Journal of Complex Networks (2020) **4**, Advance Access Publication on 23 October 2019,
Issue Publication on 14 September 2020
doi: 10.1093/comnet/cnz040

Classes of preferential attachment and triangle preferential attachment models with power-law spectra

NICOLE EIKMEIER†
Department of Computer Science, Grinnell College, 1115 8th Avenue, Grinnell, IA 50112, USA
†Corresponding author. Email: eikmeier@grinnell.edu

AND

DAVID F. GLEICH
Department of Computer Science, Purdue University, 610 Purdue Mall, West Lafayette, IN 47907, USA

Edited by: Ernesto Estrada

[Received on 21 July 2019; editorial decision on 9 September 2019; accepted on 3 October 2019]

Preferential attachment (PA) models are a common class of graph models which have been used to explain why power-law distributions appear in the degree sequences of real network data. Among other properties of real-world networks, they commonly have non-trivial clustering coefficients due to an abundance of triangles as well as power laws in the eigenvalue spectra. Although there are triangle PA models and eigenvalue power laws in specific PA constructions, there are no results that existing constructions have both. In this article, we present a specific Triangle Generalized Preferential Attachment Model that, by construction, has non-trivial clustering. We further prove that this model has a power law in both the degree distribution and eigenvalue spectra.

Keywords: power law; preferential attachment; spectra.

1. Introduction

The idea of preferential attachment (PA) has a lengthy history in explaining 'rich-get-richer' models [1, 2]. In the context of networks, a PA model suggests that when agents join a network, they form links to existing nodes with large degrees. These models offer a simple local rule that helps explain the presence of highly skewed or power-law degree distributions in real-world networks [3]. While a simple and compelling mathematical model, there are weaknesses in the relationship between PA models and real-world data. One of the most striking is the lack of clustering in PA network models. Consequently, there has been a line of work on generalized PA models that include ways to address the lack of clustering. First, Holme and Kim [4] proposed a triangle PA model, where agents arrive and link to a node based on its degree and also link to a neighbour of that node to form a triangle. Later, Ostroumova *et al.* [5] generalized a family of PA models and showed that they had power-law degree distributions and in some cases high clustering. We note that clustering coefficients and degree distributions reflect intrinsically different properties as there are networks such as a barbell graph (extremely high clustering) and a bipartite network (no clustering) which both have the same degree distribution, but radically different clustering coefficients, this has also been observed in graph models [6].

Our work follows in this vein, and in this article, we analyse a specific set of PA networks. The specific Triangle Generalized Preferential Attachment (TGPA) model will be introduced formally in Section 3. Informally, the networks grow by inserting two edges that form an idealized triangle (with an already existing 3rd edge) at each step, where the endpoints are chosen via PA. Our novel analysis of the model shows that it will produce networks with power-law degree distributions (Theorem 6.1, Corollary 7.2), and power-law spectral distributions (Theorem 6.2). In our design and analysis of TGPA, we also proved that the models presented by Avin *et al.* [7] and the Holme–Kim model [4] have power laws in the spectra as well. The result on the Holme–Kim model we get for free from our proof of the power law on TGPA. We prove the result on the Avin *et al.* model separately, in a similar method as in TGPA.

Our interest in the TGPA model stems from our recent finding on the reliable presence of power laws in the eigenvalue spectrum of the adjacency matrix [8]. Specifically, in [8], we found that real-world networks of a variety of types were more likely to have a statistically significant power law in the eigenvalues of the adjacency matrix than in the degree distribution. The focus on the existence of power laws in graph models has been primarily focused on the degree distribution in the past. While standard PA models have been shown to have power laws in the eigenvalue spectra [9, 10], this model does not have significant clustering coefficients as already mentioned.

For a few decades, power laws, or highly skewed degree distributions, were commonly accepted properties of real-world networks. As the diversity of data studied continues to grow, alternative distributions that provide superior characterizations of degree distributions of real-world networks have emerged [11, 12]. Our point with this article is not the power law results necessarily, but rather the study of simple models for networks, and an understanding of what contributes to properties of these networks reminiscent of real-world data. That is, our goal is to understand driving factors underlying simple generative mechanisms for networks and to characterize the emergent properties of those models.

In summary, the primary contributions of this manuscript are:

1. We present the TGPA model: A model which imposes higher-order structure directly into the network (Section 3).
2. We show that TGPA produces graphs with significant clustering (Section 4).
3. We extend the results presented on the Generalized Preferential Attachment (GPA) model (in [7]), to show the eigenvalues follow a power-law distribution (Section 5).
4. We conduct extensive analysis of TGPA to show that the degrees follow a power-law distribution with an exponent which can range between $(1, \infty)$ (Section 7), and that the eigenvalues follow a power-law distribution (Section 6).
5. We extend the results presented on the model presented by Holme and Kim [4] to show the eigenvalues follow a power-law distribution (Section 6.1).

In our experience establishing these results, one technical challenge involved the ideas in Lemma 6.2, where the existing approach fundamentally assumed that edges were not added in small groups. The code to reproduce our experiments is available online at https://github.com/eikmeier/TGPA.

2. Preliminaries and related work

Denote a graph G by its set of vertices V and edges E. A graph with n vertices can be represented as an $n \times n$ adjacency matrix $\boldsymbol{A}$, where $\boldsymbol{A}_{ij} = 1$ if edge (i,j) is in the graph, and $\boldsymbol{A}_{ij} = 0$ otherwise. The

Table 1 *Notation in our article*

$G_t = (V_t, E_t)$	A graph at time step t with vertices V and edges E
$\boldsymbol{A}$	The adjacency matrix
$d_t(v)$	The degree of vertex v at time t
e_t	The number of edges at time t
$m_{k,t}$	The number of nodes at time t with degree k

degree of vertex i is the number of vertices j such that $\boldsymbol{A}_{ij} = 1$. In this manuscript, we will be primarily concerned with undirected, simple, graphs, meaning $\boldsymbol{A}_{ij} = \boldsymbol{A}_{ji}$ for all i, j, and $\boldsymbol{A}_{ii} = 0$ for all i.

We will be concerned with graph models that evolve over time. There are a huge diversity of graph generation schemes, many of which have been analysed in theory and in practice. For example, latent space graphs [13] and GPA [14]. Define a non-empty initial graph G_0 with vertices, V_0, and edges, E_0. At each time step $t = 1, 2, \ldots$ perform some action on G_{t-1} (such as adding new vertices or edges) to obtain $G_t = (V_t, E_t)$. Continue until the graph is sufficiently large. Denote the degree of vertex v at time t to be $d_t(v)$. Let e_t denotes the number of edges at time t, and let $m_{k,t}$ be the number of nodes at time t with degree k.

2.1 *Preferential attachment*

PA describes a mechanism of graph evolution in which nodes with higher degree tend to continue gaining neighbours. When a new node u is added to the graph at time t, choose another existing vertex v with probability proportional to its degree. Formally, choose vertex v with probability

$$\gamma_t(v) = \frac{d_{t-1}(v)}{\sum_{w \in V_{t-1}} d_{t-1}(w)}. \tag{2.1}$$

Then add an edge connecting u to v. PA is meant to model the *power-law* behaviour that is often seen in real-world networks [15–17], that is a few vertices tend to have very large degree while most vertices have fairly low degree. A set of values $x_1, x_2, \ldots x_k$ satisfies a power law if it is drawn from a probability distribution $p(x) \propto x^{-\beta}$ for some β.

The PA graph model is found in a few different forms. In the model by Barabási and Albert [3], often called the BA model, at every new time step a new vertex is formed with m edges. Each of the edges is then connected to an existing node chosen using PA, that is based on their degrees.

In a slight variation found in [18, 19], at each time step t, a new node is added with probability p. Along with the new node is an edge between the new node and an existing node picked via PA. With probability $1 - p$ a new edge is added between two existing nodes, both chosen via PA. These two models generate slightly different distributions, but fundamentally give very similar graphs. We present our model TGPA in two forms matching these differences (Section 3).

In Sections 2.2 and 2.3, we discuss a few variations of the PA model. There exist other variations of PA [5, 20–22] which we will not detail here.

2.2 *Generalized preferential attachment*

The GPA model was defined by Avin *et al.* [7]. In this model, in addition to adding new vertices and edges, there is also an option in each time step of adding a new *component*. Furthermore, the parameters may change over time, if desired. Start with an arbitrary initial non-empty graph G_0. For time $t \geq 1$, the graph G_t is constructed by performing either a *node event* with probability $p_t \in [0, 1]$, an *edge event* with probability $r_t \in [0, 1 - p_t]$, or a *component event* with probability $q_t = 1 - p_t - r_t$. In a node event, a new vertex v is added to the graph, along with an edge (u, v) where u is chosen from G_{t-1} with probability $\gamma_t(u)$. In an edge event, a single new edge (u, w) is added, with u and w both nodes in G_{t-1}. The edge endpoint u is chosen with probability $\gamma_t(u)$ and the endpoint w is chosen with probability $\gamma_t(w)$, so the edge itself is chosen with probability $\gamma_t(u) \cdot \gamma_t(w)$, which gives a full probability distribution over edges (recall that multiedges and self-loops are allowed). And in a component event, two new nodes v_1, v_2 are added along with edge (v_1, v_2). Exactly one edge is added at each time step, so the number of edges in G_t is equal to $e_0 + t$.

The key difference in GPA over the PA model discussed in Section 2.1 is the ability to add new components to the graph. In [7], it is proved that the degree distribution follows a power law. In this manuscript, we further prove that the eigenvalues follow a power-law distribution (see Section 5).

We will also work with a slight variation of the GPA model, along the lines of the alternate version of the PA model defined in [3, 9] and discussed in Section 2.1. Start with an empty graph. Note: this model need not start as an empty graph, we just follow the convention of [3, 9]. Since all of our analysis is in asymptotics, it should have no impact. At time $t = 1, 2, \ldots$ do one of the following: with probability p add a new vertex v_t and an edge from v_t to some other vertex in u where u is chosen with probability

$$Pr[u = v_i] = \begin{cases} \frac{d_t(v_i)}{2t-1}, & \text{if } v_i \neq v_t \\ \frac{1}{2t-1}, & \text{if } v_i = v_t \end{cases}; \tag{2.2}$$

and with probability $1 - p$ add two new vertices and an edge between them. For some constant m, every m steps contract the most recent m vertices added through the PA step to form a super vertex. Notice that Equation (2.2) is not quite the same as γ_t in Equation (2.1). Equation (2.2) allows for nodes to be added with self-loops. In both versions, loops are allowed in the edge step. Regardless, the allowance of self-loops has little effect as the graph becomes large, and we remove all self-loops in our final graph for experimental analysis.

2.3 *Triad formation*

Holme and Kim [4] introduced a Triad Formation step into the BA version of the PA model (see Section 2.1). After each PA step in which a new vertex v is added and some edge is added (v, u), a triangle is closed with probability p_t by choosing a neighbour of u, u_2, and adding edge (v, u_2). An example network is shown in Figure 1 under ‘Holme’. The average number of triad closures per added vertex is $m_t = (m - 1)p_t$. It is shown in [4] that the network follows a power law in the degrees with an exponent of 3 and has clustering coefficients which can be tuned by the parameter m_t. Our model incorporates something very similar to this triad formation, but with less regular structure due to an added component step, and with a larger range of possible power-law exponents. See Section 3 for the description of our model, and Lemma 7.2 for the result on the range of degree power-law exponents.

2.4 *Higher-order features in graphs*

Recently, there has been interest in analysing the higher-order features in graphs [23–28]. One of the earlier motivations for this direction is the famous paper by Milo *et al.* [29] on the presence of motifs in real-world networks. Likewise, there are new models which aim to match these higher-order features. For example, the triad formation model described in Section 2.3 [4], and the family of PA models [5] discussed in Section 1. Another model, HyperKron, places a distribution over hyperedges and inserts motifs instead of edges [30] and is specifically shown to have higher-order clustering. The study of higher-order features is growing and complex, it is also studied in multi-layer networks, temporal data, and components of graphs. See [31] for a nice overview of work in higher-order network analysis.

3. TGPA

In this section, we present our model which we call TGPA. This model is motivated by the purpose of adding higher-order structure into the resulting graph as discussed in Section 2.4, and a recent paper [7] which shows a model of PA with any power-law exponent (Section 2.2). We present two different versions of the model. The first, in Section 3.1 follows the PA model as described by [3, 9], and the second in Section 3.2 follows the PA model as described in [7, 18]. Though these models are not the same, they share similar properties. In Sections 6 and 7, we will see each formulation is useful for the analysis of the models. Figure 1 shows some example graphs generated by TGPA compared to existing models.

3.1 *TGPA(p, q)*

Start with an empty graph. At time $t = 1, 2, \ldots$ do one of the following:

1. (node event) With probability p, add a new vertex v_t, and an edge from v_t to some other vertex u where u is chosen with probability

$$Pr[u = v_i] = \begin{cases} \frac{d_t(v_i)}{4t-2}, & \text{if } v_i \neq v_t \\ \frac{2}{4t-2}, & \text{if } v_i = v_t \end{cases}. \tag{3.1}$$

Then pick a neighbour of u, call it w, and also add an edge from v_t to w. We pick w with probability

$$Pr[w = v_i] = \begin{cases} \frac{\text{\# edges between } u,w}{d_{t-1}(u)}, & \text{if } v_i \neq u \\ \frac{2 \cdot \text{\# self-loops of } u}{d_{t-1}(u)}, & \text{if } v_i = u \end{cases}. \tag{3.2}$$

2. (component event) With probability $q = 1 - p$ add a wedge to the graph (3 new nodes with 2 edges).
3. For some constant m, every m steps contract the most recently added vertices through the PA steps (in step 1) to form a super vertex.

Note that vertex w (chosen in step 1) is *also* chosen via PA. The probability of picking w is the probability of picking u as a neighbour of w times the probability of picking w:

$$Pr[w = v_i] = \frac{\sum_{u \in N(w)} d_{t-1}(u)}{4t-2} \cdot \frac{\text{num edges between } u, w}{d_{t-1}(u)} = \frac{d_t(w)}{4t-2}.$$

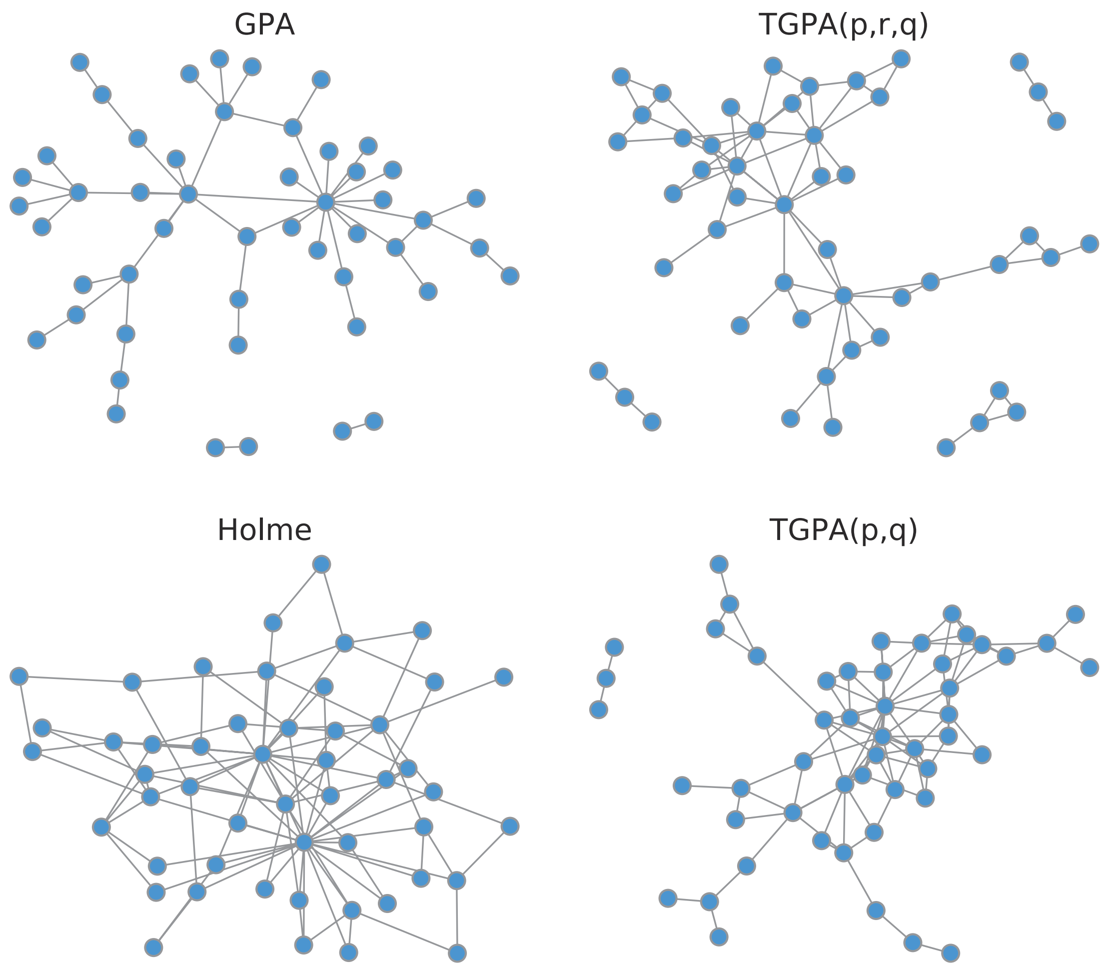

FIG. 1. Examples of 50 node graphs. The top two figures were generated using $p = 0.8, r = 0.1, q = 0.1$. The graphs on the bottom were generated using $m = 2$, and TGPA(p, q) used $p = 0.85$. See the text for the details on these parameters.

