Modularity of complex networks models

Liudmila Ostroumova Prokhorenkova[1,2], Paweł Prałat[3,4], and Andrei Raigorodskii[1,2,5,6]

[1] Moscow Institute of Physics and Technology, Moscow, Russia
[2] Yandex, Moscow, Russia
[3] Ryerson University, Toronto, ON, Canada
[4] The Fields Institute for Research in Mathematical Sciences, Toronto, ON, Canada
[5] Moscow State University, Moscow, Russia
[6] Buryat State Unversity, Ulan-Ude, Buryat Republic, Russia

Abstract. Modularity is designed to measure the strength of division of a network into clusters (known also as communities). Networks with high modularity have dense connections between the vertices within clusters but sparse connections between vertices of different clusters. As a result, modularity is often used in optimization methods for detecting community structure in networks, and so it is an important graph parameter from practical point of view. Unfortunately, many existing non-spatial models of complex networks do not generate graphs with high modularity; on the other hand, spatial models naturally create clusters. We investigate this phenomena by considering a few examples from both sub-classes. We prove precise theoretical results for the classical model of random d-regular graphs as well as the preferential attachment model, and contrast these results with the ones for the spatial preferential attachment (SPA) model that is a model for complex networks in which vertices are embedded in a metric space, and each vertex has a sphere of influence whose size increases if the vertex gains an in-link, and otherwise decreases with time.

1 Introduction and definitions

Many social, biological, and information systems can be represented by networks, whose vertices are items and links are relations between these items [2, 7, 9, 15]. That is why the evolution of complex networks attracted a lot of attention in recent years and there has been a great deal of interest in modelling of these networks [12, 21, 36]. The hyperlinked structure of the Web, citation patterns, friendship relationships, infectious disease spread, these are seemingly disparate linked data sets which have fundamentally very similar natures. Indeed, it turns out that many real-world networks have some typical properties: power-law degree distribution, small diameter, high clustering coefficient, and others [33, 35, 39]. Such properties are well-studied both in real-world networks and in many theoretical models.

Another important property of complex networks is their community structure, that is, the organization of vertices in clusters, with many edges joining

vertices of the same cluster and comparatively few edges joining vertices of different clusters [18, 23]. In social networks communities may represent groups by interest, in citation networks they correspond to related papers, in the Web communities are formed by pages on related topics, etc. Being able to identify communities in a network could help us to exploit this network more effectively. For example, clusters in citation graphs may help to find similar scientific papers, discovering users with similar interests is important for targeted advertisement, clustering can also be used for network compression and visualization.

The key ingredient for many clustering algorithms is *modularity*, which is at the same time a global criterion to define communities, a quality function of community detection algorithms, and a way to measure the presence of community structure in a network. Modularity was introduced by Newman and Girvan [37] and it is based on the comparison between the actual density of edges inside a community and the density one would expect to have if the vertices of the graph were attached at random, regardless of community structure.

Unfortunately, modularity is not a well studied parameter for the existing random graph models, at least from rigorous, theoretical point of view. We are only aware about results for binomial random graphs $G(n,p)$ and random d-regular graphs (see Section 2.3 for more details). In this paper, we continue investigating random d-regular graphs and obtain new upper bounds for their modularity. Then we move to the *preferential attachment model*, introduced by Barabási and Albert [8], which is probably the most well-studied model of complex networks. For this model no results on modularity are known and we obtain some preliminary results, both lower and upper bounds, and will investigate this model more in the journal version of this paper. In fact, the lower bound we present holds for all graphs with average degree d and sublinear maximum degree.

As expected, the models discussed above, as well as many others, have a common weakness of low modularity. One family of models which overcomes this deficiency is the family of spatial (or geometric) models, wherein the vertices are embedded in a metric space such that similar vertices are closer to each other than dissimilar ones. The underlying geometry of spatial models naturally leads to the emergence of clusters. We prove this statement rigorously for one example of a geometric model, the Spatial Preferential Attachment model introduced in [1].

The paper is structured as follows. In the next section, we formally define modularity, discuss several random graph models and present known results on modularity in these models. In Sections 3, 4 and 5 we analyze modularity in random d-regular graphs, preferential attachment and SPA models, respectively.

Due to the space limitations, proofs of our results are omitted in this short proceeding version but will be included in the longer journal one.

2 Preliminaries

2.1 Modularity

The definition of modularity was first introduced by Newman and Girvan in [37]. Since then, many popular and applied algorithms used to find clusters in large data-sets are based on finding partitions with high modularity [20, 28, 34]. The modularity function favours partitions in which a large proportion of the edges fall entirely within the parts and biases against having too few or too unequally sized parts. Formally, for a given partition $\mathcal{A} = \{A_1, \ldots, A_k\}$ of the vertex set $V(G)$, let

$$q_{\mathcal{A}} = \sum_{A \in \mathcal{A}} \left(\frac{e(A)}{|E(G)|} - \frac{\left(\sum_{v \in A} \deg(v)\right)^2}{4|E(G)|^2} \right), \tag{1}$$

where $e(A) = |\{uv \in E(G) : u, v \in A\}|$ is the number of edges in the graph induced by the set A. The first term, $\sum_{A \in \mathcal{A}} \frac{e(A)}{|E(G)|}$, is called the *edge contribution*, whereas the second one, $\sum_{A \in \mathcal{A}} \frac{\left(\sum_{v \in A} \deg(v)\right)^2}{4|E(G)|^2}$, is called the *degree tax*. The *modularity* $q^*(G)$ is then defined as the maximum of $q_{\mathcal{A}}$ over all possible partitions $\mathcal{A}$ of $V(G)$; that is,

$$q^*(G) = \max_{\mathcal{A}} q_{\mathcal{A}}(G).$$

In order to maximize $q_{\mathcal{A}}(G)$ one wants to find a partition with large edge contribution subject to small degree tax. If $q^*(G)$ approaches 1 (which is the maximum), we observe a strong community structure; conversely, if $q^*(G)$ is close to zero, we are given a graph with no community structure.

2.2 Random graph models

Classical models. Let us recall two classical models of random graphs that are extensively studied in the literature. The *binomial random graph* $\mathcal{G}(n, p)$ is the random graph G with the vertex set $[n] := \{1, 2, \ldots, n\}$ in which every pair $\{i, j\} \in \binom{[n]}{2}$ appears independently as an edge in G with probability p. Note that $p = p(n)$ may (and usually does) tend to zero as n tends to infinity.

However, in this paper we concentrate on another probability space, the probability space of *random d-regular graphs* with uniform probability distribution. This space is denoted $\mathcal{G}_{n,d}$, and asymptotics are for $n \to \infty$ with $d \geq 2$ fixed, and n even if d is odd.

We say that an event in a probability space holds *asymptotically almost surely* (or *a.a.s.*) if the probability that it holds tends to 1 as n goes to infinity. Since we aim for results that hold a.a.s., we will always assume that n is large enough.

Preferential Attachment. The *Preferential Attachment* (PA) model, introduced by Barabási and Albert [8], was an early stochastic model of complex networks. We will use the following precise definition of the model, as considered

by Bollobás and Riordan in [13] as well as Bollobás, Riordan, Spencer, and Tusnády [14].

