that some of our theoretical assumptions are probably too pessimistic and the SPA model behaves even more predictable than we have proven. We also propose an efficient algorithm for generating graphs according to the SPA model which runs much faster than the straightforward implementation.

Proofs of all theoretical results stated in this paper can be found in the journal version [13] that focuses exclusively on asymptotic results of the model. On the other hand, this proceeding version also contains results on simulated graphs and so can be viewed as a complement to the journal version.

2 Spatial Preferential Attachment model

2.1 Definition

This paper focuses on the *Spatial Preferential Attachment* (SPA) model, which was first introduced by [1]. 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 sequence of 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 $p\,|S(u, t-1)|$.

2.2 Properties 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 [1,8]. Specifically, Theorem 1.1 in [1] proves 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 $pA_2/(1 - pA_1)$, as shown in Theorem 1.3 in [1]. In [17], it was demonstrated that the SPA model give the best fit, in terms of graph structure, for a series of social networks derived from Facebook. In [18], some properties of common neighbours were used 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 [18], but [19] also investigated non-uniform distributions, which is clearly a more realistic setting. The SPA model was also used to study a duopoly market on which there is uncertainty of a product quality [20]. Finally, in [27] modularity of this model was investigated, which is a global criterion to define communities and a way to measure the presence of community structure in a network.

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., [5]). 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 [14,15] 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*, which was not previously analyzed for the SPA model. 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:

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

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 for various real-world networks and some random graph models. 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 $C(d) \sim d^{-\varphi}$ for some $\varphi > 0$. In particular, [29] shows that $C(d)$ can be well approximated by d^{-1} for four large networks, [31] obtains power-law in a real network with parameter 0.75, while [10] obtains $\varphi = 0.33$. The local clustering coefficient was also studied in several random graph models of complex networks. For instance, it was shown in [11,21,23] that some models have $C(d) \sim d^{-1}$. As we prove in this paper, similar behaviour is also 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. So, the directed clustering coefficient of a vertex v at time t and the average directed clustering for the vertices of incoming degree d are defined as

$$c^-(v,t) = \frac{|E(N^-(v,t))|}{\binom{\deg^-(v,t)}{2}}, \quad C^-(d,t) = \frac{\sum_{v:\deg^-(v,t)=d} c^-(v,t)}{|\{v : \deg^-(v,t) = d\}|}\,.$$

Note that we normalize $c^-(v,t)$ by $\binom{\deg^-(v,t)}{2}$, since in the SPA model edges can be created only from younger vertices to older ones.

4 Results

4.1 Notation

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 $f(n) = (1 + o(1))g(n)$).

First we 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 $c^-(v,n)$. 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, for a given function $\omega(n)$ that tends to infinity as $n \to \infty$, let $\hat{T}_v$ be the smallest integer t such that $\deg^-(v,t)$ exceeds

$\omega \log n$ (or $\hat{T}_v = n$ if $\deg^-(v,n) < \omega \log n$). 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. Finally,

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

and

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

where

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

4.2 Results

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 [18,19]; here, we present a slightly modified statement from [19], adjusted to our current needs.

Theorem 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 is chosen so that at this time vertex v has *a.a.s.* $(1+o(1))\omega \log n$ neighbours. 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).

Let us note that Theorem 1 immediately implies the following two corollaries.

Corollary 1 *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 $\deg^-(v,T) \geq \omega \log n$, for all times t, $T \leq t \leq n$,*

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

Corollary 2 *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 $i \geq 1$, and $i \leq t \leq n$ we have that $\deg^-(v_i, t) \leq \omega \log n \, (t/i)^{pA_1}$.*

Theorem 1 can be used to show that the contribution to $c^-(v,n)$ coming from edges to new neighbours of v is well concentrated.

Theorem 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 $c_{old}(v,n)$ 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 2 to $c^-(v,n)$.

Theorem 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 2 implies immediately the following corollary.

Corollary 3 *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

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

Moreover, despite the above 'negative' result, almost all vertices (of large enough degrees) have clustering coefficients of order $1/k$. Here is a precise statement. The conclusions in cases (i)' and (ii)' follow immediately from Theorem 2.