3.2 *TGPA*(p_t, r_t, q_t)

Start with a graph with e_0 edges. At time $t = 1, 2, \ldots$ do one of the following:

1. (Node event) With probability p_t, add a new vertex v_t, and an edge from v_t to some other vertex u where u is chosen with probability given in Equation (3.1). Then pick a neighbour of u, call it w, as in Equation (3.2). Add edge an edge from v_t to w.
2. (Wedge event) With probability r_t add a wedge to the graph by picking two nodes using PA: v_1, v_2. Pick the third node uniformly from a neighbour of v_1, call it w. Add edges (v_1, v_2) and (v_1, w).
3. (Component event) With probability q_t add a wedge to the graph (3 new nodes with 2 edges).

TABLE 2 *Clustering coefficients for three real-world networks, and generated models. TGPA is able to generate data with much larger clustering coefficients, compared to GPA*

Network name	Edges	Global clust	Local clust	HO global	HO local
Auburn (18k vertices)	974k	0.137	0.223	0.107	0.172
TGPA(18k,0.987,10,150):	640k	0.25	0.22	0.118	0.03
GPA(18k,0.001,0.999,2):	906k	0.021	0.030	0.005	0.014
Berkeley (13k vertices)	852k	0.114	0.207	0.0876	0.156
TGPA(13k,0.99, 10, 58)	502k	0.104	0.185	0.034	0.025
GPA(13k,0.001,0.999,2)	502k	0.024	0.034	0.005	0.015
Princeton (7k vertices)	293k	0.237	0.164	0.091	0.146
TGPA(7k,0.987,10,100):	207k	0.298	0.251	0.148	0.053
GPA(7k,0.001,0.999,2):	255k	0.038	0.054	0.009	0.025

4. Significant clustering coefficients

We analysed three networks from the Facebook 100 dataset [32], each of which is a set of users at a particular university. We computed the global clustering coefficient: $6|K_3|/|W|$, where $|K_3|$ is the number of triangles and $|W|$ is the number of wedges, and average local clustering coefficient: the average of $2|K_3(u)|/|W(u)|$ for all nodes u, where $K_3(u)$ denotes triangles for which u is a member. We also considered *higher-order* clustering coefficients, defined in [28] to be the fraction of appropriate motifs which are closed into 4-cliques.

To fit the TGPA(p,q) model (Section 3.1) to the real-world networks, we noted that the average degree of our model, the total degrees divided by the number of nodes, is approximately $(2m(1-p)+2m)/(m(1-p)+1)$. Choosing the average degree gives a relationship between parameters m and p. We tested various sets of parameters to obtain the best possible fit. We started both TGPA and GPA with a k-node clique. Table 2 lists the parameters we chose for the TGPA model as TGPA(n,p,k,m), which produces an n node graph starting from a k node clique. For comparison, we also fit the GPA model (Section 2.2). The parameters in Table 2 are GPA(n,p,r,k). Notice that TGPA maintains much more significant clustering coefficients across all measures.

5. Eigenvalue power law in GPA

In this section, we present results for the GPA model presented in [7] and discussed in Section 2.2, relating to the distribution of the eigenvalues of a graph formed in the model. Note that in order to get our desired result (Theorem 5.2), we also prove that the degree distribution has a power-law distribution (Theorem 5.1). This was already proven in [7], but the version of our proof is useful in order to obtain Theorem 5.2. The results and proofs mirror those in [9], but provide a useful step towards the results on the TGPA model in Section 6. All proofs are included in Appendix A.

Fix parameter p. Denote G_t^m as the GPA Graph at time t with contractions of size m.

LEMMA 5.1 Let $d_t(s)$ be the degree of vertex s in G_t^m, for any time t after s has been added to the graph. Let $a^{(k)} = a(a+1)(a+2)\cdots(a+k-1)$ be the rising factorial function. Let s' be the time at which node s arrives in the graph. Then for any positive integer k,

$$\mathbb{E}[(d_t(s))^{(k)}] \leq (2m)^{(k)} 2^{pk/2} \left(\frac{t}{s'}\right)^{pk/2}.$$

Now define a *supernode* to be a collection of nodes viewed as one. The degree of a supernode is the sum of the degrees of the vertices in the supernode.

LEMMA 5.2 Let $S = (S_1, S_2, \ldots, S_l)$ be a disjoint collection of supernodes at time t_0. Assume that the degree of S_i at time t_0 is $d_{t_0}(S_i) = d_i$. Let t be a time later than t_0. Let $p_S(\mathrm{r}; \mathrm{d}, t_0, t)$ be the probability that each supernode S_i has degree $r_i + d_i$ at time t. Let $d = \sum_{i=1}^{l} d_i, r = \sum_{i=1}^{l} r_i$. If $d = o(t^{1/2})$ and $r = o(t^{2/3})$, then

$$p_S(\mathrm{r}; \mathrm{d}, t_0, t) \leq \left(\prod_{i=1}^{l} \binom{r_i + d_i - 1}{d_i - 1}\right) \left(\frac{t_0 + 1}{t}\right)^{pd/2} \exp\left\{2 + t_0 - \frac{pd}{2} + \frac{3pr}{t^{p/2}}\right\}.$$

THEOREM 5.1 Let m, k be fixed positive integers, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_k$ denote the degrees of the k highest degree vertices of G_t^m. Then

$$\frac{t^{p/2}}{f(t)} \leq \Delta_1 \leq t^{p/2} f(t) \ \text{ and } \ \frac{t^{p/2}}{f(t)} \leq \Delta_i \leq \Delta_{i-1} - t^{p/2} f(t)$$

for $i = 1, 2, \ldots, k$ whp.

The factor of $t^{p/2}$ in Theorem 5.1 implies a power-law distribution in the largest degrees with exponent $\beta = (2+p)/p$. This can be seen by using a martingale argument, as described in [33] for instance. Notice that depending on the value chosen for p, we can obtain a power law fit with exponents ranging between 3 and ∞.

The next result relates maximum eigenvalues and maximal degrees in the GPA model. It is similar to results found in [9, 10, 34, 35]. It says that if the degrees follow a power law with exponent β, then the spectra follows a power law as well, with exponent $2\beta - 1$.

THEOREM 5.2 Let k be a fixed integer, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ be the k largest eigenvalues of the adjacency matrix of G_t^m. The for $i = 1, \ldots, k$, we have $\lambda_i = (1 + o(1))\Delta_i^{1/2}$, where Δ_i is the ith largest degree.

6. Analysis of TGPA(p, q)

In this section, we present results on the degrees and spectra of the TGPA(p, q) model, which was defined in Section 3.1. The proofs follow the proof techniques presented in Section 5. The key difference in these proofs is the fact that two edges may be added in each time step. This makes the PA much more tedious to track through graph generation. In Lemma 6.2 for example, we consider disjoint (but not *disconnected*) sets of supernodes; the probability of the supernodes increasing in degree is not independent from one other. The full proofs are in Appendix B.

Fix parameter p. Denote $G_t^m(p, q)$ as the TGPA Graph at time t with contractions of size m.

LEMMA 6.1 Let $d_t(s)$ be the degree of vertex s in $G_t^m(p,q)$, for any time t after s has been added to the graph. Let $a^{(\bar{k})} = a(a+2)(a+4)\cdots(a+k-2)$ be a modified rising factorial function. Let s' be the time at which node s arrives in the graph. Then for any positive integer k,

$$\mathbb{E}[(d_t(s))^{(\bar{k})}] \leq (4m)^{(\bar{k})} 2^{pk} \left(\frac{t}{s'}\right)^{pk}.$$

LEMMA 6.2 Let $S = (S_1, S_2, \ldots, S_l)$ be a disjoint collection of supernodes at time t_0. Assume that the degree of S_i at time t_0 is $d_{t_0}(S_i) = d_i$. Let t be a time later than t_0. Let $p_S(\mathrm{r}; \mathrm{d}, t_0, t)$ be the probability that each supernode S_i has degree $r_i + d_i$ at time t. Let $d = \sum_{i=1}^{l} d_i, r = \sum_{i=1}^{l} r_i$. If $d = o(t^{1/2})$ and $r = o(t^{2/3})$, then

$$p_S(\mathrm{r}; \mathrm{d}, t_0, t) \leq \left(\prod_{i=1}^{l} \binom{r_i + d_i - 1}{d_i - 1}\right) \left(\frac{t_0}{t-1}\right)^{pd} \exp\left\{3 + 2t_0 - pd + \frac{19pr}{4t^p}\right\}.$$

THEOREM 6.1 Let m, k be fixed positive integers, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_k$ denote the degrees of the k highest degree vertices of $G_t^m(p,q)$. Then,

$$\frac{t^p}{f(t)} \leq \Delta_1 \leq t^p f(t) \text{ and } \frac{t^p}{f(t)} \leq \Delta_i \leq \Delta_{i-1} - t^p f(t)$$

for $i = 1, 2, \ldots, k$ whp.

The factor of t^p in Theorem 6.1 implies a power-law distribution with exponent $\alpha = (1+p)/p$. This can be seen by using a martingale argument, which has been done a number of times. See for instance [33]. Notice that depending on the value chosen for p, we can obtain a power law fit with exponents ranging between 2 and ∞.

THEOREM 6.2 Let k be a fixed integer, and let $f(t)$ be a function with $f(t) \to \infty$ as $t \to \infty$. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ be the k largest eigenvalues of the adjacency matrix of $G_t^m(p,q)$. The for $i = 1, \ldots, k$, we have $\lambda_i = (1 + o(1))\Delta_i^{1/2}$, where Δ_i is the ith largest degree.

In the analysis of this proof, we restrict p to be greater than 9/32. This comes in Claim B.7, and constrains the exponent in the power law to be between 2 and 5. For the proof, we will show that with high probability G contains a star forest F, with stars of degree asymptotic to the maximum degree vertices of G. Then we show that $G \backslash F$ has small eigenvalues. Finally, we can use Rayleigh's principle to say that the large eigenvalues of G cannot be too different than the large eigenvalues of F. See Appendix B for the full proof.

6.1 *Spectral power law in Holme–Kim model*

The model which we have defined as TGPA(p,q) is equivalent to the model defined by Holme and Kim [4] with $q = 0$. We note here that then we get for free that the model by Holme and Kim also has a power law in the spectrum, by using the same analysis with $q = 0$ and $p = 1$.

7. Analysis of TGPA(p_t, r_t, q_t)

Consider *TGPA*(p_t, r_t, q_t), which was described in Section 3.2. The parameters p_t, r_t, q_t can change over time, though we will restrict the ways in which the parameters can evolve in Section 7.2.

7.1 *Recursive relation for* $m_{k,t}$

Recall that $m_{k,t}$ is the number of nodes at time t with degree k. We wish to write down a relationship for $m_{k,t+1}$ in terms of $m_{k',t}$ for $k' \leq k$. Recall also that the number of edges at time t is $e_t = e_0 + 2t$, and the total sum of degrees at any time t is $2e_t$. Note that for this reason, we need only focus on $m_{k,t}$ for $1 \leq k \leq 2e_t$.

Let $\mathscr{F}_t$ denote the σ-algebra generated by the graphs $G_0, G_1, \ldots, G_t$ ($\mathscr{F}_t$ holds the history of events up until time t). We wish to calculate the expected value of $m_{k,t+1}$, conditioned on the previous time steps, held in $\mathscr{F}_t$. Fix $k \geq 2$. Since $0 \leq d_{t+1}(v) - d_t(v) \leq 4$ for every node v and time t, we have

$$\mathbb{E}[m_{k,t+1}|\mathscr{F}] = \sum_{\{v:k-4\leq d_t(v)\leq k\}} \mathbb{P}[d_{t+1}(v) = k]. \tag{7.1}$$

Recall $\gamma_t(v)$ from Equation (2.1). Denote $\theta_t(v)$ as 2 times the number of self-loops in which v is involved divided by $\sum_{w\in V_{t-1}} d_{t-1}(w)$. (i.e. the proportion of edges which are self-loops on v). If $d_{t+1}(v) = 4$, then there are at most 5 possible values for $d_t(v)$ when $k \geq 4$:

(i) $d_t(v) = k$. In this case, there must have either been a node event not involving v (this occurs with probability $p_{t+1}(1 - 2\gamma_{t+1}(v) + \theta_{t+1}(v))$), or a wedge event not involving v (with probability $r_{t+1}(1 - \gamma_{t+1}(v))(1 - 2\gamma_{t+1}(v) + \theta_{t+1}(v))$), or a component event (with probability q_{r+1}).

(ii) $d_t(v) = k - 1$. In this case, there must have either been a node event where v is involved as the first node (probability $p_{t+1} \cdot \gamma_{t+1}(v) \cdot (1 - \theta_{t+1}(v))$), or where v is involved as the second node (probability $p_{t+1}(\gamma_{t+1}(v) - \theta_{t+1}(v))$), or a wedge event in which v is involved as the first node (with probability $r_{t+1}(\gamma_{t+1}(v) - \gamma_{t+1}(v)^2 - \theta_{t+1}(v) + \gamma_{t+1}(v) \cdot \theta_{t+1}(v))$) or as the third node (probability $r_{t+1}(1 - \gamma_{t+1}(v))(\gamma_{t+1}(v) - \theta_{t+1}(v))$).

(iii) $d_t(v) = k - 2$. In this case, there must have either been a node event in which v is picked as both nodes involved (with probability $p_{t+1} \cdot \theta_{t+1}(v)$) or there must have been a wedge event in which v is involved as the second node (with probability $r_{t+1} \cdot \theta_{t+1}(v)(1 - \gamma_{t+1}(v))$) or as the first and third nodes (with probability $r_{t+1} \cdot \gamma_{t+1}(v)(1 - \gamma_{t+1}(v) + \theta_{t+1}(v))$).

(iv) $d_t(v) = k - 3$. In this case, there must have been a wedge event where v was involved as the first and second nodes or there was a wedge event where v was involved as the second and third nodes (these events occur in combination with probability $2r_{t+1}\gamma_{t+1}(v)(\gamma_{t+1}(v) - \theta_{t+1}(v)\ )$).