Let G_1^0 be the null graph with no vertices (or let G_1^1 be the graph with one vertex, v_1, and one loop). The random graph process $(G_1^t)_{t\geq 0}$ is defined inductively as follows. Given G_1^{t-1}, we form G_1^t by adding a vertex v_t together with a single edge between v_t and v_i, where i is selected randomly with the following probability distribution:

$$\mathbb{P}(i=s)=\begin{cases}\deg(v_s,t-1)/(2t-1) & 1\leq s\leq t-1,\\ 1/(2t-1) & s=t,\end{cases}$$

where $\deg(v_s,t-1)$ denotes the degree of v_s in G_1^{t-1}. In other words, we send an edge e from v_t to a random vertex v_i, where the probability that a vertex is chosen is proportional to its degree at the time, counting e as already contributing one to the degree of v_t.

For $m\in\mathbb{N}\setminus\{1\}$, the process $(G_m^t)_{t\geq 0}$ is defined similarly with the only difference that m edges are added to G_m^{t-1} to form G_m^t (one at a time), counting previous edges as already contributing to the degree distribution. Equivalently, one can define the process $(G_m^t)_{t\geq 0}$ by considering the process $(G_1^t)_{t\geq 0}$ on a sequence $v_1', v_2', \ldots$ of vertices; the graph G_m^t if formed from G_1^{tm} by identifying vertices $v_1', v_2', \ldots, v_m'$ to form v_1, identifying vertices $v_{m+1}', v_{m+2}', \ldots, v_{2m}'$ to form v_2, and so on. Note that in this model G_m^t is in general a multigraph, possibly with multiple edges between two vertices (if $m\geq 2$) and self-loops.

It was shown in [14] that for any $m\in\mathbb{N}$ a.a.s. the degree distribution of G_m^n follows a power law: the number of vertices with degree at least k falls off as $(1+o(1))ck^{-2}n$ for some explicit constant $c=c(m)$ and large $k\leq n^{1/15}$. Also, in the case $m=1$, G_1^n is a forest. Each vertex sends an edge either to itself or to an earlier vertex, so the graph consists of components which are trees, each with a loop attached. The expected number of components is then $\sum_{t=1}^n 1/(2t-1)\sim(1/2)\log n$ and, since events are independent, we derive that a.a.s. there are $(1/2+o(1))\log n$ components in G_1^n by Chernoff's bound. Moreover, Pittel [38] essentially showed that a.a.s. the largest distance between two vertices in the same component of G_1^n is $(\gamma^{-1}+o(1))\log n$, where γ is the solution of $\gamma e^{1+\gamma}=1$. In contrast, for the case $m\geq 2$ it is known that a.a.s. G_m^n is connected and its diameter is $(1+o(1))\log n/\log\log n$ [13].

Spatial Preferential Attachment. The *Spatial Preferential Attachment* (SPA) model [1], designed as a model for the World Wide Web, combines geometry and preferential attachment, as its name suggests. Setting the SPA model apart is the incorporation of 'spheres of influence' to accomplish preferential attachment: the greater the degree of a vertex, the larger its sphere of influence, and hence the higher the likelihood of the vertex gaining more neighbours.

We now give a precise description of the SPA model. Let $S=[0,1]^m$ be the unit hypercube in $\mathbb{R}^m$, equipped with the torus metric derived from any of the L_p norms. This means that for any two points x and y in S,

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

The torus metric thus 'wraps around' the boundaries of the unit square; this metric was chosen to eliminate boundary effects. The parameters of the model consist of the *link probability* $p \in [0,1]$, and two positive constants A_1 and A_2, which, in order to avoid the resulting graph becoming too dense, must be chosen so that $pA_1 < 1$. The SPA model generates stochastic sequences of directed graphs $(G_t : t \geq 0)$, where $G_t = (V_t, E_t)$, and $V_t \subseteq S$. Let $\deg^-(v,t)$ be the in-degree of the vertex v in G_t, and $\deg^+(v,t)$ its out-degree. We define the *sphere of influence* $S(v,t)$ of the vertex v at time $t \geq 1$ to be the ball centered at v with volume $|S(v,t)|$ defined as follows:

$$|S(v,t)| = \frac{A_1 \deg^-(v,t) + A_2}{t}, \tag{2}$$

or $S(v,t) = S$ and $|S(v,t)| = 1$ if the right-hand-side of (2) is greater than 1.

The process begins at $t = 0$, with G_0 being the null graph. Time-step t, $t \geq 1$, is defined to be the transition between G_{t-1} and G_t. At the beginning of each time-step t, a new vertex v_t is chosen *uniformly at random* from S, and added to V_{t-1} to create V_t. Next, 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 $p\,|S(u, t-1)|$.

The SPA model produces scale-free networks, which exhibit many of the characteristics of real-life networks (see [1, 16]). In [25], 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. In [26], some properties of common neighbors where used to explore the underlying geometry of the SPA model and quantify vertex similarity based on distance in the space. However, the distribution of vertices in space was assumed to be uniform [26] and so in [27] non-uniform distributions were investigated which is clearly a more realistic setting.

2.3 Previous results on modularity

In this section we discuss known bounds for modularity in different random graph models.

The *isoperimetric number* of a graph G is defined as

$$j(G) = \min_{V(G)=V_1 \cup V_2} \frac{e(V_1, V_2)}{\min\{|V_1|, V_2|\}},$$

where $e(V_1, V_2) = |\{uv \in E(G) : u \in V_1, v \in V_2\}|$ is the number of edges between sets V_1 and V_2. The following result was shown by McDiarmid and Skerman in [29]. Let G be any d-regular graph on n vertices. As mentioned in [29], the following useful upper bound on the modularity is almost immediate:

$$q^*(G) \leq \max\{1 - i(G)/d, 3/4\}. \tag{3}$$

Turning to random d-regular graph, Bollobás in [11] showed that a.a.s. $i(\mathcal{G}_{n,d}) \geq (1-\eta)d/2$, where $0 < \eta < 1$ is such that $2^{4/d} < (1-\eta)^{1-\eta}(1+\eta)^{1+\eta}$ and so a.a.s.

$$q^*(\mathcal{G}_{n,d}) \leq U_1 = U_1(d) := \max\{1/2 + \eta/2, 3/4\}.$$

As a result, we get the first non-trivial upper bounds for $q^*(\mathcal{G}_{n,d})$ presented in Table 2.3 that hold a.a.s.

In [29], the bound (3) was slightly improved when the maximum size of parts in our partition is restricted. Formally, given $\delta > 0$, for a graph G with $n \geq 1/\delta$ vertices, they define $q_\delta(G)$ to be the maximum modularity of all partitions for G such that each part has size at most δn. They show that for any $\varepsilon > 0$ there exists $\delta > 0$ such any d-regular graph with at least $1/\delta$ vertices satisfies

$$q_\delta(G) \leq 1 - 2i(G)/d + \varepsilon.$$

Again, using the result of Bollobás we get that there exists $\delta > 0$ such that

$$U_2 = U_2(d) := \eta + \varepsilon$$

serves as an upper bound that holds a.a.s. for $q_\delta(\mathcal{G}_{n,d})$; again, see Table 2.3 for numerical values for small values of d. It is straightforward to see that $i(G) \geq d/2 - \sqrt{(\log 2)d}$ (see, for example, [11]) and so, in particular, U_2 can be made arbitrarily small by taking d large enough (and δ small enough). However, let us note that these upper bounds for $q_\delta(\mathcal{G}_{n,d})$, while useful, cannot be directly translated into any bound for $q^*(\mathcal{G}_{n,d})$.

d	U_1	U_2	U_3
3	0.9386	0.8771	0.8038
4	0.8900	0.7800	0.6834
5	0.8539	0.7078	0.6024
6	0.8261	0.6521	0.5435
7	0.8038	0.6076	0.4984
8	0.7855	0.5710	0.4624
9	0.7702	0.5403	0.4330
10	0.7570	0.5140	0.4083

Table 1. Upper bounds for $q^*(\mathcal{G}_{n,d})$ and for $q_\delta(\mathcal{G}_{n,d})$ (U_2)

Investigating random d-regular graphs continues in [30], a very recent paper. In fact, some of our results for this model mentioned below are obtained independently there. Moreover, they investigate the class of graphs whose product of treewidth and maximum degree is much less than the number of edges. This shows, for example, that random planar graphs typically have modularity close to 1, which is another indication that clusters naturally emerge where geometry is included.