Theorem 4. *Let $\varepsilon, \delta \in (0, 1/2)$ be any two constants, and let $k = k(n) \le n^{pA_1 - \varepsilon}$ 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 briefly discuss the undirected case. The following lemma holds.

Lemma 1. *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,*

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

Note that a weaker bound of $\log^2 n$ was proved in [1]; with Corollary 2 in hand, we can get slightly better bound but the argument remains the same.

According to the above lemma, 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 k. Hence, all results discussed in this section also hold for the clustering coefficient $c(k, n)$ defined for the undirected graph $\hat{G}_n$ obtained from G_n by considering all edges as undirected.

5 Experiments

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.

5.1 Algorithm

Let us first discuss the complexity of the straightforward (*naive*) algorithm for generating graphs according to the SPA model. At each step we add one vertex and, for each existing vertex, we check if the new vertex belongs to its sphere of influence. Then we (possibly) add new edges and update the radii for all vertices. The complexity of this procedure is $\Theta(n^2)$.

Let us now propose a more efficient algorithm. First, we describe this algorithm and provide heuristic arguments about its complexity. Then, we compare running times of the new algorithm and the naive one.

Our algorithm works in several phases, as described further in the text. For now, let us assume that we already generated a graph on n vertices according to the SPA model and we want to add one additional vertex. It is known that

$$\mathbb{E}\Big(\deg^-(v_i,t)\Big) \sim \frac{A_2}{A_1}\left(\frac{t}{i}\right)^{pA_1},$$

provided that $i \gg 1$ (see, for example, [8]). We call a vertex *heavy* if its degree is at least D for some D; otherwise, it is *light*. All heavy vertices are kept in a separate list H. Fix

$$D = \frac{A_2}{A_1}\left(\frac{n}{T}\right)^{pA_1}, \tag{2}$$

so H has expected size around T. The choice of an optimal value of T will be discussed further in this section.

Let us divide $S = [0,1]^2$ into k squares where k is some perfect square; that is, each square will have side length $1/\sqrt{k}$. (We choose the dimension $m = 2$ for our simulations. However, the ideas can easily be applied for an arbitrary m.) All light vertices are kept in k disjoint lists; let $L(i)$ be a list containing all light vertices from square i. The expected number of vertices in each list is $(n-T)/k$.

We want the following property to be satisfied:

$$\sqrt{\frac{A_1D + A_2}{\pi n}} \le \frac{1}{\sqrt{k}}. \tag{3}$$

Indeed, if this is the case, then no light vertex v_i has the area of influence that touches squares other than the square containing v_i and the 8 adjacent squares. Moreover, the same property will hold for all $t > n$ as areas of influence of light vertices decrease with time. Hence, since we aim for an integer $\sqrt{k}$ to be as large as possible:

$$k = \left\lfloor \sqrt{\frac{\pi n}{A_1 D + A_2}} \right\rfloor^2 \Rightarrow k \approx \frac{\pi n}{A_2\left(1 + (n/T)^{pA_1}\right)}. \tag{4}$$

The most expensive computational work for the algorithm is the number of comparisons needed in order to add a vertex v_{n+1} to a graph, which is of order

$$f(T) := T + 9\,\frac{n-T}{k} = T + \frac{9A_2}{\pi}(1 - T/n)\left(1 + (n/T)^{pA_1}\right). \tag{5}$$

Hence, the function $f(T)$ is minimized for

$$T = \frac{9npA_1A_2(n/T)^{pA_1}}{\pi n - 9A_2 - 9A_2(1-pA_1)(n/T)^{pA_1}}\,.$$

For large n the second and the third terms in the denominator are negligible, as $pA_1 < 1$; moreover, if pA_1 is close to 1 we will soon show that $T = \Theta(n^{1/2-\epsilon})$ for some small $\epsilon > 0$, so the approximation converges fast. Thus, we may approximate T by:

$$T \approx n^{1-1/(pA_1+1)} \left(\frac{9pA_1A_2}{\pi}\right)^{1/(pA_1+1)}. \tag{6}$$

Using this T we can calculate the recommended value of D, see (3), and the density of the $\sqrt{k} \times \sqrt{k}$ grid, see (4).