(v) $d_t(v) = k - 4$. In this case, there must have been a wedge event where v is picked for all three wedges, which happens with probability $r_{t+1} \cdot \gamma_{t+1}(v) \cdot \theta_{t+1}(v)$

Let $\alpha_{k,t} = k/(2e_t)$. Then for every v such that $d_t(v) = i$, $\gamma_{t+1}(v) = \alpha_{i,t}$. In order to ease the exposition, and for short hand, we define values A, B, C, D, E for each of the probabilities calculated above:

$$A_{k,t} = p_{t+1,k}(1 - 2\alpha_{k,t} + \theta_{t+1}(v)) + r_{t+1}(1 - \alpha_{k,t})(1 - 2\alpha_{k,t} + \theta_{t+1}(v)) + q_{t+1},$$
$$B_{k,t} = 2p_{t+1}(\alpha_{k,t} - \theta_{t+1}(v)) + 2r_{t+1}(1 - \alpha_{k,t})(\alpha_{k,t} - \theta_{t+1}(v)),$$

$$C_{k,t} = p_{t+1}\theta_{t+1}(v) + r_{t+1}(\alpha_{k,t} - \alpha_{k,t}^2 + \theta_{t+1}(v)),$$
$$D_{k,t} = 2r_{t+1}\alpha_{k,t}(\alpha_{k,t} - \theta_{t+1}(v)), \text{ and } E_{k,t} = r_{t+1}\alpha_{k,t}\theta_{t+1}(v).$$

Then $A_{k,t} + B_{k,t} + C_{k,t} + D_{k,t} + E_{k,t} = 1$ and $A_{k,t}, B_{k,t}, C_{k,t}, D_{k,t}, E_{k,t} \geq 0$ for every $0 \leq k \leq 2e_t$. Also, by Equation (7.1), for every $k \geq 4$

$$\mathbb{E}[m_{k,t+1}|\mathscr{F}] = m_{k,t}A_{k,t} + m_{k-1,t}B_{k-1,t} + m_{k-2,t}C_{k-2,t} + m_{k-3,t}D_{k-3,t} + m_{k-4,t}E_{k-4,t}. \tag{7.2}$$

And for remaining values of k, we have

$$\begin{aligned} \mathbb{E}[m_{3,t+1}|\mathscr{F}] &= m_{3,t}A_{3,t} + m_{2,t}B_{2,t} + m_{1,t}C_{1,t} \\ \mathbb{E}[m_{2,t+1}|\mathscr{F}] &= m_{2,t}A_{2,t} + m_{1,t}B_{1,t} + p_{t+1} + q_{t+1}. \\ \mathbb{E}[m_{1,t+1}|\mathscr{F}] &= m_{1,t}A_{1,t} + 2q_{t+1} \end{aligned} \tag{7.3}$$

Define

$$X_{k,t} = \begin{cases} m_{k-1,t}B_{k-1,t} + m_{k-2,t}C_{k-2,t} + m_{k-3,t}D_{k-3,t} + m_{k-4,t}E_{k-4,t} & k \geq 4 \\ m_{2,t}B_{2,t} + m_{1,t}C_{1,t} & k = 3 \\ m_{1,t}B_{1,t} + p_{t+1} + q_{t+1} & k = 2 \\ 2q_{t+1} & k = 1 \end{cases}. \tag{7.4}$$

Then, Equations (7.2) and (7.3) can be re-written as

$$\mathbb{E}[m_{k,t+1}] = \mathbb{E}[m_{k,t}] \cdot A_{k,t} + \mathbb{E}[X_{k,t}]. \tag{7.5}$$

7.2 *Degree power law in TGPA*

The following lemma is presented in [7] and is a quick generalization of a result in [18].

LEMMA 7.1 ([7]) Suppose that a sequence satisfies the recurrence relation $a_{t+1} = (1 - b_t/(t + t_1))a_t + c_t$ for $t \geq t_0$. Furthermore, let $\{s_t\}$ be a sequence of real numbers with $\lim_{t\to\infty} s_t/s_{t+1} = 1$, $d_t = t(1 - s_t/s_{t+1})$, $\lim_{t\to\infty} b_t = b$, $\lim_{t\to\infty} c_t \cdot t/s_t = c$, $\lim_{t\to\infty} d_t = d$ and $b + d > 1$. Then $\lim_{t\to\infty} a_t/s_t$ exists and $\lim_{t\to\infty} a_t/s_t = c/(b + d)$.

The following theorem and corollary prove that $TGPA(p_t, r_t, q_t)$ has a power law in the degree distribution, which we can analyse.

THEOREM 7.1 Consider TGPA(p_t, r_t, q_t). Let $y_t = p_t + 3q_t$. Assume that $\lim_{t\to\infty} y_t = y < 3$, $\sum_{t=1}^{\infty} y_t = \infty$ and $\lim_{t\to\infty} t \cdot y_{t+1}/\sum_{j=1}^{t} y_j = \Gamma > 0$. Then letting $\beta = 1 + 2\Gamma/(3 - y)$, the limit $M_k = \lim_{t\to\infty} \mathbb{E}[m_{k,t}]/\mathbb{E}[n_t]$ exists for every $k \geq 1$ and

$$M_k = \frac{\Gamma}{\Gamma + 3/2 - y/2} \prod_{j=1}^{k-1} \frac{j}{j + \beta}. \tag{7.6}$$

Proof. This proof will be an induction on k. For $k = 1$, we use Lemma 7.1 setting $(t_1, s_t, a_t, b_t, c_t) = (e_0, \mathbb{E}[n_t], \mathbb{E}[m_{1,t}], e_t(1 - A_{1,t}), y_{t+1})$. Using Equation (7.5), this gives the limits $b = 3/2 - y/2$ and $c = d = \Gamma$, which concludes the base case. Now assume the theorem holds for $k - 1$, we now prove it for k. Again use Lemma 7.1, this time with $(t_1, s_t, a_t, b_t, c_t) = (e_0, \mathbb{E}[n_t], \mathbb{E}[m_{k,t}], B_{k-1,t}\,\mathbb{E}[m_{k-1,t}] + C_{k-1,t}\,\mathbb{E}[m_{k-2,t}] + D_{k-3,t}\,\mathbb{E}[m_{k-3,t}] + E_{k-4,t}\,\mathbb{E}[m_{k-4,t}])$. Then we get $d = \Gamma$, $b = k \cdot (3/2 - y/2)$, and using the inductive hypothesis,

$$c = \lim_{t\to\infty} \frac{c_t \cdot t}{s_t} = (k-1)\left(\frac{3}{2} - \frac{y}{2}\right) M_{k-1}.$$

Therefore, M_k exists and

$$M_k = \frac{(k-1)(3/2 - y/2)M_{k-1}}{k(3/2 - y/2) + \Gamma} = \frac{k-1}{k-1+\beta} M_{k-1}.$$

□

The proof of the following corollary follows exactly from [7]. We include the proof here for completeness.

COROLLARY 7.1 Under the assumptions in Theorem 7.1, M_k is proportional to $k^{-\beta}$.

Proof. Consider Equation (7.6). It is a fact that a differentiable function f is convex if and only if $f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$ for every x_1 and x_2 [36]. Applying this with $(f(z), x_1, x_2) = (z^\beta, 1, 1 + 1/j)$, we get

$$\prod_{j=1}^{k-1} \frac{j}{j+\beta} = \prod_{j=1}^{k-1} (1 + \beta/j)^{-1} \geq \prod_{j=1}^{k-1} (1 + 1/j)^{-\beta} = k^{-\beta}. \tag{7.7}$$

Also though, applying this with $(f(z), x_1, x_2) = (z^\beta, 1, 1 - 1/j)$ we obtain

$$\prod_{j=1}^{k-1} \frac{j}{j+\beta} = \prod_{j=1}^{k-1} (1 - \beta/(j+\beta))^{-1} \leq \prod_{j=1}^{k-1} (1 - 1/(j+\beta))^{-\beta} = \left(\frac{\beta}{k-1+\beta}\right)^{\beta} \leq \beta^\beta \cdot k^{-\beta}. \tag{7.8}$$

Therefore, $c_1 k^{-beta} \leq M_k \leq c_2 k^{-beta}$ for some positive constants c_1, c_2 and every $k \geq 1$. □

Finally, we can state which power-law exponents are obtainable.

LEMMA 7.2 For any $x \in (1, \infty)$, there exists a choice of p_t, r_t, q_t such that in $TGPA(p_t, r_t, q_t)$ the resulting network follows a power law in the degree distribution with exponent $\beta = x$.

Proof. We can use three separate cases:

(i) For $x \in (5/3, \infty)$, setting $y_t = 3 - 2/(x-1)$ gives exponent $\beta = 1 + 2/(3 - (3 - 2/(x-1))) = x$.

(ii) For $x \in (1, 5/3)$, set $y_t = t^{3/2(x-5/3)}$. Then

$$\Gamma = \lim_{t\to\infty} \frac{y_{t+1} \cdot t}{\sum_{j=1}^{t} y_j} = \lim_{t\to\infty} \frac{(t+1)^{3/2(x-5/3)} \cdot t}{\sum_{j=1}^{t} (j^{3/2(x-5/3)})} = \lim_{t\to\infty} \frac{\cdot t^{3/2x-3/2}}{\int_{j=0}^{t} j^{3/2(x-5/3)} dj},$$
$$= \lim_{t\to\infty} \frac{(3/2x-3/2)t^{3/2x-3/2}}{j^{3/2x-3/2}|_{j=0}^{t}} = 3/2x - 3/2$$

and $\beta = 1 + (2\Gamma)/(3-y) = 1 + 2(3/2x - 3/2)/(3-0) = x$.

(iii) For $x = 5/3$, set $y_t = 1/\ln(t+2)$ for every t. Then we have

$$\Gamma = \lim_{t\to\infty} \frac{y_{t+1} \cdot t}{\sum_{j=1}^{t} y_j} = \lim_{t\to\infty} \frac{t/\ln(t+3)}{\sum_{j=1}^{t} 1/\ln(j+2)} = \lim_{t\to\infty} \frac{t/\ln(t+3)}{t/\ln t} = 1.$$

Then TGPA(p_t, r_t, q_t) follows a power-law degree distribution with exponent $\beta = 1 + 2\Gamma/(3-y) = 1 + 2/(3-0) = 5/3$.

□

For a final analysis, we show that the component portion is necessary to obtain the full power-law exponent range $(1, \infty)$. Lemma 7.3 comes directly from [7].

LEMMA 7.3 [7] Assume $\lim_{t\to\infty} y_t = y$ and $\lim_{t\to\infty} y_{t+1} \cdot t / \sum_{j=1}^{t} j_j = \Gamma$. Then for $y > 0$ we have $\Gamma = 1$, and for $y = 0$ we have $\Gamma \leq 1$.

COROLLARY 7.2 Consider TGPA(p_t, r_t, q_t). Assume that $\lim_{t\to\infty} q_t = 0$, $\lim_{t\to\infty} y_t = y$ and $y_{t+1}t / \sum_{j=1}^{t} y_j = \Gamma > 0$. Then the resulting graph follows a power-law degree distribution with exponent $\beta \in (1, 3]$.

Proof. By Corollary 7.1, TGPA(p_t, r_t, q_t) follows a power law in the degree distribution with exponent $\beta = 1 + 2\Gamma/(3-y) > 1$. By Lemma 7.3, for $0 < y \leq 1$, we have $\beta = 1 + 2/(3-y) \in (5/3, 3]$ and for $y = 0$, we have $\beta = 1 + 2\Gamma/3 \leq 5/3$. □

8. Conclusions and discussion

In this article, we presented a graph model called TGPA, which incorporates direct triangle formulation into a GPA model that includes possibly disconnected components. Furthermore, we provided extensive analysis of this model, showing that the degree and spectral distributions fit power-law distributions. We also extended the results for the GPA model found in [7] as well as the model defined in [4].

Our new model provides a useful platform for studying real-world network data. The importance of power laws in the spectra of real-world networks has been shown [8]; however, an explanation for why this feature occurs remains to be found. We hope that by introducing a model with this feature along with realistic clustering coefficients will lead to further explanation for the presence of highly skewed spectra. In the future, we plan to study further generalizations of higher-order PA graphs.

Funding

National Science Foundation CAREER (CCF-1149756, IIS-1422918, IIS-1546488, and CCF-1909528); National Science Foundation Center for Science of Information STC (CCF-0939370); and the Sloan Foundation.

References

1. Price, D. D. S. (1976) A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inf. Sci.*, **27**, 292–306.
2. Yule, G. U. (1925) II.—A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F. R. S. *Philos. Trans. R. Soc. Lond. B*, **213**, 21–87.
3. Barabási, A.-L. & Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
4. Holme, P. & Kim, B. J. (2002) Growing scale-free networks with tunable clustering. *Phys. Rev. E*, **65**, 026107.
5. Ostroumova, L., Ryabchenko, A. & Samosvat, E. (2013) Generalized preferential attachment: tunable power-law degree distribution and clustering coefficient. *International Workshop on Algorithms and Models for the Web-Graph* (A. Bonato, M. Mitzenmacher & P. Prałat eds). Cham: Springer, pp. 185–202.
6. Shang, Y. (2012) Distinct clusterings and characteristic path lengths in dynamic small-world networks with identical limit degree distribution. *J. Stat. Phys.*, **149**, 505–518.
7. Avin, C., Lotker, Z., Nahum, Y. & Peleg, D. (2017) Improved degree bounds and full spectrum power laws in preferential attachment networks. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, pp. 45–53.
8. Eikmeier, N. & Gleich, D. F. (2017) Revisiting power-law distributions in spectra of real world networks. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *KDD '17*. New York, NY, USA: ACM, pp. 817–826.
9. Flaxman, A., Frieze, A. & Fenner, T. (2005) High degree vertices and eigenvalues in the preferential attachment graph. *Internet Math.*, **2**, 1–19.
10. Mihail, M. & Papadimitriou, C. (2002) On the eigenvalue power law. *Randomization and Approximation Techniques in Computer Science* (J. D. P. Rolim & S. Vadhan eds). Berlin, Heidelberg: Springer, pp. 254–262.
11. Broido, A. D. & Clauset, A. (2019) Scale-free networks are rare. *Nat. Commun.*, **10**, 1017.
12. Sala, A., Zheng, H., Zhao, B. Y., Gaito, S. & Rossi, G. P. (2010) Brief announcement: revisiting the power-law degree distribution for social graph analysis. *Proceedings of the 29th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing*, *PODC '10*. New York, NY, USA: ACM, pp. 400–401.
13. Lattanzi, S. & Sivakumar, D. (2009) Affiliation networks. *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*. New York, NY, USA: ACM, pp. 427–434.
14. Shang, Y. (2014) Geometric assortative growth model for small-world networks. *Sci. World J.*, **2014**.
15. Faloutsos, M., Faloutsos, P. & Faloutsos, C. (1999) On power-law relationships of the internet topology. *SIGCOMM '99 Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*. New York, NY, USA: ACM.
16. Huberman, B. A. (2001) *The Laws of the Web*. Cambridge, MA: The MIT Press.
17. Medina, A., Matta, I. & Byers, J. (2000) On the origin of power laws in internet toplogies. *ACM SIGCOMM Comput. Commun. Rev.*, **30**, 18–28.
18. Chung, F., Chung, F. R., Graham, F. C., Lu, L., Chung, K. F. *et al.* (2006) *Complex Graphs and Networks*, vol. 107. Providence, RI, USA: American Mathematical Society.
19. Cooper, C. & Frieze, A. M. (2001) A general model of undirected web graphs. *European Symposium on Algorithms*. Berlin, Heidelberg: Springer, pp. 500–511.
20. Saramäki, J. & Kaski, K. (2004) Scale-free networks generated by random walkers. *Physica A*, **341**, 80–86.
21. Toivonen, R., Onnela, J.-P., Saramäki, J., Hyvönen, J. & Kaski, K. (2006) A model for social networks. *Physica A*, **371**, 851–860.

22. Zadorozhnyi, V. & Yudin, E. (2015) Growing network: models following nonlinear preferential attachment rule. *Physica A*, **428**, 111–132.
23. Benson, A., Gleicfh, D. F. & Lim, L.-H. (2017) The spacey random walk: a stochastic process for higher-order data. *SIAM Rev.*, **59**, 321–345.
24. Benson, A. R., Gleich, D. F. & Leskovec, J. (2016) Higher-order organization of complex networks. *Science*, **353**, 163–166.
25. Grilli, J. *et al.* (2017) Higher-order interactions stabilize dynamics in competitive network models. *Nature*, **548**, 210–213.
26. Rosvall, M. *et al.* (2014) Memory in network flows and its effects on spreading dynamics and community detection. *Nat. Commun.*, **5**.
27. Xu, J., Wickramarathne, T. L. & Chawla, N. V. (2016) Representing higher-order dependencies in networks. *Sci. Adv.*, **2**, e1600028.
28. Yin, H., Benson, A. R. & Leskovec, J. (2018) Higher-order clustering in networks. *Phys. Rev. E*, **97**, 052306.
29. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. (2002) Network motifs: simple building blocks of complex networks. *Science*, **298**, 824–827.
30. Eikmeier, N., Ramani, A. S. & Gleich, D. F. (2018) The HyperKron graph model for higher-order features. *Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, pp. 941–946.
31. Lambiotte, R., Rosvall, M. & Scholtes, I. (2019) From networks to optimal higher-order models of complex systems. *Nat. Phys.*, **15**, 313–320.
32. Traud, A. L., Mucha, P. J. & Porter, M. A. (2012) Social structure of Facebook networks. *Physica A*, **391**, 4165–4180.
33. van der Hofstad, R. (2016) *Random Graphs and Complex Networks,* vol. 1. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge, UK: Cambridge University Press.
34. Chung, F., Lu, L. & Vu, V. (2003) Eigenvalues of random power law graphs. *Ann. Combin.*, **7**, 21–33.
35. Chung, F., Lu, L. & Vu, V. (2003) Spectra of random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, **100**, 6313–6318.
36. Boyd, S. & Vandenberghe, L. (2004) *Convex Optimization.* Cambridge, UK: Cambridge University Press.
37. Golub, G. H. & Van Loan, C. F. (2013) *Matrix Computations*, 4th edn. Baltimore, MD, USA: The John Hopkins University Press.