3 Random *d*-regular graphs

3.1 Lower bound

For completeness, let us briefly discuss the following known lower bound for the modularity of $\mathcal{G}_{n,d}$. It is known that a.a.s. for any $d \in \mathbb{N} \setminus \{1, 2\}$, $\mathcal{G}_{n,d}$ is

Hamiltonian. As pointed out in [29], one can uses this fact to partition the graph such that it breaks the cycle into $\lceil\sqrt{n}\rceil$ paths of length at most $\lceil\sqrt{n}\rceil$. For this particular partition the edge contribution is $2/d - O(1/\sqrt{n})$ and the degree tax is $O(1/\sqrt{n})$. It follows then that a.a.s.

$$q^*(\mathcal{G}_{n,d}) \geq \frac{2}{d} - O(1/\sqrt{n}) = \frac{2+o(1)}{d}.$$

(Our more general lower bound that holds for graphs with average degree d implies the same—see Theorem 4 for more.) Whereas this trivial lower bound could be sharp for $d = 3$ it is definitely not the case for large d. As pointed out in [30], there exists a universal constant $c > 0$ such that a.a.s. $q^*(\mathcal{G}_{n,d}) \geq c/\sqrt{d}$.

3.2 Slightly improved, numerical upper bound

Let us consider the following, natural, approach that already improves slightly an upper bound for $q^*(\mathcal{G}_{n,d})$. Consider any d-regular graph with n vertices. For a given partition $\mathcal{A} = \{A_1, \ldots, A_k\}$ of the vertex set $V(G)$, let $x_i = |A_i|/n$ and $y_i = 2|E(A_i)|/|A_i|$; that is, set A_i has $x_i n$ vertices and induces $y_i x_i n/2$ edges. Then (1) can be simplified to

$$q_{\mathcal{A}} = \sum_{i=1}^{k} x_i \left(\frac{y_i}{d} - x_i\right). \tag{4}$$

As it is simply a weighted average, $q_{\mathcal{A}} \geq U$ would imply that there exists some set of size xn that induces $yxn/2$ edges, and $y/d - x \geq U$. Using the pairing model, we will show that a.a.s. it is not the case (for some carefully chosen $U = U(d)$) and, as a result, it will yield an upper bound for $q^*(\mathcal{G}_{n,d})$ that holds a.a.s.

For a given $d \in \mathbb{N} \setminus \{1, 2\}$, let

$$\begin{aligned} f(x,y,d) := {} & x(y/2-1)\log(x) + (1-x)(d-1)\log(1-x) + d\log(d)/2 \\ & - xy\log(y)/2 - x(d-y)\log(d-y) - (d-2xd+xy)\log(d-2xd+xy)/2. \end{aligned} \tag{5}$$

It will be clear once we establish the connection between function f and random d-regular graphs, but it is straightforward to see that for any $x \in (0,1)$ we have $f(x,d,d) < 0$ and $f(x,y,d) > 0$ for some $y \in (0,d)$. Indeed, for example note that for a fixed $x \in (0, 1/2]$, $f(x,y,d)$ is strictly concave in $y \in (0,d)$ as

$$\frac{d^2 f(x,y,d)}{dy^2} = \frac{-(d(1-2x)+y)dx}{2(d(1-2x)+xy)(d-y)y} < 0.$$

Let $y_3 = y_3(x,d)$ be largest value of $y \in (0,1)$ such that $f(x,y,d) = 0$; in particular, $f(x,y,d) \leq 0$ for any $y \in (y_3, d)$. Finally, let

$$U_3 = U_3(d) := \sup_{x \in (0,1)} \left(\frac{y_3(x,d)}{d} - x\right).$$

As usual, see Table 2.3 for numerical values for small values of d. The promised upper bound follows immediately from the following theorem.

Theorem 1 *Let $d \in \mathbb{N} \setminus \{1,2\}$ and $\varepsilon > 0$ be an arbitrarily small constant. The following property holds a.a.s. for $\mathcal{G}_{n,d}$. No set A of size xn (for any $x = x(n) \in (0,1)$) induces a graph with $yxn/2$ edges, where $y_3(x,d)+\varepsilon \le y \le d$ and $y_3(x,d)$ is defined as above. In particular, this implies that*

$$q^*(\mathcal{G}_{n,d}) \le U_3 + \varepsilon/d,$$

where $U_3 = U_3(d)$ is defined as above.

3.3 Explicit but weaker upper bound

Theorem 1 provides an upper bound that can be easily numerically computed for a given $d \in \mathbb{N} \setminus \{1,2\}$. Next, we present a slightly weaker but an explicit bound that can be obtained using the expansion properties of random d-regular graphs that follow from their eigenvalues. In particular, it will imply that a.a.s. $q^*(\mathcal{G}_{n,d}) = O(1/\sqrt{d})$ and so $q^*(\mathcal{G}_{n,d}) \to 0$ as $d \to \infty$.

Theorem 2 *Let $d \in \mathbb{N} \setminus \{1,2\}$ and $\varepsilon > 0$ be an arbitrarily small constant. The following property holds a.a.s. for $\mathcal{G}_{n,d}$. No set A of size xn induces a graph with more than $yxn/2$ edges, where $y = dx + \lambda(1-x)$. In particular, this implies that a.a.s.*

$$q^*(\mathcal{G}_{n,d}) \le \frac{\lambda}{d} \le \frac{2\sqrt{d-1}+\varepsilon}{d} \le \frac{2}{\sqrt{d}}.$$

4 The Preferential Attachment model

4.1 Constant average degree graphs

In order to obtain a lower bound for modularity of Preferential Attachment graphs, we first analyze graphs with constant average degree in general. In this section, we extend the results of [32] and we start with the analysis of trees. It was proven in [32] that trees with maximum degree $\Delta = o(\sqrt[5]{n})$ have asymptotic modularity 1. We generalize this result in two ways: first, we relax the condition on maximum degree; second, we allow our graphs to be disconnected, that is, we consider forests instead of trees. We prove the following theorem.

Theorem 3 *Let $\{F_n\}$ be a sequence of forests, F_n is a forest on n vertices with no isolated ones and $\Delta = \Delta(F_n) = o(n)$. Then $q^*(F_n) \ge 1 - O\left(\sqrt{\frac{\Delta}{n}}\right) = 1 - o(1)$ as $n \to \infty$.*

Note that it is also known that the asymptotic modularity of trees with maximum degree $\Delta = \Omega(n)$ is strictly less than 1 [32]. Hence, the assumption $\Delta = o(n)$ cannot be eliminated. Now let us use the previous theorem to get the following result for graphs of bounded average degree.

Theorem 4 *Let $\{G_n\}$ be a sequence graphs, G_n is a connected graph on n vertices with no isolated ones, with the average degree $\frac{2|E(G_n)|}{n} \le D$ for some constant D, and $\Delta = \Delta(G_n) = o(n)$. Then $q^*(G_n) \ge \frac{2}{D} - O\left(\sqrt{\frac{\Delta}{n}}\right) = \frac{2}{D} - o(1)$.*

4.2 Lower bound

The following theorem easily follows from the above result.