Below are some practical implementation details:

- It is computationally expensive to recalculate H and L division each time a new vertex is added. By empirical testing, we have found that the recalculation should be done approximately after adding $t/4$ vertices, where t is the number of vertices in already constructed graph. As a result, the number of phases is $O(\log n)$, as each time the number of vertices increases by approximately 25%.
- As we work in phases, at each step we have to check if some light vertex becomes heavy, and move it to the appropriate list, if needed. However, this operation is not expensive computationally.
- After several phases, for actually constructed graphs the optimal parameters k, T and D might deviate from the theoretical values derived above. Therefore, in the implementation we choose the optimal parameters conditional on the actual input graph structure. Namely, for each candidate value k we can calculate the corresponding D using (3) and then calculate T from the data (this is the actual number of heavy vertices given D). We choose k to optimize the number of comparisons needed to add one vertex to the actual graph, the approximation for this value is given in (5). After that we dynamically construct H and L lists.

Let us now discuss the complexity of the obtained algorithm. Equation (5) shows that T is expected to be of order $n^{pA_1/(pA_1+1)}$. So, we may derive from (4) that k is of order $n^{1-pA_1+(pA_1)^2/(pA_1+1)} = n^{1/(pA_1+1)}$. From (5) we obtain that $f(T)$ grows as $n^{pA_1/(pA_1+1)}$. So, the expected complexity of the whole process is $\Theta\left(n^{2-1/(pA_1+1)}\right) \ll n^2$.

Figure 1 presents an empirical comparison of the running time for new and naive algorithms. We also present this figure in log-log scale. The computations were performed using Julia 0.6.2 language [3] and LightGraphs [6] package on a single thread of Intel i5-5200U @ 2.20GHz processor.

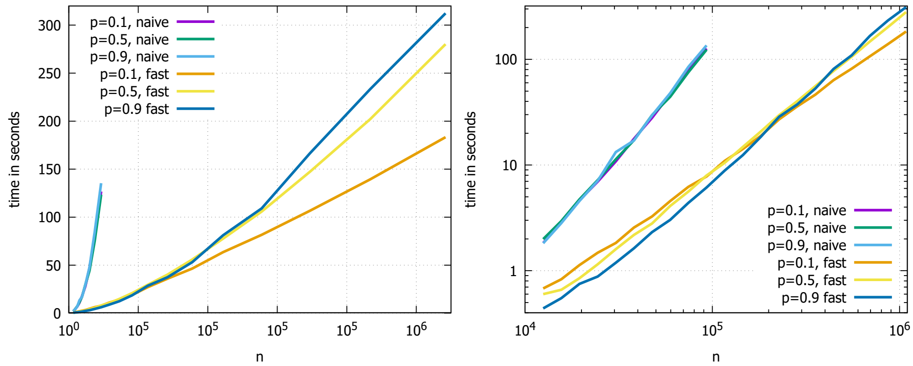

Fig. 1. Running time of the proposed and the naive algorithms.

Finally, let us mention that further improvements of the algorithm are possible. For example, one can keep more than two lists H and L. For example, $L_s(i)$ could contain vertices of degree between 2^{s-1} and 2^s that landed in region i, so the total number of lists is $O(\log n)$. Then, the running time of the algorithm would be $O(n \log n)$. Indeed, during a phase that started at time t, L_s has expected size $O(t\, 2^{-s/pA_1})$; since vertices from $L_s(i)$ are gathered from the square of area, say, $2^s/t$, the expected size of this list is $O(2^{s-s/(pA_1)}) = O(1)$. Hence, after adding one vertex, $O(\log n)$ lists are investigated and we expect only a constant number of comparisons done on each list. Of course, there is always a trade-off between the running time of an algorithm and how complicated it is to implement it. For our purpose, we decided to go for a simpler algorithm with only two lists.

5.2 Empirical analysis of the local clustering coefficient

In this section, we compare asymptotic theoretical results obtained in Section 4 with empirical results obtained for graphs with finite n. All graphs are generated according to the algorithm described in Section 5.1.