Appendix A. GPA proofs

In this appendix, we provide proofs for theoretical results from Section 5.

Proof of Lemma 5.1. Denote G_t^m as the graph at time t with contractions of size m. Let $Z_t = d_t^m(s)$ be the degree of vertex s at time t, and Y_t an indicator for the event that the edge added at time t is incident to s. Then,

$$\begin{aligned}\mathbb{E}[Z_t^{(k)}] &= \mathbb{E}[\mathbb{E}[(Z_{t-1} + Y_t)^{(k)}]|Z_{t-1}] \\ &= \mathbb{E}\left[Z_{t-1}^{(k)}\left(1 - p \cdot \frac{Z_{t-1}}{2t-1}\right) + (Z_{t-1}+1)^{(k)}\left(p \cdot \frac{Z_{t-1}}{2t-1}\right)\right] \\ &= \mathbb{E}[Z_{t-1}^{(k)}]\left(1 + \frac{pk}{2t-1}\right).\end{aligned}$$

Next we apply this relationship iteratively, down to the time when node s was added. Denote this as time s', and also note that the degree of s at time s' is bounded by $2m$ (if all m edges were added as self-loops) to get

$$\mathbb{E}(Z_t^{(k)}) = \prod_{t'=s'}^{t}\left(1+\frac{pk}{2t'-1}\right) \leq (2m)^{(k)} \prod_{t'=s'+1}^{t}\left(1+\frac{pk}{2t'-1}\right).$$

Use $1+x \leq e^x$ to write the product as a sum, and bound

$$\sum_{t'=s'+1}^{t}\frac{1}{t'-1/2} \leq \int_{x=s'}^{t}\frac{1}{x-1/2}\ \mathrm{dx} = \log\frac{t-1/2}{s'-1/2}.$$

So finally,

$$\mathbb{E}(Z_t^{(k)}) \leq (2m)^{(k)}\left(\frac{t-1/2}{s'-1/2}\right)^{pk/2} = (2m)^{(k)}\left(\frac{t}{s'}\right)^{pk/2}\left(\frac{2-1/t}{2-1/s'}\right)^{pk/2}(2m)^{(k)}\left(\frac{t}{s'}\right)^{pk/2}2^{pk/2}.$$

□

Proof of Lemma 5.2. Let $\tau^{(i)} = (\tau_1^{(i)},\ldots,\tau_{r_i}^{(i)})$, where $\tau_j^{(i)}$ is the time when we add an edge incident to S_i and increase the degree from d_i+j-1 to d_i+j. Define $\tau = (\tau_0,\tau_1,\ldots,\tau_{r+1})$ to be the ordered union of $\tau^{(i)}$, with $\tau_0 = t_0$ and $\tau_{r+1} = t$. Let $p(\tau;\mathrm{d},t_0,t)$ be the probability that S_i increases in degree at exactly the times specified by τ. Then

$$\begin{aligned} p(\tau;\mathrm{d},t_0,t) &= \left(\prod_{i=1}^{l}\prod_{k=1}^{r_i}p\frac{d_i+k-1}{2\tau_k^{(i)}-1}\right)\left(\prod_{k=0}^{r}\prod_{j=\tau_k+1}^{\tau_{k+1}-1}\left(1-p\frac{d+k}{2j-1}\right)\right) \\ &= \left(\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right)\left(\prod_{k=1}^{r}\frac{p}{2\tau_k-1}\right)\exp\left\{\sum_{k=0}^{r}\sum_{j=\tau_k+1}^{\tau_{k+1}-1}\log\left(1-p\left(\frac{d+k}{2j-1}\right)\right)\right\}. \end{aligned}$$

We can bound the inner most sum of the exponential term using a simple inequality

$$\sum_{j=\tau_k+1}^{\tau_{k+1}-1}\log\left(1-p\left(\frac{d+k}{2j-1}\right)\right) \leq \sum_{j=\tau_k+1}^{\tau_{k+1}-1}\log\left(1-\frac{p(d+k)}{2j}\right),$$

which is less than or equal to

$$\begin{aligned} \int_{\tau_k+1}^{\tau_{k+1}}\log\left(1-\frac{p(d+k)}{2x}\right)\ \mathrm{dx} = &-\tau_{k+1}\log(2\tau_{k+1}) + (\tau_k+1)\log(2\tau_k+2) \\ &- 1/2(2\tau_{k+1}-p(d+k))\log(2\tau_{k+1}-p(d+k)) \\ &- 1/2(2\tau_k+2-p(d+k))\log(2\tau_k+2-p(d+k)). \end{aligned}$$

Note that $\tau_0 = t_0$ and $\tau_{r+1} = t$. Then we can write

$$\sum_{k=0}^{r}\int_{\tau_k+1}^{\tau_{k+1}}\log\left(1-\frac{p(d+k)}{2x}\right)\ \mathrm{dx} = A + \sum_{k=1}^{r}B_k,$$

where

$$A = (t_0+1)\log(2t_0+2) - 1/2(2t_0+2-pd)\log(2t_0+2-pd) \\ - t\log(2t) + 1/2(2t-p(d+r))\log(2t-p(d+r)) \tag{A.1}$$

and

$$B_k = \tau_k \log(1+1/\tau_k) + \log(2\tau_k+2) - \frac{2-p}{2}\log(2\tau_k+p-p(d+k)) \\ + 1/2(2\tau_k+2-p(d+k))\log\left(1-\frac{2-p}{2\tau_k+2-p(d+k)}\right). \tag{A.2}$$

Bound each of A and B_k, starting with B_k. Since $1+x \le e^x$, $\tau_k\log(1+1/\tau_k) \le 1$ and $\frac{1}{2}(2\tau_k+2-p(d+k))\log\left(1-\frac{2-p}{2\tau_k+2-p(d+k)}\right) \le -p/2$. Rearranging the other two terms of Equation (A.2), we get

$$B_k \le \frac{p}{2}\log(2\tau_k+2) - \frac{2-p}{2}\log\left(1-\frac{p(d+k)+2-p}{2\tau_k+2}\right) + \frac{p}{2}.$$

Rearranging terms of A from Equation (A.1) and taking the exponential,

$$e^A = \left(1-\frac{pd}{2t_0+2}\right)^{-(t_0+1)}(2t_0+2-pd)^{pd/2}\left(1-\frac{p(d+r)}{2t}\right)^t(2t-p(d+r))^{\frac{-p(d+r)}{2}} \\ = \left(1-\frac{pd}{2t_0+2}\right)^{-(1-\frac{pd}{2(t_0+1)})(t_0+1)}\left(1-\frac{p(d+r)}{2t}\right)^{t-\frac{p(d+r)}{2}}\left(\frac{t_0+1}{t}\right)^{\frac{pd}{2}}(2t)^{\frac{-pr}{2}}.$$

Using the bound $1-x \le e^{-x-x^2/2}$ for $0 < x < 1$,

$$\left(1-\frac{p(d+r)}{2t}\right)^{t-p(d+r)/2} \le \exp\left\{-\frac{p(d+r)}{2} + \frac{p^2(d+r)^2}{8t} + \frac{p^3(d+r)^3}{16t^2}\right\}.$$

Putting the bounds on A and B_k together, we get

$$e^{A+\sum B_k} \le \left(1-\frac{pd}{2t_0+2}\right)^{-(1-\frac{pd}{2(t_0+1)})(t_0+1)}\exp\left\{-\frac{p(d+r)}{2} + \frac{p^2(d+r)^2}{8t} + \frac{p^3(d+r)^3}{16t^2}\right\} \\ \times\left(\frac{t_0+1}{t}\right)^{pd/2}(2t)^{-pr/2}\prod_{k=1}^{r}\left(\left(1-\frac{p(d+k)+2-p}{2\tau_k+2}\right)^{-(2-p)/2}(2\tau_k+2)^{p/2}\right)e^{pr/2}. \tag{A.3}$$

Define $\mathrm{err}(r,d,t_0,t)$

$$\mathrm{err}(r,d,t_0,t) = \left(1-\frac{pd}{2t_0+2}\right)^{-(1-\frac{pd}{2(t_0+1)})(t_0+1)}\exp\left\{-\frac{pd}{2} + \frac{p^2(d+r)^2}{8t} + \frac{p^3(d+r)^3}{16t^2}\right\},$$

then we can rewrite Equation (A.3) as

$$\operatorname{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{\frac{pd}{2}}(2t)^{\frac{-pr}{2}}\prod_{k=1}^{r}\left(\left(1-\frac{p(d+k)+2-p}{2\tau_k+2}\right)^{\frac{-(2-p)}{2}}(2\tau_k+2)^{\frac{p}{2}}\right).$$

So, we finally finish with the bound on $p(\tau;\mathrm{d},t_0,t)$ by substituting Equation (A.3) into Equation (A.1) and rearranging terms:

$$p(\tau;\mathrm{d},t_0,t)=\left(\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right)\operatorname{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd/2}(2t)^{-pr/2}$$
$$\times\prod_{k=1}^{r}\left(p(2\tau_k+p-p(d+k))^{-(2-p)/2}\left(1+\frac{3}{2\tau_k-1}\right)\right).$$

Now sum $p(\tau;\mathrm{d},t_0,t)$ over all ordered choices of τ to get

$$p(\mathrm{r};\mathrm{d},t_0,t)\le\sum_{\tau^{(1)},\ldots,\tau^{(l)}}p(\tau;\mathrm{d},t_0,t)$$

$$\le\binom{r}{r_1,\ldots,r_l}\sum_{t_0+1\le\tau_1<\ldots<\tau_r\le t}\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\operatorname{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{\frac{pd}{2}}$$
$$\times(2t)^{\frac{-pr}{2}}p\prod_{k=1}^{r}(2\tau_k+p-p(d+k))^{-(2-p)/2}\left(1+\frac{3}{2\tau_k-1}\right)$$
$$=r!\left(\prod_{i=1}^{l}\binom{r_i+d_i-1}{d_i-1}\right)\operatorname{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd/2}(2t)^{-pr/2}$$
$$\times\sum_{t_0+1\le\tau_1<\ldots<\tau_r\le t}p\prod_{k=1}^{r}(2\tau_k+p-p(d+k))^{-(2-p)/2}\left(1+\frac{3}{2\tau_k-1}\right).\tag{A.4}$$

Now let $\tau_k'=\tau_k-\lceil p(d+k)/2\rceil$. Since $d\ge 1$ and $k\ge 1$, we have $2\lceil p(d+k)/2\rceil\ge 2$. So Equation (A.4) is less than or equal to

$$\sum_{(t_0-p\lceil d/2\rceil+1)\le\tau_1'\le\ldots\le\tau_r'\le(t-p\lceil(d+r)/2\rceil)}\left(p\prod_{k=1}^{r}(2\tau_k'+p)^{-(2-p)/2}\left(1+\frac{3}{2\tau_k'+1}\right)\right)$$
$$\le\frac{p}{r!}\left(\sum_{\tau'=(t_0-p\lceil d/2\rceil+1)}^{t-p\lceil(d+r)/2\rceil}\left(2\tau'+p\right)^{-(2-p)/2}+3(2\tau'+1)^{-(4-p)/2}\right)^{r}$$

$$= \frac{1}{r!}(2t)^{pr/2} \underbrace{\left(1 - \frac{p(d+r)-p}{2t}\right)^{pr/2}}_{\leq \exp\left\{-\frac{rp(p(d+r)-p)}{4t}\right\}} \underbrace{\left(1 + \frac{3p}{(2t-p(d+r)+p)^{p/2}}\right)^{r}}_{\leq \exp\left\{\frac{3pr}{(2t-p(d+r)+p)^{p/2}}\right\}}.$$

Where the last inequalities come from $1 + x \leq e^x$. So finally,

$$p_S(\mathrm{r}; \mathrm{d}, t_0, t) \leq \left(\prod_{i=1}^{l} \binom{r_i + d_i - 1}{d_i - 1}\right) \mathrm{err}(r, d, t_0, t) \left(\frac{t_0+1}{t}\right)^{pd/2}$$
$$\exp\left\{\frac{-rp((d+r)-p)}{4t} + \frac{3pr}{(2t-p(d+r)+p)^{p/2}}\right\}.$$

Using $d = o(t^{1/2})$ and $r = o(t^{2/3})$ gives the final bound, and this concludes the proof. □

Proof of Theorem 5.1. Partition the vertices into those added before time t_0, before time t_1, and after t_1, with $t_0 = \log\log\log f(t), t_1 = \log\log f(t)$. We will argue about the maximum degree vertices in each set.

CLAIM A.1 In G_t^m, the degree of the supernode of vertices added before time t_0 is at least $t_0^{1/3} t^{p/2}$ whp.

Proof. Consider all vertices added before time t_0 as a supernode. Let A_1 denote the event that this supernode has degree less than $t_0^{1/3} t^{p/2}$ at time t. Use Lemma 5.2 with $l = 1$, and $d = 2t_0$ to bound

$$Pr[A_1] \leq \sum_{r_1=0}^{t_0^{1/3} t^{p/2} - 2t_0} \binom{r_1 + 2t_0 - 1}{2t_0 - 1} \left(\frac{t_0+1}{t}\right)^{pd/2} e^{2+t_0-pd/2+3pr/t^{p/2}}.$$

Then, since $r \leq t_0^{1/3} t^{1/2} - 2t_0$,

$$Pr[A_1] \leq \sum_{r_1=0}^{t_0^{1/3} t^{p/2} - 2t_0} \binom{t_0^{1/3} t^{p/2} - 1}{2t_0 - 1} \left(\frac{t_0+1}{t}\right)^{pt_0} e^{2+t_0-pt_0+3pt_0^{1/3}-6pt_0/t^{p/2}}.$$

Replacing the sum with $(t_0^{1/3} t^{p/2} - 2t_0)$, and the using the definition of the combination,

$$Pr[A_1] \leq (t_0^{1/3} t^{p/2} - 2t_0) \frac{(t_0^{1/3} t^{p/2} - 1)!}{(2t_0 - 1)!(t_0^{1/3} t^{p/2} - 2t_0)!} \left(\frac{t_0+1}{t}\right)^{pt_0} e^{2+t_0(1-p)+3pt_0^{1/3}-6pt_0/t^{p/2}}.$$

And finally, by writing out $(t_0^{1/3} t^{p/2} - 1)!/(t_0^{1/3} t^{p/2} - 2t_0)!$, we can reduce the expression further as

$$Pr[A_1] \leq t_0^{1/3} t^{p/2} \frac{(t_0^{1/3} t^{p/2})^{2t_0-1}}{(2t_0 - 1)!} \left(\frac{t_0+1}{t}\right)^{pt_0} e^{2+t_0(1-p)+3pt_0^{1/3}-6pt_0/t^{p/2}}.$$

Using $1/x \leq e^x/x^x$ and rearranging terms, $P[A_1]$ goes to 0

$$Pr[A_1] \leq \frac{e^{1+(3-p)t_0+3pt_0^{1/3}-6pt_0/t^{p/2}}}{(2t_0-1)^{t_0(4/3-p)-1}}.$$ □

CLAIM A.2 In G_t^m, no vertex added after time t_1 has degree exceeding $t_0^{-2}t^{p/2}$ whp.

Proof. Let A_2 denote the event that some vertex added after time t_1 has degree exceeding $t_0^{-2}t^{p/2}$. Write the probability of A_2 occurring as

$$\begin{aligned} Pr[A_2] &\leq \sum_{s=t_1}^{t} \Pr[d_t(s) \geq t_0^{-2}t^{p/2}] = \sum_{s=t_1}^{t} \Pr[(d_t(s))^{(l)} \geq (t_0^{-2}t^{p/2})^{(l)}] \\ &\leq \sum_{s=t_1}^{t} t_0^{2l} t^{-lp/2} \mathbb{E}[(d_t(s))^{(l)}] = \sum_{s=t_1}^{t} t_0^{2l} t^{-lp/2} (2m)^{(l)} 2^{lp/2} \left(\frac{t}{s}\right)^{lp/2} \\ &= 2^{lp/2}(2m)^{(l)} t_0^{2l} \int_{t_1-1}^{t} x^{-lp/2} \text{ dx}. \end{aligned} \tag{A.5}$$

Compute the integral in Equation (A.5)

$$\int_{t_1-1}^{t} x^{-lp/2} \text{ dx} = \left. \frac{x^{-lp/2+1}}{-lp/2+1} \right|_{t_1-1}^{t} = (-lp/2+1)^{-1} \left(t^{-lp/2+1} - (t_1-1)^{-lp/2+1}\right). \tag{A.6}$$

Choose $l > 2/p$. Then the integral in Equation (A.6) is less than or equal to $(lp/2-1)^{-1}(t_1-1)^{-lp/2+1}$, and plugging in the computation from Equation (A.6) into Equation (A.5),

$$Pr[A_2] \leq \frac{2^{lp/2}(2m)^{(l)}t_0^{2l}}{(lp/2-1)(t_1-1)^{lp/2-1}}$$

which goes to 0 as t increases. □

CLAIM A.3 In G_t^m, no vertex added before time t_1 has degree exceeding $t_0^{1/6}t^{p/2}$ whp.