Theorem 5 *For any $\varepsilon > 0$ a.a.s.*

$$q^*(G_m^n) \geq \frac{1}{m} - O\left(n^{-1/4+\varepsilon}\right) = \frac{1}{m} - o(1).$$

As in the case of random d-regular graphs, it is natural to conjecture that the above lower bound is not sharp. Let $c \in (0,1)$ and consider the following partition: $A_1 = \{v_1, \ldots, v_{cn}\}$, $A_2 = V(G_m^n) \setminus A_1 = \{v_{cn+1}, \ldots, v_n\}$. Using martingales, it is possible to show that a.a.s. $\sum_{v \in A_1} \deg(v,n) \sim 2mn\sqrt{c}$ (and so $\sum_{v \in A_2} \deg(v,n) \sim 2mn(1-\sqrt{c})$). Clearly, $e(A_1) = mcn$ and so a.a.s. $e(A_1, A_2) \sim 2mn(\sqrt{c} - c)$ and $e(A_2) \sim mn(1 + c - 2\sqrt{c})$. The edge contribution and the degree tax are then both asymptotic to $1 + 2c - 2\sqrt{c}$. Not surprisingly, such partition cannot be used to get a non-trivial lower bound for the modularity but, similarly to the situation for random d-regular graphs, we may try to use it as a starting point to get slightly better partition. The basic idea is very simple: one can start with a given partition (or partition the vertices randomly into two classes), and if a vertex has more neighbours in the other class than in its own, then we randomly decide whether to shift it to the other class or leave it where it is. This approach proved to be useful to get a bound for the bisection width in random d-regular graphs [3] which, in turn, yields a lower bound for the modularity [30]. We plan to investigate it further in the journal version of this paper.

4.3 Upper bound

The *edge expansion* $\rho = \rho(G)$ of a graph G is defined as follows:

$$\rho = \min_{S \subset V(G), |S| \leq |V|/2} \frac{e(S, V \setminus S)}{|S|}.$$

In [31] it was shown that a.a.s. $\rho(G_m^n) \geq \alpha$, provided that $2(m-1) - 4\alpha - 1 > 0$. In other words, for any $\varepsilon > 0$ we have that a.a.s.

$$\rho(G_m^n) \geq \frac{m}{2} - \frac{3+\varepsilon}{4}.$$

Using this observation one can easily obtain a non-trivial upper bound for $q^*(G_m^n)$.

Let $\varepsilon > 0$ be an arbitrary small constant. Consider any partition $\mathcal{A} = \{A_1, \ldots, A_k\}$ of the vertex set $V(G_m^n)$. If $|A_i| > n/2$ for some i, then the degree tax is at least

$$\frac{(\sum_{v \in A_i} \deg(v,n))^2}{4|E(G_m^n)|} \geq \frac{(m|A_i|)^2}{4(mn)^2} = \frac{1}{16}.$$

On the other hand, if $|A_i| \le n/2$ for all i, then a.a.s. the number of edges between parts is equal to

$$\frac{1}{2}\sum_{i=1}^{k} e(A_i, V \setminus A_i) \ge \frac{1}{2}\sum_{i=1}^{k} \rho|A_i| = \frac{\rho n}{2} \ge \left(\frac{m}{4} - \frac{3+\varepsilon}{8}\right) n,$$

and so the edge contribution is a.a.s. at most

$$1 - \left(\frac{1}{4} - \frac{3+\varepsilon}{8m}\right) = \frac{3}{4} + \frac{3+\varepsilon}{8m} \le \frac{15+\varepsilon}{16},$$

for any $m \ge 2$. The following result holds.

Theorem 6 *For any $\varepsilon > 0$ a.a.s.*

$$q^*(G_2^n) \le \frac{15+\varepsilon}{16}.$$

Moreover, for any $m \ge 3$ a.a.s.

$$q^*(G_m^n) \le \frac{15}{16}.$$

Much stronger expansion property was recently obtained in [19]. We are currently working on using this property to obtain general upper bound for $q^*(G_m^n)$ that holds for any integer m as well as specific stronger bounds for small values of m. Details will be provided in the journal version of this paper.

5 The Spatial Preferential Attachment model

Consider $G_n = (V_n, E_n)$, a graph generated by the SPA model. As the modularity is defined for undirected graphs, we consider $\hat{G}_n$ that is a graph obtained from G_n by replacing each directed edge (u, v) by undirected edge uv. (As edges in G_n are always from 'older' to 'younger' vertices, there is no problem with generating multigraph; $\hat{G}_n$ is a simple graph.) Let us recall that $V_n \subseteq S$ where S is the unit hypercube $[0,1]^m$. We will use the geometry of the model to obtain a suitable partition that yields high modularity of G_n. The following two properties (proved many times; see, for example, [1, 16]) are the only properties of the model that will be used in the proof: a.a.s. for every pair i, t such that $1 \le i \le t \le n$ we have that

$$\deg^-(v_i, t) = O\Big((t/i)^{pA_1} \log^2 n\Big) \tag{6}$$

$$\deg^+(v_i, t) = O\Big(\log^2 n\Big), \tag{7}$$

and $|E(G_n)| = \Theta(n)$. Since we aim for a result that holds a.a.s., we may assume in the proof below that these properties hold deterministically. Now, we are ready to state our result for the SPA model.

Theorem 7 *Let $p \in (0,1]$, $A_1, A_2 > 0$, and suppose that $pA_1 < 1$. Then, the following holds a.a.s.:*

$$q^*(\hat{G}_n) = 1 - O\left(n^{\max\{-1/m, -1+pA_1\}/2} \log^{9/2} n\right) = 1 - o(1).$$

Acknowledgements

This work is supported by Russian Science Foundation (grant number 16-11-10014), NSERC, The Tutte Institute for Mathematics and Computing, and Ryerson University.

References

1. W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat, A spatial web graph model with local influence regions, Internet Mathematics 5 (2009), 175–196.
2. R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Reviews of modern physics, vol. 74 (2002), 47–97.
3. N. Alon, On the edge-expansion of graphs, Combinatorics, Probability and Computing 6 (1997), 145–152.
4. N. Alon and F.R.K. Chung, Explicit construction of linear sized tolerant networks, Discrete Math., 72 (1988), 15–19.
5. N. Alon and V.D. Milman, λ_1, isoperimetric inequalities for graphs and superconcentrators, J. Combinatorial Theory, Ser. B 38 (1985), 73–88.
6. N. Alon and J.H. Spencer, The Probabilistic Method, Wiley, 1992 (Second Edition, 2000).
7. S. Bansal, S. Khandelwal, L.A. Meyers, Exploring biological network structure with clustered random networks, BMC Bioinformatics, 10:405 (2009)
8. A.L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* **286** (1999) 509–512.
9. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, Physics reports, vol. 424(45) (2006), 175-308.
10. B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, European J. Combin. 1 (1980), no. 4, 311–316.
11. B. Bollobás, The isoperimetric number of random regular graphs, Europ. J. Combinatorics 9 (1988), 241–244.
12. B. Bollobás, O.M. Riordan, Mathematical results on scale-free random graphs, Handbook of Graphs and Networks: From the Genome to the Internet (2003), 1–34.
13. B. Bollobás, O. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24** (2004) 5–34.
14. B. Bollobás, O. Riordan, J. Spencer, G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* **18** (2001) 279–290.
15. C. Borgs, M. Brautbar, J. Chayes, S. Khanna, B. Lucier, The power of local information in social networks, Internet and Network Economics, LNCS vol. 7695 (2012), 406–419.
16. C. Cooper, A. Frieze, and P. Prałat, Some typical properties of the Spatial Preferred Attachment model, Internet Mathematics 10 (2014), 27–47.
17. A. Flaxman, A. Frieze, and T. Fenner, High degree vertices and eigenvalues in the preferential attachment graph, Internet Mathematics 2(1) (2005), 1–19.
18. S. Fortunato, Community detection in graphs, Physics Reports, vol. 486(3?5) (2010), 75–174.
19. A. Frieze, X. Pérez-Giménez, P. Prałat, and B. Reiniger, Perfect matchings and Hamiltonian cycles in the preferential attachment model, preprint.