It is proven in Theorem 4 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 $p \in \{0.1, 0.2, \ldots, 0.9\}$, $A_1 = 1$, $A_2 = 10(1-p)/p$ (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.

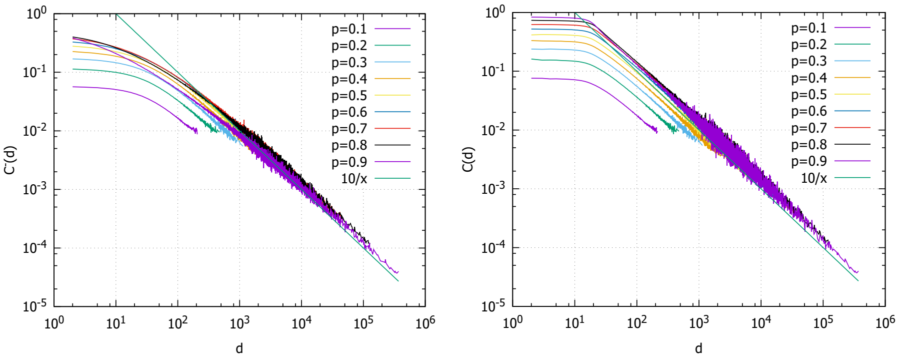

Fig. 2. Average local clustering coefficient for directed (left) and undirected (right) graphs.

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 that the statement of Theorem 4, 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.

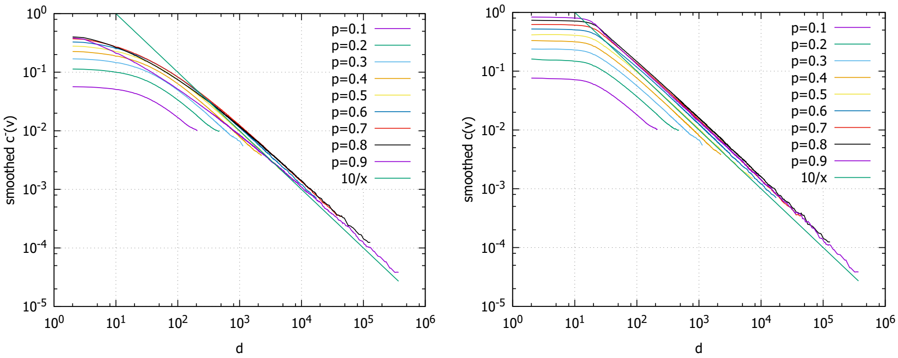

Fig. 3. Local clustering coefficient for directed (left) and undirected (right) graphs averaged over X_d.

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 4.2, where we analyzed new and old edges separately. In our experiments, we split $c^-(v,n)$ into "old" and "new" parts as in (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 $C^-_{old}(d)$.

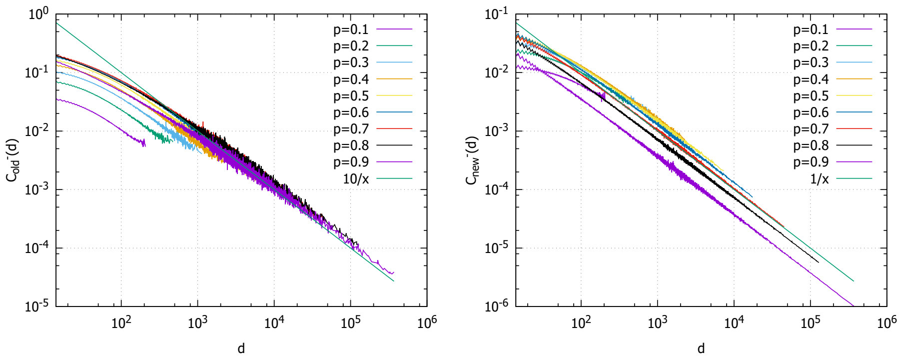

Fig. 4. Comparison of "new" and "old" parts of the average local clustering coefficient.

Finally, Figure 5 shows the distribution of the individual local clustering coefficients for one graph generated with $p = 0.7$. Theorem 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. 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 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 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.