Proof. Use same technique as in Claim A.2. □

CLAIM A.4 The k highest degree vertices of G_t^m are added before time t_1 and have degree Δ_i bounded by $t_0^{-1}t^{p/2} \leq \Delta_i \leq t_0^{1/6}t^{p/2}$.

Proof. If the lower bound does not hold, then one of the top k vertices has degree less than $t_0^{-1}t^{p/2}$ and the total degree of vertices added before time t_0 is bounded by

$$\begin{aligned} &(k-1)t_0^{1/6}t^{p/2} + \left(\frac{t_0}{m} - k + 1\right)\left(t_0^{-1}t^{p/2}\right) \leq kt_0^{1/6}t^{p/2} + t_0(t_0^{-1}t^{p/2}) \\ &= kt_0^{1/6}t^{p/2} + t^{p/2} = t^{p/2}(kt_0^{1/6}+1) \leq t^{p/2}(2kt_0^{1/6}) \leq t^{p/2}t_0^{1/3}. \end{aligned}$$

Since we have the lower bound, and we know that $t^{p/2}/t_0 \geq t^{p/2}/t_0^2$, none of the largest degree vertices could be added after time t_1. □

CLAIM A.5 The k highest degree vertices have $\Delta_i \leq \Delta_{i-1} - \frac{t^{p/2}}{f(t)}$ whp.

Proof. Let A_4 denote the event that there are two vertices among the first t_1 time steps with degrees exceeding $t_0^{-1}t^{p/2}$ and within $t^{p/2}/f(t)$ of each other. Let $\overline{A}_3$ be the opposite of event A_3 from Claim A.3. Let

$$p_{l,s_1,s_2} = \Pr\left[d_t(s_1) - d_t(s_2) = l \mid \overline{A}_3\right], \text{for } |l| \leq t^{p/2}/f(t). \tag{A.7}$$

Then,

$$\Pr\left[A_4|\overline{A}_3\right] \leq \sum_{1\leq s_1<s_2\leq t_1} \sum_{l=-t^{p/2}/f(t)}^{t^{p/2}/f(t)} p_{l,s_1,s_2}$$

$$p_{l,s_1,s_2} \leq \sum_{r_1=t_0^{-1}t^{p/2}}^{t_0^{1/6}t^{p/2}} \sum_{d_1,d_2=1}^{2t_1} \underbrace{p_{(s_1,s_2)}\left((r_1, r_1 - l); (d_1, d_2), t_1, t\right)}_{\text{Notation from Lemma 5.2.}}$$

$$\leq \sum_{r_1=t_0^{-1}t^{p/2}}^{t_0^{1/6}t^{p/2}} \sum_{d_1,d_2=1}^{2t_1} \binom{r_1+d_1-1}{d_1-1}\binom{r_1-l+d_2-1}{d_2-1}\left(\frac{t_1+1}{t}\right)^{\frac{p(d_1+d_2)}{2}} e^{\left\{2+t_1-\frac{p(d_1+d_2)}{2}+\frac{3p(r_1-l)}{t^{p/2}}\right\}}$$

$$\leq t_0^{1/6}t^{p/2} \sum_{d_1,d_2=1}^{2t_1} (2t_0^{1/6}t^{p/2})^{d_1+d_2-2}(t_1+1)^{2pt_1}t^{-p(d_1+d_2)/2}e^{2+t_1+3pt_0^{1/6}}$$

$$= t^{-p/2}t_0^{1/6}(2t_1)^2 2^{4t_1} t_0^{2t_1/3}(t_1+1)^{2pt_1}e^{2+t_1+3pt_0^{1/6}}.$$

Denote the last equation as $h(t)$ and note $h(t)$ is a polynomial in $\log(f(t))$ times a factor of $t^{-p/2}$. Then going back to Equation (A.7),

$$\Pr\left[A_4|\overline{A}_3\right] \leq \binom{t_1}{2}2\frac{t^{p/2}}{f(t)}h(t) = \binom{t_1}{2}2\frac{\text{poly}(\log(f(t)))}{f(t)},$$

which goes to 0 as t increases. □

Finishing that final Claim finishes the proof of the theorem. □

Proof of Theorem 5.2. Let $G = G_t^m$. We will show that with high probability G contains a star forest F, with stars of degree asymptotic to the maximum degree vertices of G. Then show that $G\backslash F$ has small eigenvalues. Then we can use Rayleigh's principle to say that the large eigenvalues of G cannot be too different than the large eigenvalues of F.

Let S_i be the vertices added after time t_{i-1} and at or before time t_i, for $t_0 = 0, t_1 = t^{1/8}, t_2 = t^{9/16}, t_3 = t$. We start by finding bounds on the degrees of G.

CLAIM A.6 For any $\varepsilon > 0$, and any $f(t)$ with $f(t) \to \infty$ as $t \to \infty$ the following holds whp: for all s with $f(t) \le s \le t$, for all vertices $v \in G_s$, if v was added at time r, then $d_s(v) \le s^{p/2+\varepsilon} r^{-p/2}$.

Proof. We prove this by contradiction. Bound the probability that $d_s^m(r)$ is greater than $s^{p/2+\varepsilon} r^{-p/2}$:

$$\Pr\left[\cup_{s=f(t)}^{t} \cup_{r=1}^{s} \left\{d_s^m(r) \ge s^{p/2+\varepsilon} r^{-p/2}\right\}\right] \le \sum_{s=f(t)}^{t} \sum_{r=1}^{s} \Pr\left[d_s^m(r) \ge s^{p/2+\varepsilon} r^{-p/2}\right]$$

$$= \sum_{s=f(t)}^{t} \sum_{r=1}^{s} \Pr\left[(d_s^m(r))^{(l)} \ge (s^{p/2+\varepsilon} r^{-p/2})^{(l)}\right],$$

which is bounded using Markov:

$$\le \sum_{s=f(t)}^{t} \sum_{r=1}^{s} s^{-l(p/2+\varepsilon)} r^{pl/2} \mathbb{E}\left[(d_s^m(r))^{(l)}\right],$$

and which we can bound using Lemma 5.1

$$\le \sum_{s=f(t)}^{t} \sum_{r=1}^{s} s^{-l(p/2+\varepsilon)} r^{lp/2} (2m)^{(l)} 2^{lp/2} \left(\frac{s}{r}\right)^{lp/2} = (2m)^{(l)} 2^{lp/2} \sum_{s=f(t)}^{t} s^{1-\varepsilon l}.$$

Take $l \ge 3/\varepsilon$. Then we can bound the sum by an integral,

$$\sum_{s=f(t)}^{t} s^{1-\varepsilon l} \le \int_{f(t)-1}^{\infty} x^{1-\varepsilon l}\, \mathrm{dx} = \frac{1}{2-\varepsilon l} x^{2-\varepsilon l}\Big|_{f(t)-1}^{\infty} = \frac{1}{\varepsilon l - 2}(f(t)-1)^{2-\varepsilon l},$$

which goes to zero as t increases, since $l \ge 3/\varepsilon$. □

CLAIM A.7 Let S_3' be the set of vertices in S_3 that are adjacent to more than one vertex of S_1 in G. Then $|S_3'| \le t^{7p/16}$ with high probability.

Proof. Let B_1 be the event that the conditions of Claim A.6 hold with $f(t) = t_2$ and $\varepsilon = 1/16$. Then for a vertex $v \in S_3$ added at time s, the probability that v picks at least one neighbour in S_1 is less than or equal to

$$\frac{\sum_{w \in S_1} d_s(w)}{2s-1} \le \frac{\sum_{w \in S_1} s^{p/2+\varepsilon}}{2s-1} = \frac{t_1 s^{p/2+\varepsilon}}{2s-1}.$$

Then the probability of having two or more neighbours in S_1 can be bounded by

$$\Pr\left[\, |N(v) \cap S_1| \ge 2 \mid B_1 \right] \le \left(\frac{t_1 s^{p/2+\varepsilon}}{2s-1}\right)^2 \cdot \binom{m}{2} \le m^2 t^{1/4} s^{(-15+8p)/8}.$$

Let X denote the number of $v \in S_3$ adjacent to more than one vertex of S_1. Then

$$\mathbb{E}[X|B_1] \leq \sum_{t_2+1}^{t} m^2 s^{(-15+8p)/8} t^{1/4} \leq m^2 t^{1/4} \int_{t_2}^{t} x^{(-15+8p)/8} \, \mathrm{dx}$$
$$= m^2 t^{1/4} \left[\frac{8}{7+8p} x^{(-7+8p)/8} \Big|_{t_2}^{t} \right] \leq \frac{8m^2 t^{1/4}}{-7+8p} t^{(-7+8p)/8}.$$

Then by Markov,

$$\Pr\left[X \geq t^{7p/16} | B_1\right] \leq \frac{\mathbb{E}[X|B_1]}{t^{7p/16}} \leq \frac{8m^2}{8p-7} \frac{t^{p-5/8}}{t^{7p/16}} = \frac{8m^2}{8p-7} \frac{t^{-5/8}}{t^{-9p/16}},$$

and $\frac{t^{-5/8}}{t^{-9p/16}} = \frac{t^{9p/16}}{t^{5/8}} \leq \frac{t^{9/16}}{t^{5/8}}$ which goes to zero. □

Let $F \subseteq G$ be the star forest consisting of edges between S_1 and $S_3 \backslash S_3'$.

Claim A.8 Let $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_k$ denote the degrees of the k highest degree vertices of G. Then $\lambda_i(F) = (1-o(1))\Delta_i^{1/2}$.

Proof. Denote K_{1,d_i} to be a star of degree d_i. Let H be the star forest $H = K_{1,d_1} \cup \ldots \cup K_{1,d_k}$ with $d_1 \geq d_2 \geq \ldots \geq d_k$. Then for $i = 1,\ldots,k$, $\lambda_i(H) = d_i^{1/2}$. So it will be sufficient to show that $\Delta_i(F) = (1-o(1))\Delta_i(G)$. Within the proof of Theorem 5.1, we show that the k highest degree vertices G are added before time t_1 (specifically in Claim A.4). So these vertices are all in F. The only edges to those vertices that are not in F are those added before time t_2 and those incident to S_3'.

By Theorem 5.1, we can choose $f(t)$ such that $\Delta_1(G_{t_2}^m) \leq t_2^{p/2} f(t) \leq t^{7p/16}$. Additionally by Theorem 5.1, we get $\Delta_i(G) \geq t^{p/2}/\log t$. Finally, Claim A.7 says that $|S_3'| \leq t^{7p/16}$ whp. So with high probability,

$$\Delta_i(F) \geq \Delta_i(G) - t^{7p/16} - mt^{7p/16} \geq \frac{t^{p/2}}{\log t} - t^{7p/16}(1+m) = \frac{t^{p/2}}{\log t}\left[1 - t^{7p/16}(1+m)\frac{\log t}{t^{p/2}}\right]$$
$$= \frac{t^{p/2}}{\log t}\left[1-(1+m)\frac{\log t}{t^{p/2-7p/16}}\right] = \frac{t^{p/2}}{\log t}\left[1-(1+m)\frac{\log t}{t^{p/16}}\right] = (1-o(1))\Delta_i(G).$$ □

Let $H = G\backslash F$. Denote $\boldsymbol{A}_G$, $\boldsymbol{A}_F$ and $\boldsymbol{A}_H$ to be the adjacency matrices for graphs G, F and H. In the following claim, we'll show that $\lambda_1(\boldsymbol{A}_H)$ is $o(\lambda_k(\boldsymbol{A}_F))$. Consider the fact that if $\boldsymbol{A}$ and $\boldsymbol{A}+\boldsymbol{E}$ are symmetric n by n matrices, then $\lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{E}) \leq \lambda_k(\boldsymbol{A}) + \lambda_1(\boldsymbol{E})$ (see for instance [37]). That implies that for any subspace L,

$$\max_{\mathrm{x}\in L, \mathrm{x}\neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_G \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} = \max_{\mathrm{x}\in L, \mathrm{x}\neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_F \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} \pm O\left(\max_{\mathrm{x}\neq 0} \frac{\mathrm{x}^T \boldsymbol{A}_H \mathrm{x}}{\mathrm{x}^\top \mathrm{x}}\right).$$

This is enough to finish the proof because by the Courant–Fischer Minimax Theorem ([37], Theorem 8.1.2), $\lambda_i(\boldsymbol{A}_G) = \lambda_i(\boldsymbol{A}_F)(1 \pm o(1))$.

CLAIM A.9 $\lambda_1(\boldsymbol{A}_H) \le 6mt^{15/64}$ whp.

Proof. Let H_i denotes the subgraph of H induced by S_i, and let H_{ij} denotes the subgraph of H containing only edges with one vertex in S_i and the other in S_j. That is, write $\boldsymbol{A}_H$ in the following way:

$$\boldsymbol{A}_H = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_{12} & \boldsymbol{H}_{13} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_2 & \boldsymbol{H}_{23} \\ \boldsymbol{H}_{31} & \boldsymbol{H}_{32} & \boldsymbol{H}_3 \end{bmatrix},$$

and use this to bound the maximal eigenvalue of $\boldsymbol{A}_H$ as

$$\begin{aligned} \lambda_1(\boldsymbol{A}_H) &= \lambda_1\left(\begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_{12} & \boldsymbol{H}_{13} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_2 & \boldsymbol{H}_{23} \\ \boldsymbol{H}_{31} & \boldsymbol{H}_{32} & \boldsymbol{H}_3 \end{bmatrix}\right) \\ &\le \lambda_1(\boldsymbol{H}_1) + \lambda_1(\boldsymbol{H}_2) + \lambda_1(\boldsymbol{H}_3) + \lambda_1(\boldsymbol{H}_{12}) + \lambda_1(\boldsymbol{H}_{23}) + \lambda_1(\boldsymbol{H}_{13}). \end{aligned}$$

Note that the maximum eigenvalue of a graph is at most the maximum degree of a graph. By Claim A.6 with $f(t) = t_1$ and $\varepsilon = 1/64$,

$$\begin{aligned} \lambda_1(\boldsymbol{H}_1) &\le \Delta_1(\boldsymbol{H}_1) = \max_{v \le t_1}\{d^m_{t_1}(v)\} \le t_1^{p/2+\varepsilon} \le t^{33/512} \\ \lambda_1(\boldsymbol{H}_2) &\le \Delta_1(\boldsymbol{H}_2) \le \max_{t_1 \le v \le t_2}\{d^m_{t_2}(v)\} \le t_2^{p/2+\varepsilon}/t_1^{p/2} \le t^{233/1024} \\ \lambda_1(\boldsymbol{H}_3) &\le \Delta_1(\boldsymbol{H}_3) \le \max_{t_2 \le v \le t_3}\{d^m_{t_3}(v)\} \le t_3^{p/2+\varepsilon}/t_2^{p/2} \le t^{15/64}. \end{aligned}$$

To bound $\lambda_1(\boldsymbol{H}_{ij})$, start with $m = 1$. For $i < j$, this implies that each vertex in S_j has at most one edge in $\boldsymbol{H}_{ij}$, that is H_{ij} is a star forest. Then we have a bound on $\boldsymbol{H}_{ij}$ by Claim A.8. For $m > 1$, let G' be one of our generated graphs with t edges and $m = 1$. Think now of contracting vertices in G' (only the ones added using PA) into a single vertex. We can write $\boldsymbol{A}_G$ in terms of $\boldsymbol{A}'_G$: $\boldsymbol{A}_G = \boldsymbol{C}^\top \boldsymbol{A}'_G \boldsymbol{C}$, where $\boldsymbol{C}$ is a contraction matrix with t rows and the number of columns equal to the number of vertices in $\boldsymbol{A}_G$ (at most t/m). The ith column is equal to 1 at indices j in which (i,j) are identified. Similarly, we can write $\boldsymbol{H}_{ij}$ in terms of $\boldsymbol{H}'_{ij}$.