20. A. Clauset, M.E.J. Newman, C. Moore, Finding community structure in very large networks, Phys. Rev. E 70 (2004) 066111.
21. L. da F. Costa, F. A. Rodrigues, G. Travieso, P. R. U. Boas, Characterization of complex networks: A survey of measurements, Adv. Phys. 56 (2007), 167–242.
22. J. Friedman, A proof of Alon's second eigenvalue conjecture, accepted to the Memoirs of the A.M.S., 118 pp.
23. M. Girvan and M. E. Newman, Community structure in social and biological networks, Proceedings of the National Academy of Sciences, 99(12) (2002), 7821–7826.
24. S. Hoory, N. Linial, and A. Wigderson, Expander graphs and their applications, Bull. Amer. Math. Soc. (N.S.) 43 (2006), no. 4, 439–561.
25. J. Janssen, M. Hurshman, and N. Kalyaniwalla. Model selection for social networks using graphlets. Internet Mathematics, 8(4) (2013), 338–363.
26. J. Janssen, P. Prałat, and R. Wilson, Geometric Graph Properties of the Spatial Preferred Attachment model, Advances in Applied Mathematics 50 (2013), 243–267.
27. J. Janssen, P. Prałat, and R. Wilson, Non-Uniform Distribution of Nodes in the Spatial Preferential Attachment Model, Internet Mathematics 12(1-2) (2016), 121–144.
28. A. Lancichinetti and S. Fortunato, Limits of modularity maximization in community detection, Phys. Rev. E 84 (2011) 066122.
29. C. McDiarmid and F. Skerman, Modularity in random regular graphs and lattices, Electronic Notes in Discrete Mathematics 43 (2013) 431–437.
30. C. McDiarmid and F. Skerman, Modularity of tree-like and random regular graphs, preprint.
31. M. Mihail, C. Papadimitriou and A. Saberi, On certain connectivity properties of the Internet topology, Proc. Conf. on Foundations of Computer Science (2003) 28–35.
32. F. De Montgolfier, M. Soto, and L. Viennot, Asymptotic modularity of some graph classes, International Symposium on Algorithms and Computation (2011), 435–444.
33. M.E.J. Newman, Assortative mixing in networks, Phys. Rev. Letter, 89 (2002) 208701.
34. M.E.J. Newman, Fast algorithm for detecting community structure in networks, Phys. Rev. E 69 (2004) 066133.
35. M.E.J. Newman, Power laws, Pareto distributions and Zipf's law, Contemporary Physics, 46, N5 (2005), 323–351.
36. M.E.J. Newman, The structure and function of complex networks, SIAM review, 45(2) (2003), 167–256.
37. M.E.J. Newman and M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2004) 026–113.
38. B. Pittel, Note on the heights of random recursive trees and random m-ary search trees, *Random Structures and Algorithms* **5** (1994) 337–347.
39. D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998), 440–442.
40. N.C. Wormald, Models of random regular graphs, Surveys in combinatorics, 1999 (Canterbury), London Math. Soc. Lecture Note Ser., vol. 267, Cambridge Univ. Press, Cambridge, 1999, pp. 239–298.

To be included in the journal version of this paper

5.1 Paring model

Instead of working directly in the uniform probability space of random regular graphs on n vertices $\mathcal{G}_{n,d}$, we use the *pairing model* (also known as the *configuration model*) of random regular graphs, first introduced by Bollobás [10], which is described next. Suppose that dn is even, as in the case of random regular graphs, and consider dn points partitioned into n labelled buckets $v_1, v_2, \ldots, v_n$ of d points each. A *pairing* of these points is a perfect matching into $dn/2$ pairs. Given a pairing P, we may construct a multigraph $G(P)$, with loops allowed, as follows: the vertices are the buckets $v_1, v_2, \ldots, v_n$, and a pair $\{x, y\}$ in P corresponds to an edge $v_i v_j$ in $G(P)$ if x and y are contained in the buckets v_i and v_j, respectively. It is an easy fact that the probability of a random pairing corresponding to a given simple graph G is independent of the graph, hence the restriction of the probability space of random pairings to simple graphs is precisely $\mathcal{G}_{n,d}$. Moreover, it is well known that a random pairing generates a simple graph with probability asymptotic to $e^{-(d^2-1)/4}$ depending on d, so that any event holding a.a.s. over the probability space of random pairings also holds a.a.s. over the corresponding space $\mathcal{G}_{n,d}$. For this reason, asymptotic results over random pairings suffice for our purposes. For more information on this model, see, for example, the survey of Wormald [40].

5.2 Some ideas for an upper bound for random d-regular graphs

In this section, we discuss briefly a few ideas for improvement that, unfortunately, did not give any substantial improvement.

Idea 1: Recall that in order to get the current best upper bound we showed that a.a.s. no set of size xn induces more than $y_3(x,d)xn/2$ edges. As a result the largest value of $y_i/d - x_i$ in (4) is at most $U_3(d)$. For example, for $d = 3$ the optimal choice that maximizes $U_3(3)$ is: $x = \hat{x} \approx 0.0225$, $y = \hat{y} \approx 2.4789$, and so $U_3(3) \approx 0.8038$ as reported in Table 2.3. However, clearly it is impossible to partition a graph precisely into parts of size $\hat{x}n$. It is possible to show that the following upper bound holds, which is clearly not larger than the previous one:

$$U_4 = U_4(d) := \max_{k \in \mathbb{N} \setminus \{1\}} \left(\frac{y_3(1/k, d)}{d} - 1/k \right).$$

Unfortunately, this maximum value is achieved for $k = 45$ (which corresponds to parts of size roughly $0.0222n$, and no improvement is achieved: $U_4(3) \approx 0.8038$. The reason this idea fails is that the optimal value of $\hat{x}$ is small so that rounding to the nearest integer for k does not improve the bound much.

Idea 2: Let us look at (4) again but this time let us order the terms so that

$$\left(\frac{y_1}{d} - x_1\right) \geq \left(\frac{y_2}{d} - x_2\right) \geq \ldots \geq \left(\frac{y_k}{d} - x_k\right).$$

It follows that

$$q_{\mathcal{A}} = \sum_{i=1}^{k} x_i \left(\frac{y_i}{d} - x_i\right) \geq x_1 \left(\frac{y_1}{d} - x_1\right) + (1 - x_1)\left(\frac{y_2}{d} - x_2\right).$$

It is slightly more tedious than before, but one can get an improvement by considering (ordered) disjoint pairs of vertices X_1, X_2 with $|X_1| = x_1 n$, $|X_2| = x_2 n$, $e(X_1) = y_1 x_1 n/2$, $e(X_2) = y_2 x_2 n/2$, and $e(X_1, X_2) = zn$. Unfortunately, this idea also does not provide any reasonable improvement. For $d = 3$, the expected number of pairs of sets for the following vector $(x_1, x_2, y_1, y_2, z) = (\hat{x_1}, \hat{x_2}, \hat{y_1}, \hat{y_2}, \hat{z}) \approx (0.0239, 0.0225, 2.4830, 2.4790, 0.000037)$ and, again, no substantial improvement is achieved: $U_5(3) \approx 0.8038$.