Acknowledgements

This work is supported by Russian President grant MK-527.2017.1 and NSERC.

References

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

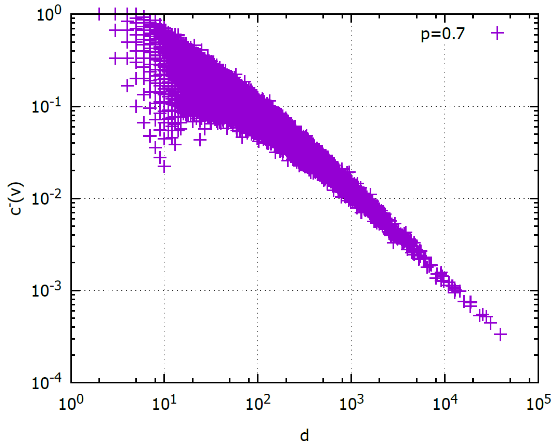

Fig. 5. The distribution of individual local clustering coefficients.

2. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. science 286(5439), 509–512 (1999)
3. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.: Julia: A fresh approach to numerical computing. SIAM Review 69, 65–98 (2017)
4. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex networks: Structure and dynamics. Physics reports 424(4), 175–308 (2006)
5. Bollobás, B., Riordan, O.M.: Mathematical results on scale-free random graphs. Handbook of graphs and networks: from the genome to the internet pp. 1–34 (2003)
6. Bromberger, S., other contributors: Juliagraphs/lightgraphs.jl: Lightgraphs v0.10.5 (Sep 2017)
7. Chung, F., Lu, L.: Complex Graphs and Networks. American Mathematical Society, Providence, Rhode Island, USA (2006)
8. Cooper, C., Frieze, A., Prałat, P.: Some typical properties of the spatial preferred attachment model. Internet Mathematics 10, 27–47 (2014)
9. Costa, L.d.F., Rodrigues, F.A., Travieso, G., Villas Boas, P.R.: Characterization of complex networks: A survey of measurements. Advances in physics 56(1), 167–242 (2007)
10. Csányi, G., Szendrői, B.: Structure of a large social network. Physical Review E 69(3), 036131 (2004)
11. Dorogovtsev, S.N., Goltsev, A.V., Mendes, J.F.F.: Pseudofractal scale-free web. Physical review E 65(6), 066122 (2002)
12. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the internet topology. In: ACM SIGCOMM computer communication review. vol. 29, pp. 251–262. ACM (1999)
13. Iskhakov, L., Mironov, M., Ostroumova Prokhorenkova, L., Pralat, P.: Local clustering coefficient of spatial preferential attachment model. arXiv preprint arXiv:1711.06846 (2017)