Note that if $\mathrm{y} = \boldsymbol{C}\mathrm{x}$, then $\mathrm{y}^\top\mathrm{y} = \mathrm{x}^\top\boldsymbol{C}^\top\boldsymbol{C}\mathrm{x}$, where $\boldsymbol{C}^\top\boldsymbol{C}$ is a diagonal matrix with $1's$ and m's on the diagonal. So $\mathrm{x}^\top\mathrm{x} \le \mathrm{y}^\top\mathrm{y} \le m\mathrm{x}^\top\mathrm{x}$ which we use to bound $\lambda_1(H_{ij})$ as

$$\begin{aligned} \lambda_1(\boldsymbol{H}_{ij}) &= \max_{\mathrm{x}\ne 0} \frac{\mathrm{x}^\top\boldsymbol{H}_{ij}\mathrm{x}}{\mathrm{x}^\top\mathrm{x}} = \max_{\mathrm{x}\ne 0} \frac{\mathrm{x}^\top\boldsymbol{C}^\top\boldsymbol{H}'_{ij}\boldsymbol{C}\mathrm{x}}{\mathrm{x}^\top\mathrm{x}} = \max_{\mathrm{x}\ne 0, \mathrm{y}=\boldsymbol{C}\mathrm{x}} \frac{\mathrm{y}^\top\boldsymbol{H}'_{ij}\mathrm{y}}{\mathrm{x}^\top\mathrm{x}} \\ &= \max_{\mathrm{x}\ne 0, \mathrm{y}=\boldsymbol{C}\mathrm{x}} \frac{m\mathrm{y}^\top\boldsymbol{H}'_{ij}\mathrm{y}}{m\mathrm{x}^\top\mathrm{x}} \le \max_{\mathrm{x}\ne 0, \mathrm{y}=\boldsymbol{C}\mathrm{x}} \frac{m\mathrm{y}^\top\boldsymbol{H}'_{ij}\mathrm{y}}{\mathrm{y}^\top\mathrm{y}}. \end{aligned} \tag{A.7}$$

Now using Claim A.6 with $f(t) = t_1$ and $\varepsilon = 1/64$,

$$\begin{aligned} \Delta_1(\boldsymbol{H}'_{12}) &= \max_{\mathrm{v} \le t_2}\{d'_{t_2}(v)\} \le t_2^{p/2+\varepsilon} \le t^{297/1024} \\ \Delta_1(\boldsymbol{H}'_{23}) &= \max_{t_1 \le v \le t_3}\{d'_{t_3}(v)\} \le t_3^{p/2+\varepsilon}/t_1^{p/2} \le t^{29/64}. \end{aligned} \tag{A.8}$$

Finally, all edges in $\boldsymbol{H}'_{13}$ are between S_1 and S'_3, so Claim A.7 shows $\Delta_1(\boldsymbol{H}'_{13}) \le t^{p-9/16} \le t^{7/16}$ whp. Putting together Equations (A.7) and (A.8), we get $\lambda_1(\boldsymbol{H}_{ij}) \le m\lambda_1(\boldsymbol{H}'_{ij}) \le m\Delta_1(\boldsymbol{H}'_{ij})^{1/2} \le mt^{15/64}$. And so we get the final bound

$$\lambda_1(\boldsymbol{A}_H) \le \sum_{i=1}^{3} \lambda_1(\boldsymbol{H}_i) + \sum_{i<j} \lambda_1(\boldsymbol{H}_{ij}) \le 6mt^{15/64}.$$

This shows that $\lambda_i(\boldsymbol{A}_H)$ is $o(\lambda_k(\boldsymbol{A}_F))$, which implies $\lambda_i(\boldsymbol{A}_G) = \lambda_i(\boldsymbol{A}_G) = \lambda_i(\boldsymbol{A}_F)(1 \pm o(1))$. This concludes the final claim. □

And this concludes the proof. □

Appendix B. TGPA spectral power law proofs

In this appendix, we provide proofs for theoretical results from Section 6.

Proof of Lemma 6.1. Denote G_t^m as the graph at time t with contractions of size m. Let $Z_t = d_t^m(s)$ be the degree of vertex s at time t. Let Y_t be an indicator for the event that only one edge added at time t is incident to s, and let X_t be an indicator variable for the event that both of the edges added at time t are incident to s. First, let us calculate the probability of placing exactly one edge incident to node s at time t:

$$p\left[\underbrace{\frac{d_t(s)}{4t-2}\left(1-\frac{2(\text{num of self-loops})}{d_t(s)}\right)}_{\text{probability of picking s first, and then not picking it second}} + \underbrace{\frac{\sum_{u\in N(s),u\neq s} d_t(u)}{4t-2}\left(\frac{\text{num edges btwn u, s}}{d_t(u)}\right)}_{\text{probability of not picking s first, but picking a neighbour, and then picking s second}}\right]$$

$$= p\left[\frac{d_t(s)}{2t-1} - \frac{2(\text{num self-loops})}{2t-1}\right].$$

Also the probability of placing two (both) edges incident to node s at time t:

$$p\frac{d_t(s)}{4t-2}\cdot \Pr\left[\text{picking it second} \mid \text{picked it first}\right] = \frac{d_t(s)}{4t-2}\cdot\frac{2(\text{num self-loops})}{d_t(s)} = \frac{\text{num self-loops}}{2t-1}.$$

Then, we can write the expectation of Z_t in terms of Z_{t-1} using the above calculations:

$$\begin{aligned}
\mathbb{E}[Z_t^{(\bar{k})}] &= \mathbb{E}[\mathbb{E}[(Z_{t-1}+Y_t+2X_t)^{(\bar{k})}]|Z_{t-1}] \\
&= \mathbb{E}\left[(Z_{t-1}+2)^{(\bar{k})}p\cdot\left(\frac{\text{num self-loops}}{2t-1}\right) + (Z_{t-1}+1)^{(\bar{k})}p\cdot\left(\frac{d_t(s)}{2t-1}-\frac{2(\text{num self-loops})}{2t-1}\right)\right. \\
&\qquad \left. + Z_{t-1}^{(\bar{k})}\left(1-p\frac{d_t(s)}{2t-1}+p\frac{\text{num self-loops}}{2t-1}\right)\right] \\
&\le \mathbb{E}\left[(Z_{t-1}+2)^{(\bar{k})}p\cdot\left(\frac{\text{num self-loops}}{2t-1}\right) + (Z_{t-1}+2)^{(\bar{k})}p\cdot\left(\frac{d_t(s)}{2t-1}-\frac{2(\text{num self-loops})}{2t-1}\right)\right.
\end{aligned}$$

$$
\begin{aligned}
&+Z_{t-1}^{(\bar{k})}\left(1-p\frac{d_t(s)}{2t-1}+p\frac{\text{num self-loops}}{2t-1}\right)\Bigg]\\
&=\mathbb{E}\left[(Z_{t-1}+2)^{(\bar{k})}p\cdot\left(\frac{d_t(s)}{2t-1}-\frac{\text{num self-loops}}{2t-1}\right)+Z_{t-1}^{(\bar{k})}\left(1-p\frac{d_t(s)}{2t-1}+p\frac{\text{num self-loops}}{2t-1}\right)\right]\\
&=\mathbb{E}\left[Z_{t-1}^{(\bar{k})}\left(1-p\frac{Z_{t-1}}{2t-1}+p\frac{\text{num self-loops}}{2t-1}+\frac{Z_{t-1}+k}{Z_{t-1}}\left(p\frac{Z_{t-1}}{2t-1}-p\frac{\text{num self-loops}}{2t-1}\right)\right)\right]\\
&=\mathbb{E}\left[Z_{t-1}^{(\bar{k})}\left(1-p\frac{Z_{t-1}-\text{num self-loops}}{2t-1}+\frac{Z_{t-1}+k}{Z_{t-1}}\left(p\frac{Z_{t-1}-\text{num self-loops}}{2t-1}\right)\right)\right]\\
&\leq\mathbb{E}\left[Z_{t-1}^{(\bar{k})}\left(1-p\frac{Z_{t-1}-\text{num self-loops}}{2t-1}+p\frac{Z_{t-1}+k}{2t-1}\right)\right]\\
&=\mathbb{E}\left[Z_{t-1}^{(\bar{k})}\left(1+\frac{p}{2t-1}(k+\text{num self-loops})\right)\right].
\end{aligned}
$$

Now if $k \geq$ number self-loops we can move on to:

$$
\mathbb{E}[Z_t^{(\bar{k})}]\leq\mathbb{E}\left[Z_{t-1}^{(\bar{k})}(1+\frac{2pk}{2t-1}).\right] \tag{B.1}
$$

Apply this relationship iteratively, down to the time when node s was added (recall we denoted that time as s'). Also note that the degree of s at time s' is bounded by $4m$ (if all m edges were added as self-loops). Thus:

$$
\mathbb{E}(Z_t^{(\bar{k})})=\prod_{t'=s'}^{t}\left(1+\frac{2pk}{2t'-1}\right)\leq(4m)^{(\bar{k})}\prod_{t'=s'+1}^{t}\left(1+\frac{2pk}{2t'-1}\right). \tag{B.2}
$$

Use $1+x\leq e^x$ to write the product as a sum, and bound the sum with an integral:

$$
\sum_{t'=s'+1}^{t}\frac{1}{t'-1/2}\leq\int_{x=s'}^{t}\frac{1}{x-1/2}\ \mathrm{dx}=\log\frac{t-1/2}{s'-1/2}. \tag{B.3}
$$

So finally,

$$
\begin{aligned}
\mathbb{E}(Z_t^{(\bar{k})})&\leq(4m)^{(\bar{k})}\left(\frac{t-1/2}{s'-1/2}\right)^{pk}\\
&=(4m)^{(\bar{k})}\left(\frac{t}{s'}\right)^{pk/2}\left(\frac{2-1/t}{2-1/s'}\right)^{pk}\\
&\leq(4m)^{(\bar{k})}\left(\frac{t}{s'}\right)^{pk}2^{pk}.
\end{aligned} \tag{B.4}
$$

□

Proof of Lemma 6.2. Let $\tau^{(i)} = (\tau_1^{(i)}, \ldots, \tau_{r_i}^{(i)})$, where $\tau_j^{(i)}$ is the time when we add an edge incident to S_i and increase the degree of S_i from $d_i + j - 1$ to $d_i + j$. Note that we could have repeated times in $\tau^{(i)}$. Define $\tau = (\tau_0, \tau_1, \ldots, \tau_{r+1})$ to be the ordered union of $\tau^{(i)}$, with $\tau_0 = t_0$ and $\tau_{r+1} = t$. Again, there may be up to two moves per time step. Let $p(\tau; \mathrm{d}, t_0, t)$ be the probability that supernodes S_i increase in degree at exactly the times specified by τ between time t_0 and t. Define all time steps to be $T := \{t_0, t_1, t_1, t_2, t_2, \ldots, t_r, t_r\}$. Time steps involving the sets S_i we defined to be τ. So the remaining time steps are $T - \tau$. Then

$$p(\tau; \mathrm{d}, t_0, t) \leq \underbrace{\left(\prod_{i=1}^{l}\prod_{k=1}^{r_i} 2p\frac{d_i + k - 1}{4\tau_k^{(i)} - 2}\right)}_{\substack{\text{for each supernode } S_i\text{, the prob.}\\ \text{of } \tau \text{ aligning with } \tau^{(i)}.}} \underbrace{\left(\prod_{k=0}^{r}\prod_{\substack{j\in T-\tau\\ j\geq\tau_k\\ j<\tau_{k+1}}}\left(1 - 2p\frac{d+k}{4j-2}\right)\right)}_{\substack{\text{for each time step in between the relevant ones,}\\ \text{the probability of picking any edge outside of}\\ S_1, \ldots, S_l.}} \tag{B.5}$$

$$= \left(\prod_{i=1}^{l}\frac{(r_i + d_i - 1)!}{(d_i - 1)!}\right)\left(\prod_{k=1}^{r}\frac{p}{2\tau_k - 1}\right)\exp\left\{\sum_{k=0}^{r}\sum_{\substack{j\in T-\tau\\ j\geq\tau_k\\ j<\tau_{k+1}}}\log\left(1 - p\left(\frac{d+k}{2j-1}\right)\right)\right\}.$$

Now we can bound the inner most sum of the exponential term as

$$\sum_{\substack{j\in T-\tau\\ j\geq\tau_k\\ j<\tau_{k+1}}}\log\left(1 - p\left(\frac{d+k}{2j-1}\right)\right) \leq 2\sum_{j=\tau_k+1}^{\tau_{k+1}-1}\log\left(1 - \frac{p(d+k)}{2j}\right), \tag{B.6}$$

which is less than or equal to

$$\begin{aligned} 2\int_{\tau_k+1}^{\tau_{k+1}}\log\left(1 - \frac{p(d+k)}{2x}\right)\,\mathrm{d}x = &-2\tau_{k+1}\log(2\tau_{k+1}) + 2(\tau_k + 1)\log(2\tau_k + 2)\\ &+ (2\tau_{k+1} - p(d+k))\log(2\tau_{k+1} - p(d+k))\\ &- (2\tau_k + 2 - p(d+k))\log(2\tau_k + 2 - p(d+k)). \end{aligned} \tag{B.7}$$

Note that $\tau_0 = t_0$ and $\tau_{r+1} = t$. Then we can write

$$\sum_{k=0}^{r}\int_{\tau_k+1}^{\tau_{k+1}}\log\left(1 - \frac{p(d+k)}{2x}\right)\,\mathrm{d}x = A + \sum_{k=1}^{r}B_k, \tag{B.8}$$

where

$$\begin{aligned} A = &\,2(t_0 + 1)\log(2t_0 + 2) - (2t_0 + 2 - pd)\log(2t_0 + 2 - pd)\\ &- 2t\log(2t) + (2t - p(d+r))\log(2t - p(d+r)) \end{aligned} \tag{B.9}$$

and

$$B_k = 2\tau_k \log(1 + 1/\tau_k) + 2\log(2\tau_k + 2) - (2-p)\log(2\tau_k + p - p(d+k)) + (2\tau_k + 2 - p(d+k))\log\left(1 - \frac{2-p}{2\tau_k + 2 - p(d+k)}\right). \tag{B.10}$$

We will bound each of A and B_k, starting with B_k. Since $1 + x \le e^x$, $1\tau_k \log(1 + 1/\tau_k) \le 2$ and $(2\tau_k + 2 - p(d+k))\log\left(1 - \frac{2-p}{2\tau_k+2-p(d+k)}\right) \le p - 2$. Rearranging the other two terms of Equation (B.10) and combining with these inequalities we get

$$B_k \le p\log(2\tau_k + 2) - (2-p)\log\left(1 - \frac{p(d+k) + 2 - p}{2\tau_k + 2}\right) + p. \tag{B.11}$$

Now rearranging terms of A from Equation (B.9),

$$\begin{aligned} A &= -2(t_0+1)\log\left(1 - \frac{pd}{2t_0+2}\right) + pd\log(2t_0 + 2 - pd) \\ &\quad + 2t\log\left(1 - \frac{p(d+r)}{2t}\right) - p(d+r)\log(2t - p(d+r)) \\ e^A &= \left(1 - \frac{pd}{2t_0+2}\right)^{-2(t_0+1)} (2t_0 + 2 - pd)^{pd} \left(1 - \frac{p(d+r)}{2t}\right)^{2t} (2t - p(d+r))^{-p(d+r)} \\ &= \left(1 - \frac{pd}{2t_0+2}\right)^{-2(1-\frac{pd}{2(t_0+1)})(t_0+1)} \left(1 - \frac{p(d+r)}{2t}\right)^{2t-p(d+r)} \left(\frac{t_0+1}{t}\right)^{pd} (2t)^{-pr}. \end{aligned} \tag{B.12}$$

Using the bound $1 - x \le e^{-x-x^2/2}$ for $0 < x < 1$,

$$\left(1 - \frac{p(d+r)}{2t}\right)^{2t-p(d+r)} \le \exp\left\{-p(d+r) + \frac{p^2(d+r)^2}{4t} + \frac{p^3(d+r)^3}{8t^2}\right\}. \tag{B.13}$$

Putting the bounds on A and B_k together, we get

$$\begin{aligned} e^{A+\sum B_k} &\le \left(1 - \frac{pd}{2t_0+2}\right)^{-2(1-\frac{pd}{2(t_0+1)})(t_0+1)} \exp\left\{-p(d+r) + \frac{p^2(d+r)^2}{4t} + \frac{p^3(d+r)^3}{8t^2}\right\} \\ &\quad \times \left(\frac{t_0+1}{t}\right)^{pd} (2t)^{-pr} \prod_{k=1}^{r} \left(\left(1 - \frac{p(d+k)+2-p}{2\tau_k+2}\right)^{-(2-p)} (2\tau_k+2)^p\right) e^{pr}. \end{aligned} \tag{B.14}$$