Idea 3: As before, let us concentrate on the case $d = 3$ but similar ideas can be used for any integer $d \geq 3$. We can try to use the fact that $G_{n,3}$ can be constructed by putting a random matching on the vertices of a Hamiltonian cycle. Let us fix any set of vertices of size xn that induces zn components (paths) by restricting only to edges of the Hamiltonian cycle. Each such set can be represented by the following triple: vertex v, vector $(a_1 - 1, \ldots, a_{zn} - 1)$, and vector $(b_1 - 1, \ldots, b_{zn} - 1)$: v starts some path, a_i is the number of vertices on path i, b_i is the number of vertices not in the set and right after path i. The number of such sets is at most $n\binom{xn}{zn}\binom{(1-x)n}{zn}$. The number of edges within this set that are part of the Hamiltonian cycle is $xn - zn$. Hence, in order for the set to induce $yxn/2$ edges, $(yx/2 - x + z)n$ edges must be coming from the matching.

The hope is (that is, *was*) that for small values of z, there are only a few sets to consider. On the other hand, if z is closer to x, then less edges are "for free" (edges of the Hamiltonian cycle). Unfortunately, again this idea does not lead to any substantial improvement. Concentrating on $d = 3$, $\hat{x} = 0.0225$, $\hat{y} = 2.4789$, and tuning $\hat{z} \approx 0.00392$, the expected number of such sets is tending to infinity as $n \to \infty$.

Conclusion: The lack of improvement is disappointing but perhaps should not be surprising. Looking at one or two parts of a partition maximizing q^* is not enough (local property). Having one large term $y_i/d - x_i$ in (4) might be possible but having all of them to be large perhaps is not. So in order to improve the upper bound, one needs to consider all parts at the same time (global property).

5.3 Proof of Theorem 1

Proof. Consider $\mathcal{G}_{n,d}$ for some $d \in \mathbb{N} \setminus \{1, 2\}$ and let $\varepsilon > 0$ be an arbitrarily small constant. Our goal is to show that the expected number of sets S such that $|S| = xn$ and $e(S) = yxn/2$ with $y \geq y_3(x, d) + \varepsilon$ is $o(n^{-2})$. (For simplicity, we do not round numbers that are supposed to be integers either up or down; this is justified since these rounding errors are negligible to the asymptomatic calculations we will make.) This, together with the first moment principle, implies that a.a.s. no such set exists for any $x \in (0, 1)$ and $y \in [y_3(x, d) + \varepsilon, d]$ (as there are $O(n)$ possible sizes of S and $O(n)$ possible values of $e(S)$ that we need to consider).

The desired upper bound for the modularity will follow immediately by the argument discussed earlier.

Let $x = x(n)$ and $y = y(n)$ be any functions of n such that $0 < x < 1$ and $y_3(x,d) + \varepsilon < y < d$. Let $X(x,y)$ be the expected number of sets S such that $|S| = xn$ and $e(S) = yxn/2$. Using the paring model, it is clear that

$$X(x,y) = \binom{n}{xn}\binom{dxn}{yxn}\binom{d(1-x)n}{(d-y)xn}((d-y)xn)!M\left(yxn\right) \cdot M\left(d(1-x)n-(d-y)xn\right)/M(dn),$$

where $M(i)$ is the number of perfect matchings on i vertices, that is,

$$M(i) = \frac{i!}{(i/2)!2^{i/2}}.$$

(Each time we deal with perfect matchings, i is assumed to be an even number.) After simplification we get

$$X(x,y) = n!\,(dxn)!\,(d(1-x)n)!\,(yxn)!(dn/2)!\;2^{dn/2}\cdot \Bigg[(xn)!\,((1-x)n)!\,(yxn)!\,((d-y)xn)!\left(\frac{yx}{2}n\right)!\;2^{\frac{yx}{2}n}\left(\frac{d-2dx+yx}{2}n\right)!\;2^{\frac{d-2dx+yx}{2}n}(dn)!\Bigg]^{-1}.$$

Using Stirling's formula ($i! \sim \sqrt{2\pi i}(i/e)^i$) and focusing on the exponential part we obtain

$$X(x,y) = \Theta(n^{-1})e^{f(x,y,d)n},$$

where $f(x,y,d)$ is defined in (5). It follows immediately from the definition of $y_3(x,d)$ that $f(x,y,d) < 0$ for any pairs of integers xn and $yxn/2$ under consideration, and so for any pair we get $X(x,y) = o(n^{-2})$ and the proof is finished.

5.4 Proof of Theorem 2

The adjacency matrix $A = A(G)$ of a given a d-regular graph G with n vertices, is an $n \times n$ real and symmetric matrix. Thus, the matrix A has n real eigenvalues which we denote by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. It is known that certain properties of a d-regular graph are reflected in its spectrum but, since we focus on expansion properties, we are particularly interested in the following quantity: $\lambda = \lambda(G) = \max(|\lambda_2|, |\lambda_n|)$. In words, λ is the largest absolute value of an eigenvalue other than $\lambda_1 = d$ (for more details, see the general survey [24] about expanders, or [6], Chapter 9).

The value of λ for random d-regular graphs has been studied extensively. A major result due to Friedman [22] is the following:

Lemma 8 ([22]) *For every fixed $\varepsilon > 0$ and for $G \in \mathcal{G}_{n,d}$, a.a.s.*

$$\lambda(G) \leq 2\sqrt{d-1} + \varepsilon.$$

The number of edges $|E(S,T)|$ between sets S and T is expected to be close to the expected number of edges between S and T in a random graph of edge density d/n, namely $d|S||T|/n$. A small λ (or large spectral gap) implies that this deviation is small. The following useful bound is essentially proved in [4] (see also [6]):

Lemma 9 (Expander Mixing Lemma) *Let G be a d-regular graph with n vertices and set $\lambda = \lambda(G)$. Then for all $S, T \subseteq V$*

$$\left| |E(S,T)| - \frac{d|S||T|}{n} \right| \leq \lambda \sqrt{|S||T|}\,.$$

Note that $S \cap T$ does not have to be empty; in general, $|E(S,T)|$ is defined to be the number of edges between $S \setminus T$ to T plus twice the number of edges that contain only vertices of $S \cap T$. For our purpose here it is better to apply a slightly stronger lower estimate for $|E(S, V \setminus S)|$, namely,

$$|E(S, V \setminus S)| \geq \frac{(d-\lambda)|S||V \setminus S|}{n} \tag{8}$$

for all $S \subseteq V$. This is proved in [5], see also [6]. Using this inequality we get immediately that for any S of size xn we have

$$e(S) = \frac{d|S| - |E(S, V \setminus S)|}{2} \leq \frac{dxn - (d-\lambda)x(1-x)n}{2} = \frac{dx + \lambda(1-x)}{2}. \tag{9}$$

After combining (8), (4), Lemma 8, and Lemma 9 the desired result holds.

5.5 Proof of Theorem 3

Proof. Consider a forest F_n. For any subset A of the vertex set $V(F_n)$ let $\mathrm{vol}(A) = \sum_{v \in A} \deg(v)$. Here $\deg(v)$ is the degree of a vertex v in F_n. We also use the notation $\mathrm{vol}(F') := \mathrm{vol}(V(F'))$, where F' is a subgraph of F_n.