14. Jacob, E., Mörters, P.: A spatial preferential attachment model with local clustering. In: International Workshop on Algorithms and Models for the Web-Graph. pp. 14–25. Springer (2013)
15. Jacob, E., Mörters, P., et al.: Spatial preferential attachment networks: Power laws and clustering coefficients. The Annals of Applied Probability 25(2), 632–662 (2015)
16. Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley, New York, USA (2000)
17. Janssen, J., Hurshman, M., Kalyaniwalla, N.: Model selection for social networks using graphlets. Internet Mathematics 8(4), 338–363 (2013)
18. Janssen, J., Prałat, P., Wilson, R.: Geometric graph properties of the spatial preferred attachment model. Advances in Applied Mathematics 50, 243–267 (2013)
19. Janssen, J., Prałat, P., Wilson, R.: Non-uniform distribution of nodes in the spatial preferential attachment model. Internet Mathematics 12(1–2), 121–144 (2016)
20. Kamiński, B., Olczak, T., Prałat, P.: 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. pp. 42–59. Springer (2017)
21. Krot, A., Prokhorenkova, L.O.: Local clustering coefficient in generalized preferential attachment models. In: International Workshop on Algorithms and Models for the Web-Graph. pp. 15–28. Springer (2015)
22. Leskovec, J.: Dynamics of large networks (proquest). Ann Arbor (2008)
23. Newman, M.E.: Properties of highly clustered networks. Physical Review E 68(2), 026121 (2003)
24. Newman, M.E.: The structure and function of complex networks. SIAM review 45(2), 167–256 (2003)
25. Newman, M.E.: Power laws, pareto distributions and zipf's law. Contemporary physics 46(5), 323–351 (2005)
26. Ostroumova Prokhorenkova, L.: Global clustering coefficient in scale-free weighted and unweighted networks. Internet Mathematics 12(1-2), 54–67 (2016)
27. Ostroumova Prokhorenkova, L., Prałat, P., Raigorodskii, A.: Modularity of complex networks models, preprint (2017)
28. Pittel, B., Spencer, J., Wormald, N.: Sudden emergence of a giant k-core in a random graph. Journal of Combinatorial Theory, Series B 67, 111–151 (1996)
29. Ravasz, E., Barabási, A.L.: Hierarchical organization in complex networks. Physical Review E 67(2), 026112 (2003)
30. Serrano, M.Á., Boguna, M.: Clustering in complex networks. i. general formalism. Physical Review E 74(5), 056114 (2006)
31. Vázquez, A., Pastor-Satorras, R., Vespignani, A.: Large-scale topological and dynamical properties of the internet. Physical Review E 65(6), 066130 (2002)
32. Wormald, N.: The differential equation method for random graph processes and greedy algorithms. Lectures on Approximation and Randomized Algorithms pp. 73–155 (1999)

Appendix

Proof of Lemma 1

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 2 (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.

Proof of Theorem 1

We will use the following version of the Chernoff bound that can be found, for example, in [16, p. 27, Corollary 2.3].

Lemma 2. *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 $X_i \in \mathrm{Be}(p_i)$ with (possibly) different $p_i = \mathbb{P}(X_i = 1) = \mathbb{E}X_i$. If $\varepsilon \le 3/2$, then*

$$\mathbb{P}(|X - \mathbb{E}X| \ge \varepsilon \mathbb{E}X) \le 2 \exp\left(-\frac{\varepsilon^2 \mathbb{E}X}{3}\right). \tag{7}$$

Let us start with the following key lemma.

Lemma 3. *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, $T \leq t \leq 2T$,*

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

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.

We use the following stopping time

$$T_0 = \min \left\{ t \geq T : \left(\deg^-(v,t) > d \cdot \left(\frac{t}{T}\right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{t}{T} \sqrt{d \log n} \right) \ \vee \ (t = 2T+1) \right\}.$$

Note that if $T_0 = 2T + 1$, then the in-degree of v remained bounded as required during the entire time interval $T \leq t \leq 2T$. Hence, in order to prove the bound, we need to show that with probability $1 - o(n^{-6})$ we have $T_0 = 2T + 1$.

Suppose that $T_0 \leq 2T$. Note that for $t \geq T$ up to and including time-step $T_0 - 1$, the random variable $\deg^-(v,t)$ is (deterministically) bounded from above. Hence, the number of new neighbours accumulated during this phase of the process, $\deg^-(v,T_0) - d$, can be (stochastically) bounded from above by the sum $X = \sum_{t=T}^{T_0-1} X_t$ of independent indicator random variables X_t, where

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

Clearly, since $pA_1 < 1$,

$$\begin{aligned}
\mathbb{E}X &= \sum_{t=T}^{T_0-1} \mathbb{E}X_t \\
&= pA_1 d T^{-pA_1} \left(\sum_{t=T}^{T_0-1} t^{pA_1-1} \right) + \frac{T_0 - T}{T} 5\sqrt{d \log n} + O(1) \\
&= d \left(\frac{T_0}{T}\right)^{pA_1} - d \left(\frac{T}{T}\right)^{pA_1} + \frac{T_0 - T}{T} 5\sqrt{d \log n} + O(1) \\
&= d \left(\frac{T_0}{T}\right)^{pA_1} - d + \frac{T_0 - T}{T} 5\sqrt{d \log n} + O(1).
\end{aligned}$$