Using

$$\mathrm{err}(r,d,t_0,t) = \left(1 - \frac{pd}{2t_0+2}\right)^{-2(1-\frac{pd}{2(t_0+1)})(t_0+1)} \exp\left\{-pd + \frac{p^2(d+r)^2}{4t} + \frac{p^3(d+r)^3}{8t^2}\right\}, \tag{B.15}$$

we can write Equation (B.14) as

$$e^{A+\sum B_k} \leq \mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd}(2t)^{-pr}\prod_{k=1}^{r}\left(\left(1-\frac{p(d+k)+2-p}{2\tau_k+2}\right)^{-(2-p)}(2\tau_k+2)^p\right). \quad \text{(B.16)}$$

So, we finally finish with the bound on $p(\tau;\mathrm{d},t_0,t)$ by substituting Equation (B.16) into Equation (B.5):

$$\begin{aligned} p(\tau;\mathrm{d},t_0,t) \leq &\left(\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right)\mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd}(2t)^{-pr} \\ &\times\prod_{k=1}^{r}\left(\left(1-\frac{p(d+k)+2-p}{2\tau_k+2}\right)^{-(2-p)}(2\tau_k+2)^p\frac{p}{2\tau_k-1}\right), \end{aligned} \quad \text{(B.17)}$$

which can be rearranged as

$$\begin{aligned} &=\left(\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\right)\mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd}(2t)^{-pr} \\ &\times\prod_{k=1}^{r}\left(p(2\tau_k+p-p(d+k))^{-(2-p)}\left(2\tau_k+5+\frac{9}{2\tau_k-1}\right)\right). \end{aligned} \quad \text{(B.18)}$$

Now we will sum $p(\tau;\mathrm{d},t_0,t)$ over all ordered choices of τ:

$$\begin{aligned} p(\mathrm{r};\mathrm{d},t_0,t) &\leq \sum_{\tau^{(1)},\ldots,\tau^{(l)}} p(\tau;\mathrm{d},t_0,t) \\ &\leq \binom{r}{r_1,\ldots,r_l}\sum_{t_0+1\leq\tau_1<\ldots<\tau_r\leq t}\prod_{i=1}^{l}\frac{(r_i+d_i-1)!}{(d_i-1)!}\mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd} \\ &\quad\times(2t)^{-pr}p\prod_{k=1}^{r}(2\tau_k+p-p(d+k))^{-(2-p)}\left(2\tau_k+5+\frac{9}{2\tau_k-1}\right) \\ &= r!\left(\prod_{i=1}^{l}\binom{r_i+d_i-1}{d_i-1}\right)\mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd}(2t)^{-pr} \\ &\quad\times\sum_{t_0+1\leq\tau_1<\ldots<\tau_r\leq t}p\prod_{k=1}^{r}(2\tau_k+p-p(d+k))^{-(2-p)}\left(2\tau_k+5+\frac{9}{2\tau_k-1}\right). \end{aligned} \quad \text{(B.19)}$$

Now let $\tau_k' = \tau_k - \lceil p(d+k)/2\rceil$. Since $d \geq 1$ and $k \geq 1$, we have $2\lceil p(d+k)/2\rceil \geq 2$. Then the last term in Equation (B.19) is less than or equal to

$$\sum_{(t_0-p\lceil d/2\rceil+1)\leq\tau_1'\leq\ldots\leq\tau_r'\leq(t-p\lceil(d+r)/2\rceil)}\left(p\prod_{k=1}^{r}(2\tau_k'+p)^{-(2-p)}\left(2\tau_k'+5+\frac{9}{2\tau_k'+1}\right)\right)$$

$$\le \frac{p}{r!}\left(\sum_{\tau'=(t_0-p\lceil d/2\rceil+1)}^{t-p\lceil (d+r)/2\rceil}\left(9(2\tau'+p)^{-(3-p)}+(2\tau_k'+5)(2\tau_k'+p)^{-(2-p)}\right)\right)^r$$
$$\le \frac{p}{r!}\left(\int_0^{t-p(d+r)/2}\left(9(2x+p)^{-(3-p)}+(2\tau_k'+5)(2x+1)^{-(2-p)}\right)\mathrm{d}x\right)^r$$
$$\le \frac{p}{r!}\left(\frac{9}{2(2-p)p^{2-p}}+\frac{5}{(1-p)p^{1-p}}+\frac{2}{(1-p)p}(2t-p(d+r)+1)^p\right)^r \tag{B.20}$$
$$\le \frac{2p}{r!(1-p)p^{2-p}}\left(\frac{19}{4}+(2t-p(d+r)+p)^p\right)^r$$
$$=\frac{2}{r!(1-p)p^{1-p}}\left((2t)^p\left(1-\frac{p(d+r)-p}{2t}\right)^p\left(1+\frac{19p}{4(2t-p(d+r)+p)^p}\right)\right)^r$$
$$\le= \frac{2}{r!(1-p)p^{1-p}}(2t)^{pr}\underbrace{\left(1-\frac{p(d+r)-p}{2t}\right)^{pr}}_{\le\exp\left\{-\frac{pr(p(d+r)-p)}{2t}\right\}}\underbrace{\left(1+\frac{19p}{4(2t-p(d+r)+p)^p}\right)^r}_{\le\exp\left\{\frac{19pr}{4(2t-p(d+r)+p)^p}\right\}},$$

where the last inequalities come from $1+x\le e^x$. So finally,

$$p_S(\mathrm{r};\mathrm{d},t_0,t)\le\left(\prod_{i=1}^{l}\binom{r_i+d_i-1}{d_i-1}\right)\mathrm{err}(r,d,t_0,t)\left(\frac{t_0+1}{t}\right)^{pd}$$
$$\times\exp\left\{\frac{-pr((d+r)-p)}{2t}+\frac{19pr}{4(2t-p(d+r)+p)^p}\right\}.$$

Since $d=o(t^{1/2})$ and $r=o(t^{2/3})$,

$$\mathrm{err}(r,d,t_0,t)\exp\left\{\frac{-rp((d+r)-p)}{2t}+\frac{19pr}{4(2t-p(d+r)+p)^p}\right\}$$
$$\le\left(1-\frac{pd}{2(t_0+1)}\right)^{-2(1-pd/2(t_0+1))(t_0+1)}\exp\left\{1-pd-\frac{r^2}{4t}+\frac{19pr}{4t^p}\right\}$$
$$\le \underbrace{e^{2(t_0+1)}}_{\text{since } x^{-x}\le e}\exp\left\{1-pd+\frac{19pr}{4t^p}\right\}$$
$$=\exp\left\{3+2t_0-pd+\frac{19pr}{4t^p}\right\}$$

This concludes the proof. □

Proof of Theorem 6.1. Partition the vertices into those added before time t_0, before time t_1 and after t_1, with $t_0=\log\log\log f(t), t_1=\log\log f(t)$. We will argue about the maximum degree vertices in each set.

CLAIM B.1 In G_t^m, the degree of the supernode of vertices added before time t_0 is at least $t_0^{(1-p)/2}t^p$ whp.

Proof. Consider all vertices added before time t_0 as a supernode. Let A_1 denote the event that this supernode has degree less than $t_0^{(1-p)/2}t^p$ at time t. We will use Lemma 6.2 with $l = 1$, and $d = 4t_0$ (because the supernode has all edges at time t_0). Calculate $Pr[A_1]$ as

$$
\begin{aligned}
Pr[A_1] &\leq \sum_{r_1=0}^{t_0^{(1-p)/2}t^p - 4t_0} \binom{r_1 + 4t_0 - 1}{4t_0 - 1} \left(\frac{t_0+1}{t}\right)^{pd} e^{3+2t_0 - pd + 19pr/4t^p} \\
&\leq \underbrace{\sum_{r_1=0}^{t_0^{(1-p)/2}t^p - 4t_0} \binom{t_0^{(1-p)/2}t^p - 1}{4t_0 - 1}}_{\text{By substituting } r_1 = t_0^{(1-p)/2}t^p} \left(\frac{t_0+1}{t}\right)^{4pt_0} \underbrace{e^{3+2t_0 - 4pt_0 + (19/4)pt_0^{1/3} - \frac{19pt_0}{t^p}}}_{\text{because } r = r_1 \text{ and } d = 4t_0} \\
&= (t_0^{(1-p)/2}t^p - 4t_0) \frac{(t_0^{(1-p)/2}t^p - 1)!}{(4t_0-1)!(t_0^{(1-p)/2}t^p - 4t_0)!} \left(\frac{t_0+1}{t}\right)^{4pt_0} e^{3+2t_0(1-2p)+(19/4)pt_0^{1/3} - \frac{19pt_0}{t^p}} \qquad \text{(B.21)} \\
&\leq t_0^{(1-p)/2}t^p \frac{(t_0^{(1-p)/2}t^p)^{4t_0-1}}{(4t_0-1)!} \left(\frac{t_0+1}{t}\right)^{4pt_0} e^{3+2t_0(1-2p)+(19/4)pt_0^{1/3} - 19pt_0/t^p} \\
&\leq t_0^{2(1-p)t_0} \underbrace{\frac{e^{4t_0-1}}{(4t_0-1)^{4t_0-1}}}_{\text{since } 1/x! \leq e^x/x^x} (t_0+1)^{4pt_0} e^{3+2t_0(1-2p)+(19/4)pt_0^{1/3} - 19pt_0/t^p} \\
&\leq \frac{e^{2+2t_0(3-2p)+(19/4)pt_0^{1/3} - 19pt_0/t^p}}{(4t_0-1)^{2t_0(1-p)-1}},
\end{aligned}
$$

which goes to 0 as t goes to infinity. Thus A_1 does not hold with high probability, and the claim is proved.□

CLAIM B.2 In G_t^m, no vertex added after time t_1 has degree exceeding $t_0^{-2}t^p$ whp.

Proof. Let A_2 denote the event that some vertex added after time t_1 has degree exceeding $t_0^{-2}t^p$. Bound $Pr[A_2]$ as

$$
\begin{aligned}
Pr[A_2] &\leq \sum_{s=t_1}^{t} \Pr[d_t(s) \geq t_0^{-2}t^p] = \sum_{s=t_1}^{t} \Pr[(d_t(s))^{(\bar{l})} \geq (t_0^{-2}t^p)^{(\bar{l})}] \leq \underbrace{\sum_{s=t_1}^{t} t_0^{2l} t^{-lp} \mathbb{E}[(d_t(s))^{(\bar{l})}]}_{\text{by Markov}} \\
&= \underbrace{\sum_{s=t_1}^{t} t_0^{2l} t^{-lp} (4m)^{(\bar{l})} 2^{lp} \left(\frac{t}{s}\right)^{lp}}_{\text{by Lemma 6.1}} = 2^{lp}(4m)^{(\bar{l})} t_0^{2l} \int_{t_1-1}^{t} x^{-lp}\, \mathrm{dx},
\end{aligned} \qquad \text{(B.22)}
$$

and compute the integral in Equation (B.22),

$$
\int_{t_1-1}^{t} x^{-lp}\, \mathrm{dx} = \left.\frac{x^{-lp+1}}{-lp+1}\right|_{t_1-1}^{t} = (-lp+1)^{-1}\left(t^{-lp+1} - (t_1-1)^{-lp+1}\right). \qquad \text{(B.23)}
$$

We want to choose l so that $-lp+1$ is less than 0. So choose $l > 1/p$. Then the integral in Equation (B.23) is less than or equal to $(lp-1)^{-1}(t_1-1)^{-lp+1}$, and plugging in the computation from Equation (B.23) into Equation (B.22),

$$Pr[A_2] \leq \frac{2^{lp}(4m)^{(\bar{l})}t_0^{2l}}{(lp-1)(t_1-1)^{lp-1}}, \tag{B.24}$$

which goes to 0 as t increases. □

CLAIM B.3 In G_t^m, no vertex added before time t_1 has degree exceeding $t_0^{(1-p)/4}t^p$ whp.

Proof. Let A_3 denote the event that some vertex added before time t_1 has degree exceeding $t_0^{(1-p)/4}t^p$. Using the exact argument as in Claim B.2, Pr$[A_3]$ goes to 0 as t increases. □

CLAIM B.4 The k highest degree vertices of G_t^m are added before time t_1 and have degree Δ_i bounded by $t_0^{-1}t^p \leq \Delta_i \leq t_0^{(1-p)/4}t^p$.

Proof. First lets summarize the results of the last three claims:

- Bound on degrees of vertices added after time t_1: $t_0^{-2}t^p$
- Bound on degrees of vertices added before time t_1: $t_0^{(1-p)/4}t^p$
- Sum of all degrees added before time t_0 is at least: $t_0^{(1-p)/2}t^p$

So the upper bound of the claim is immediately clear from the second item. Suppose that the lower bound does not hold. Then one of the top k vertices has degree less than $t_0^{-1}t^p$ and the total degree of vertices added before time t_0 is bounded by

$$\underbrace{(k-1)t_0^{(1-p)/4}t^p}_{\text{largest possible degrees of }(k-1)\text{ vertices}} + \underbrace{\left(\frac{t_0}{m}-k+1\right)\left(t_0^{-1}t^p\right)}_{\text{largest possible degrees of remaining vertices}} \leq kt_0^{(1-p)/4}t^p + t_0(t_0^{-1}t^p) \tag{B.25}$$
$$= kt_0^{(1-p)/4}t^p + t^p = t^p(kt_0^{(1-p)/4}+1) \leq t^p(2kt_0^{(1-p)/4}) \leq t^p t_0^{(1-p)/2},$$

which contradicts the third bulleted item. Finally, since we have the lower bound, and we know that $t^p/t_0 \geq t^p/t_0^2$, then none of the largest degree vertices could be added after time t_1. □

CLAIM B.5 The k highest degree vertices of G_t^m have $\Delta_i \leq \Delta_{i-1} - \frac{t^p}{f(t)}$ whp.

Proof. Let A_4 denote the event that there are two vertices among the first t_1 time steps with degrees exceeding $t_0^{-1}t^p$ and within $t^p/f(t)$ of each other. Define

$$p_{l,s_1,s_2} = \Pr\left[d_t(s_1) - d_t(s_2) = l \mid \overline{A}_3\right], \text{for } |l| \leq t^p/f(t),$$

where $\overline{A}_3$ is defined to be the opposite of event A_3 from Claim B.3. Then

$$\Pr\left[A_4|\overline{A}_3\right] \leq \sum_{1\leq s_1<s_2\leq t_1} \sum_{l=-t^p/f(t)}^{t^p/f(t)} p_{l,s_1,s_2}. \tag{B.26}$$

Now

$$p_{l,s_1,s_2} \leq \sum_{r_1=t_0^{-1}t^p}^{t_0^{(1-p)/4}t^p} \sum_{d_1,d_2=1}^{4t_1} \underbrace{p_{(s_1,s_2)}\left((r_1, r_1-l);(d_1,d_2),t_1,t\right)}_{\text{Notation from Lemma 6.2.}} \tag{B.27}$$

Using Lemma 6.2,

$$\begin{aligned} \leq \sum_{r_1=t_0^{-1}t^p}^{t_0^{(1-p)/4}t^p} \sum_{d_1,d_2=1}^{4t_1} \binom{r_1+d_1-1}{d_1-1}\binom{r_1-l+d_2-1}{d_2-1}\left(\frac{t_1+1}{t}\right)^{p(d_1+d_2)} \\ \times e^{\left\{3+2t_1-p(d_1+d_2)+\frac{19p(r_1-l)}{4t^p}\right\}} \end{aligned} \tag{B.28}$$

$$\begin{aligned} &\leq t_0^{(1-p)/4}t^p \sum_{d_1,d_2=1}^{4t_1} \binom{t_0^{(1-p)/4}t^p+d_1-1}{d_1-1}\binom{t_0^{(1-p)/4}t^p-l+d_2-1}{d_2-1}\left(\frac{t_1+1}{t}\right)^{p(d_1+d_2)} \\ &\qquad \times e^{\left\{3+2t_1-p(d_1+d_2)+\frac{19pt_0^{(1-p)/4}t^p}{4t^p}\right\}} \\ &\leq t_0^{(1-p)/4}t^p \sum_{d_1,d_2=1}^{4t_1} \binom{2t_0^{(1-p)/4}t^p}{d_1-1}\binom{2t_0^{(1-p)/4}t^p}{d_2-1}\left(\frac{t_1+1}{t}\right)^{p(d_1+d_2)} \\ &\qquad \times e^{3+2t_1+(19/4)pt_0^{(1-p)/4}} \\ &\leq t_0^{(1-p)/4}t^p \sum_{d_1,d_2=1}^{4t_1} (2t_0^{(1-p)/4}t^p)^{d_1+d_2-2}(t_1+1)^{8pt_1}t^{-p(d_1+d_2)}e^{3+2t_1+(19/4)pt_0^{(1-p)/4}} \\ &= t^{-p}t_0^{(1-p)/4}(4t_1)^2 2^{8t_1} t_0^{2t_1(1-p)}(t_1+1)^{8pt_1}e^{3+2t_1+(19/4)pt_0^{(1-p)/4}}. \end{aligned} \tag{B.29}$$

Denote the last equation as $h(t)$ and note $h(t)$ is a polynomial in $\log(f(t))$ times a factor of t^{-p}. Then going back to Equation (B.26),

$$\Pr\left[A_4|\overline{A}_3\right] \leq \binom{t_1}{2}2\frac{t^p}{f(t)}h(t) = \binom{t_1}{2}2\frac{\text{poly}(\log(f(t)))}{f(t)}, \tag{B.30}$$

which goes to 0 as t increases. This concludes the proof of this final claim. □

And this concludes the proof of the theorem. □

Proof of Theorem 6.2. Let S_i be the vertices added after time t_{i-1} and at or before time t_i, for $t_0 = 0,\ t_1 = t^{1/8},\ t_2 = t^{9/16},\ t_3 = t$. Let $G = G_t$. We start by finding bounds on the degrees and co-degrees of G.