Let us take $h = \sqrt{n\Delta}$. The idea of the proof is to partition $V(F_n)$ into $A_1, \ldots, A_k$ such that for each i: $\mathrm{vol}(A_i) \leq h$ and a subgraph induced by A_i is a tree.

Our forest F_n may already contain trees $T_1, \ldots, T_\ell$ with $\mathrm{vol}(T_i) \leq h$. It remains to decompose the remaining trees into smaller ones. The way we do this decomposition is in a sense similar to the clustering algorithm *greedy-decompose*$_{\leq h}$ from [32]. Namely, we first redefine a notion of a centroid edge of a subtree T of the initial forest F_n. The removal of any edge of a tree T splits T into two parts T_1 and T_2. A centroid edge of T is an edge chosen to maximize $\min\{\mathrm{vol}(T_1), \mathrm{vol}(T_2)\}$. Our algorithm is then the following: as long as our forest contains a tree T with $\mathrm{vol}(T) > h$, it finds a centroid edge e of T and removes it. After this decomposition, we obtain additional trees $T_{\ell+1}, \ldots, T_k$ and we set $A_i = V(T_i)$ for $1 \leq i \leq k$. Note that for each i we have $\mathrm{vol}(A_i) \leq h$.

We will also need the following claim that is similar to Lemma 6 from [32].

Claim 1 $\mathrm{vol}(A_i) > \frac{h}{\Delta} - 1$ *for all* $l+1 \le i \le k$.

Proof (Proof of the Claim). Consider any step of our decomposition procedure. We take a tree T with $\mathrm{vol}(T) = h' > h$, remove its centroid edge e, and obtain two trees T_1 and T_2. Without loss of generality we may assume that $\mathrm{vol}(T_1) \le \mathrm{vol}(T_2)$. Let $s = \mathrm{vol}(T_1)$, $s \le h'/2$. Let x be the vertex incident with e and belonging to T_2. For every edge e' incident with x, for the part T' of $T - e'$ not containing x we have $\mathrm{vol}(T') \le s$ (otherwise e is not a centroid edge). As x has degree at most Δ, then $h' \le \Delta s + \Delta$ (at most s for each of the $\le \Delta$ parts plus the degree of x itself). So, $s \ge \frac{h'-\Delta}{\Delta} > \frac{h}{\Delta} - 1$ and the claim holds.

Now we have the partition $\mathcal{A} = \{A_1, \ldots, A_k\}$ of the vertex set $V(F_n)$. Our goal is to estimate $q_{\mathcal{A}}$. First, we estimate the edge contribution. According to Lemma 1, $\mathrm{vol}(A_i) \ge \frac{h}{\Delta} - 1$ for $\ell + 1 \le i \le k$. Also, as F_n is a forest $\sum_i \mathrm{vol}(A_i) = \mathrm{vol}(F_n) \le 2n - 2$. Therefore, $k - l \le (2n-2)/(\frac{h}{\Delta} - 1)$. Finally, using that $|E(F_n)| \ge \frac{n}{2}$, we get

$$\sum_{A \in \mathcal{A}} \frac{e(A)}{|E(F_n)|} \ge 1 - \frac{k-l-1}{|E(F_n)|} \ge 1 - \frac{4(n-1)}{n(\frac{h}{\Delta} - 1)} = 1 - O\left(\sqrt{\frac{\Delta}{n}}\right).$$

It remains to estimate the degree tax. Recall that $\mathrm{vol}(A_i) \le h$ for $1 \le i \le k$ and $\sum_i \mathrm{vol}(A_i) \le 2n - 2$. Therefore,

$$\sum_{A \in \mathcal{A}} \frac{\mathrm{vol}^2(A)}{4|E(F_n)|^2} \le \frac{h^2 \frac{n}{h}}{n^2} = O\left(\sqrt{\frac{\Delta}{n}}\right),$$

and so the proof is finished.

5.6 Proof of Theorem 4

Proof. As in the previous proof, we fix $h = \sqrt{\Delta n}$ and our goal is to partition $V(G_n)$ into $A_1, \ldots, A_k$ such that $\mathrm{vol}(A_i) \le h$ for each i. For a graph G_n let us consider a spanning tree T_n. For T_n we apply the decomposition procedure described in the proof of Theorem 3. In this case, we define $\mathrm{vol}(A) = \sum_{v \in A} \deg(v)$, where $\deg(v)$ is the degree of a vertex v in the *initial graph* G_n. The decomposition procedure gives us a partition $A_1, \ldots, A_k$.

In order to estimate $q_{\mathcal{A}}$ we first deal with the edge contribution. Similarly to Claim 1, we have $\mathrm{vol}(A_i) > \frac{h}{\Delta} - 1$ for all i. Also, $\sum_i \mathrm{vol}(A_i) = \mathrm{vol}(G_n) \le nD$. Therefore, $k \le nD/(\frac{h}{\Delta} - 1)$. The number of intracluster edges in T_n is $n - k$, and clearly this is the lower bound for $\sum_{A \in \mathcal{A}} e(A)$. Finally,

$$\sum_{A \in \mathcal{A}} \frac{e(A)}{|E(G_n)|} \ge \frac{n-k}{nD/2} \ge \frac{2}{D} - \frac{2}{\frac{h}{\Delta} - 1} = \frac{2}{D} - O\left(\sqrt{\frac{\Delta}{n}}\right).$$

It remains to estimate the degree tax. Recall that $\mathrm{vol}(A_i) \le h$ for all i and $\sum_i \mathrm{vol}(A_i) \le nD$. Therefore,

$$\sum_{A\in\mathcal{A}} \frac{\mathrm{vol}^2(A)}{4|E(F_n)|^2} \le \frac{h^2\frac{nD}{h}}{n^2D^2} = O\left(\sqrt{\frac{\Delta}{n}}\right),$$

and so the proof is finished.

5.7 Proof of Theorem 5

Proof. Let $\varepsilon > 0$. It is well-known that a.a.s. $\Delta(G_m^n) = O\left(n^{\frac{1}{2}+2\varepsilon}\right)$ (see, e.g., [17] and Theorem 17 in [12]). Also, clearly the average degree of G_m^n equals $2m$. In addition, for $m \ge 2$ a.a.s. G_m^n is connected [13]. So, the statement of Theorem 5 follows directly from Theorems 3 and 4.

5.8 Proof of Theorem 7

Proof. Let $\omega = \left\lceil n^{\max\{1/m,1-pA_1\}/2}\log n^{-1/2}\right\rceil$. Note that $\omega \ge n^{\varepsilon}$ for some $\varepsilon > 0$ that depends on the parameters of the model. Let us partition the space S into ω parts as follows: for each $1 \le r \le \omega$,

$$S_r = \left\{ s = (s_1,\ldots,s_m) \in S : \frac{r-1}{\omega} \le s_1 < \frac{r}{\omega} \right\}.$$

This partition of S naturally gives us a partition $\mathcal{A}$ of the vertex set: for each $1 \le r \le \omega$, $A_r = V_n \cap S_r$. We will show that a.a.s.

$$q_{\mathcal{A}}(\hat{G}_n) = 1 - O\left(n^{\max\{-1/m,-1+pA_1\}/2}\log^{9/2} n\right),$$

which will finish the proof as $q^*(\hat{G}_n) \ge q_{\mathcal{A}}(\hat{G}_n)$ and always $q^*(\hat{G}_n) \le 1$.