Since $T_0 \le 2T$, the in-degree of v at time T_0 failed the desired condition, which implies that

$$\begin{aligned}
X &\ge \deg^-(v, T_0) - d \\
&\ge \left(d \cdot \left(\frac{T_0}{T}\right)^{pA_1} + \frac{5}{pA_1} \cdot \frac{T_0}{T} \sqrt{d \log n} \right) - d \\
&= \mathbb{E}X + \frac{5}{pA_1} \cdot \frac{T_0}{T} \sqrt{d \log n} - \frac{T_0 - T}{T} 5\sqrt{d \log n} + O(1) \\
&\ge \mathbb{E}X + 5\sqrt{d \log n},
\end{aligned}$$

using again that it is assumed that $pA_1 < 1$. It follows from the Chernoff bound (7) that

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

where $\varepsilon = 5\sqrt{d \log n}/\mathbb{E}X$. The maximum value of $\mathbb{E}X$ corresponds to $T_0 = 2T$ and so

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

So $\varepsilon \ge 5\sqrt{d^{-1} \log n}$. Therefore, the probability that $T_0 \le 2T$ is at most $2 \exp(-(25/3) \log n) = o(n^{-6})$ and the proof is finished.

Now, with Lemma 3 in hand we can get Theorem 1.

Proof (Proof of Theorem 1). Let $\omega = \omega(n)$ be a function going to infinity with n. Let v be a vertex with final degree $k \ge \omega \log n$. 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 3 that, with probability $1 - o(n^{-6})$,

$$\begin{aligned}
&d \left(\frac{t}{T}\right)^{pA_1} \left(1 - \frac{5}{pA_1}\sqrt{d^{-1} \log n}\right) \\
&\qquad\qquad \le \deg^-(v, t) \le d \left(\frac{t}{T}\right)^{pA_1} \left(1 + \frac{5}{pA_1}\sqrt{d^{-1} \log n}\right)
\end{aligned}$$

for $T \le t \le 2T$. 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 3 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 $1 \le i < i_{\max} = \lfloor \log_2 n \rfloor + 1$, let $d_i = \deg^-(v, 2^i T)$. Then by Lemma 3, we have for $i > 1$ that $d_i \le d_{i-1} 2^{pA_1}(1+\varepsilon_i)$, 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 $2^{i-1}T \le t < 2^i T$, we have that

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

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

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

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.

Proof of Theorem 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 [1] 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

$$\mathbb{E}(N_{0,t+1} - N_{0,t} \mid G_t) = b - N_{0,t}p\frac{A_2}{t}, \tag{8}$$

$$\mathbb{E}(N_{i,t+1} - N_{i,t} \mid G_t) = N_{i-1,t}p\frac{A_1(i-1)+A_2}{t} - N_{i,t}p\frac{A_1 i + A_2}{t}. \tag{9}$$

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 [7].

Lemma 4. *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 $i \le i_f$ (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 prove that deviations from the c-Lipschitz condition occur with very small probability. Lemma 1 gives a deterministic bound for $|N_{i,t+1} - N_{i,t}|$ which holds with high probability. Indeed, it is obvious that $|N_{i,t+1} - N_{i,t}| \le \max\{\deg^+(v_{t+1}, t+1), 1\}$.

Let us 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. [28], as described in [32].

Lemma 5. *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 $\alpha > 0$,

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

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

Theorem 5. *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 $N_{0,s}$ $(1 \le s \le t)$ 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 (8)), 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

$$H(x, y) = x^{pA_2} y - \frac{bx^{1+pA_2}}{1 + pA_2} \tag{10}$$

(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 $(H(\mathbf{w}_s) : 1 \le s \le t)$. The second-order partial derivatives of H evaluated at $\mathbf{w}_s$ are all $O(s^{pA_2-1})$. Moreover, it follows from Lemma 1 that we may assume that

$$|N_{0,s+1} - N_{0,s}| \le \omega \log n. \tag{11}$$

Therefore, we have

$$H(\mathbf{w}_{s+1}) - H(\mathbf{w}_s) = (\mathbf{w}_{s+1} - \mathbf{w}_s) \cdot \text{grad } H(\mathbf{w}_s) + O(s^{pA_2-1}\omega^2 \log^2 n), \tag{12}$$

where "$\cdot$" denotes the inner product and 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