CLAIM B.6 For any $\varepsilon > 0$, and any $f(t)$ with $f(t) \to \infty$ as $t \to \infty$ the following holds whp: for all s with $f(t) \leq s \leq t$, for all vertices $v \in G_s$, if v was added at time r, then $d_s(v) \leq s^{p+\varepsilon} r^{-p}$.

Proof.

$$\begin{aligned}\Pr\left[\cup_{s=f(t)}^{t} \cup_{r=1}^{s} \left\{d_s^m(r) \geq s^{p+\varepsilon} r^{-p}\right\}\right] &\leq \sum_{s=f(t)}^{t} \sum_{r=1}^{s} \Pr\left[d_s^m(r) \geq s^{p+\varepsilon} r^{-p}\right] \\ &= \sum_{s=f(t)}^{t} \sum_{r=1}^{s} \Pr\left[(d_s^m(r))^{(\bar{l})} \geq (s^{p+\varepsilon} r^{-p})^{(\bar{l})}\right]\end{aligned} \tag{B.31}$$

which is bounded using Markov:

$$\leq \sum_{s=f(t)}^{t} \sum_{r=1}^{s} s^{-l(p+\varepsilon)} r^{pl} \mathbb{E}\left[(d_s^m(r))^{(\bar{l})}\right].$$

We can bound the preceding equation using Lemma 6.1

$$\leq \sum_{s=f(t)}^{t} \sum_{r=1}^{s} s^{-l(p+\varepsilon)} r^{lp} (4m)^{(\bar{l})} 2^{lp} \left(\frac{s}{r}\right)^{lp} = (4m)^{(l)} 2^{lp} \sum_{s=f(t)}^{t} s^{1-\varepsilon l}.$$

Take $l \geq 3/\varepsilon$. Then we can bound the sum by an integral,

$$\sum_{s=f(t)}^{t} s^{1-\varepsilon l} \leq \int_{f(t)-1}^{\infty} x^{1-\varepsilon l}\, \mathrm{dx} = \frac{1}{2-\varepsilon l} x^{2-\varepsilon l}\Big|_{f(t)-1}^{\infty} = \frac{1}{\varepsilon l - 2} (f(t)-1)^{2-\varepsilon l}, \tag{B.32}$$

which goes to zero as t increases, since $l \geq 3/\varepsilon$. □

CLAIM B.7 Let S_3' be the set of vertices in S_3 that are adjacent to more than one vertex of S_1 in G. Then $|S_3'| \leq t^{2p-9/16}$ with high probability.

Proof. Let B_1 be the event that the conditions of Claim B.6 hold with $f(t) = t_2$ and $\varepsilon = 1/16$. Then for a vertex $v \in S_3$ added at time s, the probability that v picks at least one neighbour in S_1 is less than or equal to

$$\frac{2\sum_{w \in S_1} d_s(w)}{4s-2} \leq \frac{\sum_{w \in S_1} s^{p+\varepsilon}}{2s-1} = \frac{t_1 s^{p+\varepsilon}}{2s-1}. \tag{B.33}$$

So, the probability of having two or more neighbours in S_1 can be bounded by

$$\Pr\left[\, |N(v) \cap S_1| \geq 2 \mid B_1\right] \leq \left(\frac{t_1 s^{p+\varepsilon}}{2s-1}\right)^2 \cdot \binom{2m}{2} \leq m^2 t^{1/4} s^{(16p-15)/8}. \tag{B.34}$$

Let X denote the number of $v \in S_3$ adjacent to more than one vertex of S_1. Then

$$\begin{aligned}\mathbb{E}[X|B_1] &\leq \sum_{t_2+1}^{t} m^2 s^{(-15+16p)/8} t^{1/4} \leq m^2 t^{1/4} \int_{t_2}^{t} x^{(-15+16p)/8} \, \mathrm{dx} \\ &= m^2 t^{1/4} \left[\frac{8}{-7+16p} x^{(-7+16p)/8} \Big|_{t_2}^{t} \right] \leq \frac{8m^2 t^{1/4}}{-7+16p} t^{(-7+16p)/8}.\end{aligned} \tag{B.35}$$

Then by Markov,

$$\Pr\left[X \geq t^{2p-9/16} | B_1\right] \leq \frac{\mathbb{E}[X|B_1]}{t^{2p-9/16}} \leq \frac{8m^2}{16p-7} \frac{t^{2p-5/8}}{t^{2p-9/16}} = \frac{8m^2}{16p-7} \frac{t^{-5/8}}{t^{-9/16}}, \tag{B.36}$$

and $\frac{t^{-5/8}}{t^{-9/16}} = \frac{t^{9/16}}{t^{10/16}}$ which goes to zero. □

Let $F \subseteq G$ be the star forest consisting of edges between S_1 and $S_3 \backslash S_3'$.

Claim B.8 Let $\Delta_1 \geq \Delta_2 \geq \ldots \geq \Delta_k$ denote the degrees of the k highest degree vertices of G. Then $\lambda_i(F) = (1 - o(1))\Delta_i^{1/2}$.

Proof. Denote K_{1,d_i} to be a star of degree d_i. Let H be the star forest $H = K_{1,d_1} \cup \ldots \cup K_{1,d_k}$ with $d_1 \geq d_2 \geq \ldots \geq d_k$. Then for $i = 1, \ldots, k$, $\lambda_i(H) = d_i^{1/2}$. So, it will be sufficient to show that $\Delta_i(F) = (1 - o(1))\Delta_i(G)$. Claim B.4 shows that the k highest degree vertices G are added before time t_1. So these vertices are all in F. The only edges to those vertices that are not in F are those added before time t_2 and those incident to S_3'.

By Theorem 6.1, we can choose $f(t)$ such that, $\Delta_1(G_{t_2}^m) \leq t_2^p f(t) \leq t^{7p/16}$. Also by Theorem 6.1, $\Delta_i(G) \geq t^p / \log t$. Finally, Claim B.7 says that $|S_3'| \leq t^{7p/16}$ whp. And so, with high probability,

$$\begin{aligned}\Delta_i(F) &\geq \Delta_i(G) - t^{7p/16} - mt^{7p/16} \geq \frac{t^p}{\log t} - t^{7p/16}(1+m) = \frac{t^{p/2}}{\log t}\left[1 - t^{7p/16}(1+m)\frac{\log t}{t^p}\right] \\ &= \frac{t^p}{\log t}\left[1 - (1+m)\frac{\log t}{t^{p-7p/16}}\right] = \frac{t^p}{\log t}\left[1 - (1+m)\frac{\log t}{t^{9p/16}}\right] = (1 - o(1))\Delta_i(G) \qquad .\end{aligned} \tag{B.37}$$

□

Let $H = G \backslash F$. Denote $\boldsymbol{A}_G$, $\boldsymbol{A}_F$ and $\boldsymbol{A}_H$ to be the adjacency matrices for graphs G, F and H. In the following claim, we'll show that $\lambda_1(\boldsymbol{A}_H)$ is $o(\lambda_k(\boldsymbol{A}_F))$. Consider the fact that $\boldsymbol{A}_F$ and $\boldsymbol{A}_F + \boldsymbol{A}_H = \boldsymbol{A}_G$ are symmetric matrices, which implies $\lambda_k(\boldsymbol{A}_G) \leq \lambda_k(\boldsymbol{A}_F) + \lambda_1(\boldsymbol{A}_H)$ (see for instance theorem 8.1.5 in [37]). That implies that for any subspace L,

$$\max_{\mathrm{x} \in L, \mathrm{x} \neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_G \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} = \max_{\mathrm{x} \in L, \mathrm{x} \neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_F \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} \pm O\left(\max_{\mathrm{x} \neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_H \mathrm{x}}{\mathrm{x}^\top \mathrm{x}}\right).$$

We will use the Courant–Fischer Minimax Theorem, which states

$$\lambda_i(H) = \min_{S} \max_{x \in S, x \neq 0} \frac{\mathrm{x}^\top \boldsymbol{A}_H \mathrm{x}}{\mathrm{x}^\top \mathrm{x}},$$

where $S \subset \mathbb{R}^i$ (see, e.g. [37] Theorem 8.1.2). That will be enough to finish the proof with $\lambda_i(\boldsymbol{A}_G) = \lambda_i(\boldsymbol{A}_F)(1 \pm o(1))$.

Claim B.9 $\lambda_1(\boldsymbol{A}_H) \leq 6mt^{29/64}$ whp.

Proof. Let H_i denotes the subgraph of H induced by S_i, and let H_{ij} denotes the subgraph of H containing only edges with one vertex in S_i and the other in S_j. That is, write $\boldsymbol{A}_H$ in the following way:

$$\boldsymbol{A}_H = \begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_{12} & \boldsymbol{H}_{13} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_2 & \boldsymbol{H}_{23} \\ \boldsymbol{H}_{31} & \boldsymbol{H}_{32} & \boldsymbol{H}_3 \end{bmatrix}.$$

We will use this to bound the maximal eigenvalue of $\boldsymbol{A}_H$ as

$$\begin{aligned} \lambda_1(\boldsymbol{A}_H) &= \lambda_1\left(\begin{bmatrix} \boldsymbol{H}_1 & \boldsymbol{H}_{12} & \boldsymbol{H}_{13} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_2 & \boldsymbol{H}_{23} \\ \boldsymbol{H}_{31} & \boldsymbol{H}_{32} & \boldsymbol{H}_3 \end{bmatrix}\right) \\ &\leq \lambda_1(\boldsymbol{H}_1) + \lambda_1(\boldsymbol{H}_2) + \lambda_1(\boldsymbol{H}_3) + \lambda_1(\boldsymbol{H}_{12}) + \lambda_1(\boldsymbol{H}_{23}) + \lambda_1(\boldsymbol{H}_{13}). \end{aligned} \tag{B.38}$$

Note that the maximum eigenvalue of a graph is at most the maximum degree of a graph. By Claim B.6 with $f(t) = t_1$ and $\varepsilon = 1/64$,

$$\begin{aligned} \lambda_1(\boldsymbol{H}_1) &\leq \Delta_1(\boldsymbol{H}_1) = \max_{v \leq t_1}\{d^m_{t_1}(v)\} \leq t_1^{p+\varepsilon} \leq t^{65/512}, \\ \lambda_1(\boldsymbol{H}_2) &\leq \Delta_1(\boldsymbol{H}_2) \leq \max_{t_1 \leq v \leq t_2}\{d^m_{t_2}(v)\} \leq t_2^{p+\varepsilon}/t_1^p \leq t^{457/1024}, \\ \lambda_1(\boldsymbol{H}_3) &\leq \Delta_1(\boldsymbol{H}_3) \leq \max_{t_2 \leq v \leq t_3}\{d^m_{t_3}(v)\} \leq t_3^{p+\varepsilon}/t_2^p \leq t^{29/64}. \end{aligned} \tag{B.39}$$

To bound $\lambda_1(\boldsymbol{H}_{ij})$, start with $m = 1$. For $i < j$, this implies that each vertex in S_j has at most one edge in $\boldsymbol{H}_{ij}$, that is H_{ij} is a star forest. Then we have a bound on $\boldsymbol{H}_{ij}$ by Claim B.8. For $m > 1$, let G' be one of our generated graphs with t edges and $m = 1$. Think now of contracting vertices in G' (only the ones added using PA) into a single vertex. We can write $\boldsymbol{A}_G$ in terms of $\boldsymbol{A}'_G$: $\boldsymbol{A}_G = \boldsymbol{C}^\top \boldsymbol{A}'_G \boldsymbol{C}$, where $\boldsymbol{C}$ is a contraction matrix with t rows and the number of columns equal to the number of vertices in $\boldsymbol{A}_G$ (at most t/m). The ith column is equal to 1 at indices j in which (i, j) are identified. Similarly, we can write $\boldsymbol{H}_{ij}$ in terms of $\boldsymbol{H}'_{ij}$.

Note that if $\mathrm{y} = \boldsymbol{C}\mathrm{x}$, then $\mathrm{y}^\top \mathrm{y} = \mathrm{x}^\top \boldsymbol{C}^\top \boldsymbol{C} \mathrm{x}$, where $\boldsymbol{C}^\top \boldsymbol{C}$ is a diagonal matrix with $1's$ and $m's$ on the diagonal. So $\mathrm{x}^\top \mathrm{x} \leq \mathrm{y}^\top \mathrm{y} \leq m\mathrm{x}^\top \mathrm{x}$. We use this inequality to bound the largest eigenvalue as

$$\begin{aligned} \lambda_1(\boldsymbol{H}_{ij}) &= \max_{\mathrm{x} \neq 0} \frac{\mathrm{x}^\top \boldsymbol{H}_{ij} \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} = \max_{\mathrm{x} \neq 0} \frac{\mathrm{x}^\top \boldsymbol{C}^\top \boldsymbol{H}'_{ij} \boldsymbol{C} \mathrm{x}}{\mathrm{x}^\top \mathrm{x}} = \max_{\mathrm{x} \neq 0, \mathrm{y} = \boldsymbol{C}\mathrm{x}} \frac{\mathrm{y}^\top \boldsymbol{H}'_{ij} \mathrm{y}}{\mathrm{x}^\top \mathrm{x}} \\ &= \max_{\mathrm{x} \neq 0, \mathrm{y} = \boldsymbol{C}\mathrm{x}} \frac{m\mathrm{y}^\top \boldsymbol{H}'_{ij} \mathrm{y}}{m\mathrm{x}^\top \mathrm{x}} \leq \max_{\mathrm{x} \neq 0, \mathrm{y} = \boldsymbol{C}\mathrm{x}} \frac{m\mathrm{y}^\top \boldsymbol{H}'_{ij} \mathrm{y}}{\mathrm{y}^\top \mathrm{y}}. \end{aligned} \tag{B.40}$$

Now using Claim B.6 with $f(t) = t_1$ and $\varepsilon = 1/64$,

$$\Delta_1(\boldsymbol{H}'_{12}) = \max_{v \le t_2}\{d'_{t_2}(v)\} \le t_2^{p+\varepsilon} \le t^{37/64},$$
$$\Delta_1(\boldsymbol{H}'_{23}) = \max_{t_1 \le v \le t_3}\{d'_{t_3}(v)\} \le t_3^{p+\varepsilon}/t_1^p \le t^{29/64}. \tag{B.41}$$

Finally, all edges in $\boldsymbol{H}'_{13}$ are between S_1 and S'_3, so Claim B.7 shows $\Delta_1(\boldsymbol{H}'_{13}) \le t^{p-9/16} \le t^{7/16}$ whp. Putting together Equations (B.40) and (B.41), we get $\lambda_1(\boldsymbol{H}_{ij}) \le m\lambda_1(\boldsymbol{H}'_{ij}) \le m\Delta_1(\boldsymbol{H}'_{ij})^{1/2} \le mt^{29/64}$. And so we get the final bound,

$$\lambda_1(\boldsymbol{A}_H) \le \sum_{i=1}^{3} \lambda_1(\boldsymbol{H}_i) + \sum_{i<j} \lambda_1(\boldsymbol{H}_{ij}) \le 6mt^{29/64}.$$

This shows that $\lambda_1(\boldsymbol{A}_H)$ is $o(\lambda_k(\boldsymbol{A}_F))$, which implies $\lambda_i(\boldsymbol{A}_G) = \lambda_i(\boldsymbol{A}_F)(1 \pm o(1))$. □

With the completion of the proof of Claim B.9, that finishes the proof. □