First, let us start with estimating the edge contribution. In order to do that, we need to estimate the number of edges between different parts. So, let us focus on any part A_r. We will investigate how many *bad* edges in G_n connect vertices outside of A_r with vertices inside A_r by counting (independently) bad edges directed to vertices of similar age. (Note that for convenience we consider here directed graph G_n instead of $\hat{G}_n$.) For a given integer k such that $0 \le k \le \lfloor \log n \rfloor$, let

$$\begin{aligned}
V^{(k)} &= \{v_i \in V_n : e^k \le i < \min\{e^{k+1}, n+1\}\},\\
E^{(k)} &= \{(v_j,v_i) \in E_n : v_i \in V^{(k)}, v_j \in V_n, \text{ and } i < j \le n\}\\
C^{(k)} &= \{(v_j,v_i) \in E_n : v_i \in V^{(k)} \cap A_r, v_j \in V_n \setminus A_r, \text{ and } i < j \le n\} \subseteq E^{(k)}.
\end{aligned}$$

It is clear that $\{V^{(k)} : 0 \le k \le \lfloor \log n \rfloor\}$ and $\{E^{(k)} : 0 \le k \le \lfloor \log n \rfloor\}$ are partitions of the vertex set and the edge set (both in $\hat{G}_n$ and G_n), respectively, and so $\{C^{(k)} : 0 \le k \le \lfloor \log n \rfloor\}$ is a partition of the bad edges we want to count. It remains to estimate the size of $C^{(k)}$ for a given value of k.

Fix $0 \le k \le \lfloor \log n \rfloor$, and let us concentrate on any $v_i \in V^{(k)}$. It follows from (6) that the maximum volume of a sphere of influence of v_i is $O(i^{-1}\log^2 n) = O(e^{-k}\log^2 n)$ (during the whole process) and so the maximum radius of influence of v_i is $O((e^{-k}\log^2 n)^{1/m})$. Therefore, if there is an edge in the cut directed to $v_i = (s_1, \ldots, s_m)$, then v_i must fall not only into A_r but also into a strip within distance $O((e^{-k}\log^2 n)^{1/m})$ from one of the two cutting hyperplanes separating A_r from neighbouring parts; that is, $|s_1 - \frac{r-1}{\omega}| = O((e^{-k}\log^2 n)^{1/m})$ or $|s_1 - \frac{r}{\omega}| = O((e^{-k}\log^2 n)^{1/m})$. Since $|V^{(k)}| = O(e^k)$, we get that

$$O((e^{-k}\log^2 n)^{1/m}) \cdot |V^{(k)}| = O(e^{k(1-1/m)}(\log n)^{2/m})$$

vertices of $V^{(k)}$ are expected to appear in these two strips during the whole process. Hence, it follows from Chernoff bound that with probability at least $1 - \exp(-\Theta(\log^2 n))$ there are $O(e^{k(1-1/m)}\log^2 n)$ vertices in these strips at the end of the process. Note that the exponent of $\log n$ has changed from $2/m$ to 2 in order to guarantee the claimed upper bound is at least $\log^2 n$ which is required for a bound to hold with the desired probability. Using (6) one more time, we get that all vertices introduced in this time period have (final) in-degree $O((n/e^k)^{pA_1}\log^2 n)$. Hence, there are

$$|C^{(k)}| = O\left((e^{k(1-1/m)}\log^2 n) \cdot (n/e^k)^{pA_1}\log^2 n\right) = O\left(n^{pA_1}e^{k(1-1/m-pA_1)}\log^4 n\right)$$

edges in the cut with probability at least $1-\exp(-\Theta(\log^2 n))$ and so this property holds a.a.s. for all parts A_r and all values of k. It follows that a.a.s. the number of bad edges involving A_r is at most

$$\begin{aligned}\xi_r &= \sum_{k=0}^{\lfloor \log n \rfloor} |C^{(k)}| = \sum_{k=0}^{\lfloor \log n \rfloor} O(n^{pA_1}e^{k(1-1/m-pA_1)}\log^4 n)\\ &\le \begin{cases} \log n \cdot O\left(n^{pA_1}n^{1-1/m-pA_1}\log^4 n\right), & \text{if } pA_1 < 1-1/m;\\ \log n \cdot O(n^{pA_1}\log^4 n), & \text{otherwise,}\end{cases}\\ &= O(n^{\max\{1-1/m,pA_1\}}\log^5 n).\end{aligned}$$

Finally, we get an estimate for the edge contribution: a.a.s.

$$\begin{aligned}\sum_{r=1}^{\omega}\frac{e(A_r)}{|E(G_n)|} &= 1 - \sum_{r=1}^{\omega}\frac{\xi_r}{|E(G_n)|} = 1 - \frac{\omega \cdot O(n^{\max\{1-1/m,pA_1\}}\log^5 n)}{\Theta(n)}\\ &= 1 - O\left(n^{\max\{-1/m,-1+pA_1\}/2}\log^{9/2} n\right).\end{aligned}$$

It remains to estimate the degree tax. In order to do that we need to, for a given r under consideration, estimate $\sum_{v\in A_r}\deg(v)$ in $\hat{G}_n$; that is, $\sum_{v\in A_r}(\deg^-(v)+\deg^+(v))$ in G_n. As before, we partition the vertices of A_r into sets containing vertices of similar age. Let k_0 be the largest integer k such that $(k-1)\omega\log^2 n < n$. Clearly, $k_0 = O(n/(\omega\log^2 n))$. This time, for a given integer k such that $1 \le k \le k_0$, let

$$V^{(k)} = \{v_i \in V_n : (k-1)\omega\log^2 n < i \le \min\{k\omega\log^2 n, n\}\},$$

and our goal is to estimate the size of $A_r \cap V^{(k)}$. The expected number of vertices of $V^{(k)}$ that fall into A_r is $|V^{(k)}|/\omega \le \log^2 n + 1$ and if follows from Chernoff's bound that with probability at least $1-\exp(-\Theta(\log^2 n))$ it is $O(\log^2 n)$. Using (6) for the last time, we get that all vertices introduced in this time period have (final) in-degree $O((n/(k\omega \log^2 n))^{pA_1} \log^2 n)$. It follows that with the desired probability

$$\begin{aligned}\sum_{v \in A_r} \deg^-(v) &= \sum_{k=1}^{k_0} O((n/(k\omega \log^2 n))^{pA_1} \log^2 n) \cdot O(\log^2 n) \\ &= O((n/(\omega \log^2 n))^{pA_1} \log^4 n) \cdot O(k_0^{1-pA_1}) = O(n \log^2 n/\omega),\end{aligned}$$

and so it holds a.a.s. for all r. Similarly, using Chernoff's bound and (7) we get that a.a.s. for all r we have $|A_r| \sim n/\omega$ and so

$$\sum_{v \in A_r} \deg^+(v) = O(n/\omega) \cdot O(\log^2 n) = O(n \log^2 n/\omega).$$

Finally, we are able to get an estimate for the degree tax in $\hat{G}_n$: a.a.s.

$$\begin{aligned}\sum_{r=1}^{\omega} \frac{(\sum_{v \in A_r} \deg(v))^2}{4|E(G_n)|^2} &= \frac{\omega \cdot O((n \log^2 n/\omega)^2)}{\Theta(n^2)} \\ &= O(\log^4 n/\omega) = O\left(n^{\max\{-1/m, -1+pA_1\}/2} \log^{9/2} n\right),\end{aligned}$$

and the proof is finished.