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

Hence, taking the expectation of (12) 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 (12) and (11), noting that

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

we have that

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

Our goal is to apply Lemma 5 to the sequence $(H(\mathbf{w}_s) : 1 \le s \le t)$ to get an upper bound for $H(\mathbf{w}_s)$. A symmetric argument applied to $(-H(\mathbf{w}_s) : 1 \le s \le t)$ 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 $H(\mathbf{w}_s) \ge 0$, 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

$$\sum_{s=1}^{t}\beta_s = O\left(\sum_{s=1}^{t} s^{pA_2-1}\omega^2\log^2 n\cdot(b+s^{-1})\right) \tag{13}$$
$$= O\left(b\,\omega^2\log^2 n\sum_{s=1}^{t} s^{pA_2-1}\right)+O\left(\omega^2\log^2 n\sum_{s=1}^{t} s^{pA_2-2}\right)$$
$$= O\left(b\,t^{pA_2}\omega^2\log^2 n\right)+O\left(t^{pA_2-1}\omega^2\log^2 n\right)=O\left(b\,t^{pA_2}\omega^2\log^2 n\right),$$

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

$$\sum_{s=1}^{t}\gamma_s^2 = O\left(\sum_{s=1}^{t}(s^{pA_2}\omega\log n)^2\cdot(b+s^{-1})\right)=O\left(b\,t^{1+2pA_2}\omega^2\log^2 n\right). \tag{14}$$

Finally, we are ready to apply Lemma 5 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 (10) 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.

We may repeat (recursively) the argument as in the proof of Theorem 5 for $N_{i,t}$ with $i\geq 1$. Since the expected change for $N_{i,t}$ is slightly different now (see (9)), 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 (13) and (14)) 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. *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 2.

Proof (Proof of Theorem 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 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 2 (applied with, say, $\sqrt{\omega}$ instead of ω) we may assume that any vertex v_i born after time T_v satisfies (for any $T_v \le t \le n$)

$$\deg(v_i,t) \le \sqrt{\omega} \log n \left(\frac{t}{i}\right)^{pA_1} = o\left(\omega \log n \left(\frac{t}{i}\right)^{pA_1}\right) = o(\deg(v,t)); \qquad (15)$$

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 $t \geq t_{\ell+1}$, 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

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

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 (15) to get that

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

The fact that, with the desired probability, $\alpha = O(k)$ follows from a standard martingale argument (for example, one could use Lemma 5).

Similarly, we can deal with γ. It follows from (15) that no vertex born after time

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

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

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

Finally, we need to deal with β. This time, we need to use Theorem 6 to count (independently) the number of vertices in $B_{\ell-1}$ of a certain degree. We may apply this theorem as for any $L' \le \ell \le L-1$, we have

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

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

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

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 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-neigbours 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.

Proof of Theorem 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

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

provided that

$$\max\left\{\beta \log \alpha, \beta^2 \log(1/p)\right\} \leq \frac{1}{5} \log n. \tag{16}$$

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

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

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

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

provided that

$$\max\left\{\beta \log \alpha, \beta \log \beta\right\} \leq \frac{1}{5} \log n. \tag{17}$$

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 3, 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

$$\beta \geq \omega^3\, (\log\log n)^2 (\log\log\log n). \tag{18}$$

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 1 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 (16) 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 (17) and (18) 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 (17) and (18) are satisfied, and the proof is finished.

5.3 Proof of Theorem 4

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 $1+1/(pA_1)$ [1]. Let

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

Note that $T \ge n^{\varepsilon/(pA_1)}$, as $k \le n^{pA_1-\varepsilon}$. It follows from Theorem 1 that a.a.s., for each $v \in X_k$,

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

In particular, for n large enough,

$$\deg^-(v,T) > \omega\log n \tag{19}$$

(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 2 to be able to assume that for any $u \le t \le n$,

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

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

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

(Here we used Theorem 1 and (19).) 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

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

Since the expected number of such points is at most

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

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 3 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.