Published paper

Jordan, J. (2011) *Randomised reproducing graphs*, Electronic Journal of Probability, 16, pp. 1549-1562
http://128.208.128.142/~ejpecp/index.php

Randomised reproducing graphs

Jonathan Jordan
University of Sheffield *

July 26, 2011

Abstract

We introduce a model for a growing random graph based on simultaneous reproduction of the vertices. The model can be thought of as a generalisation of the reproducing graphs of Southwell and Cannings and Bonato et al to allow for a random element, and there are three parameters, α, β and γ, which are the probabilities of edges appearing between different types of vertices. We show that as the probabilities associated with the model vary there are a number of phase transitions, in particular concerning the degree sequence. If $(1+\alpha)(1+\gamma) < 1$ then the degree distribution converges to a stationary distribution, which in most cases has an approximately power law tail with an index which depends on α and γ. If $(1+\alpha)(1+\gamma) > 1$ then the degree of a typical vertex grows to infinity, and the proportion of vertices having any fixed degree d tends to zero. We also give some results on the number of edges and on the spectral gap.

AMS 2000 Subject Classification: Primary 05C82; secondary 60G99, 60J10

Key words and phrases: reproducing graphs, random graphs, degree distribution, phase transition

Submitted to EJP 21 December 2010, final version accepted 18 July 2011.

*School of Mathematics and Statistics, University of Sheffield, Hounsfield Road, Sheffield, Yorkshire, S3 7RH, U.K., jonathan.jordan@shef.ac.uk

1 Introduction

In this paper we introduce a new model for a growing random graph based on simultaneous reproduction of the vertices in the graph, with edges being formed between the new vertices and each other and between the new vertices and the existing ones according to a random mechanism conditioned on the pattern of edges between the vertices in the previously existing graph. The model is a generalisation of the models introduced by Southwell and Cannings [10, 11, 12] and the Iterated Local Transitivity (ILT) model of [2], introducing stochasticity, which causes the regular structure found in the graphs of [10, 11, 12] to be lost, and which may make them more suitable for modelling in areas such as social networks; the authors of [2] particularly suggest their model as a model for online social networks, mentioning Facebook and Twitter among other examples.

We will show that our model, which depends on three parameters α, β and γ, which are to be thought of as probabilities, exhibits a number of phase transitions as the parameters vary; for example for some values of the parameters we will show that the degree distribution of the graph converges to a limiting probability distribution, while for other choices of the parameters the degree of a randomly chosen (in an appropriate sense) vertex in G_n can be shown to tend to infinity as $n \to \infty$. We will also show that for certain choices of the parameter values the model exhibits a power-law-like decay of the degree distribution, which is a property reported for many "real world" networks, and is also associated with other random graph models such as preferential attachment.

We start with a graph G_0, and form a new graph G_{n+1} by adding a "child" vertex for every vertex of G_n. As in [10, 11, 12] we denote the vertices by binary strings, writing $v0$ for the "child" of vertex $v \in V(G_n)$ and $v1$ for the continuation of vertex v as a vertex of G_{n+1}. The edges of G_{n+1} are then obtained according to the following mechanism. For each n define independent (of each other and of the random variables at other stages of the construction) Bernoulli random variables $a^{(n)}_{\{u,v\}} \sim Ber(\alpha)$ for each unordered pair $\{u, v\}$ of vertices of G_n, $b^{(n)}_u \sim Ber(\beta)$ for each vertex in G_n, $c^{(n)}_{(u,v)} \sim Ber(\gamma)$ for each ordered pair (u, v) of vertices of G_n, and connect vertices as follows:

(a) $u1$ is connected to $v1$ in G_{n+1} if and only if u and v are connected in G_n, that is existing edges are retained, and no further edges are formed between existing vertices.

(b) $u0$ is connected to $u1$ in G_{n+1} if and only if $b^{(n)}_u = 1$, so each child is connected to its parent with probability β.

(c) $u0$ is connected to $v1$ in G_{n+1} if and only if $c^{(n)}_{(u,v)} = 1$ and u and v are connected in G_n, so each child is connected to each of its parent's neighbours with probability γ.

(d) $u0$ is connected to $v0$ in G_{n+1} if and only of $a^{(n)}_{\{u,v\}} = 1$ and u and v are connected in G_n, so each child is connected to each of its parent's neighbours' children with probability α.

The models introduced in [10, 11, 12] have $\alpha, \beta, \gamma \in \{0, 1\}$, so are deterministic. Additionally the case where $\alpha = 0, \beta = 1, \gamma = 1$ is the ILT model, introduced in [2] as a model for online social networks. The ILT(p) model introduced in [2] as a stochastic generalisation of the ILT model adds extra random edges between the child vertices without regard to whether their parents were connected, and thus cannot be seen as a special case of our model. In addition as defined in [2] the ILT(p) model always has at least the edges found in the basic ILT model, so is not suited to producing relatively sparse graphs.

The model differs from duplication graphs, for example those considered in [3, 5], in that in those models only one vertex, chosen at random, duplicates at any one time step, whereas in the models considered here and in [2, 10, 11, 12] all vertices simultaneously duplicate.

Our main results concern the degree distribution. We will deal with the cases where $\beta = 0$ and $\beta > 0$ separately, as the behaviour of the model when $\beta = 0$ is potentially quite different, with large numbers of isolated vertices.

Theorem 1. *Let $\beta = 0$. Then, if $(1+\gamma)(\alpha+\gamma) \leq 1$, the probability that a randomly chosen vertex in the graph G_n is isolated tends to 1 as $n \to \infty$, and the proportion of vertices in the graph with degree zero tends to 1, almost surely. If $(1+\gamma)(\alpha+\gamma) > 1$, then the probability that a randomly chosen vertex in the graph G_n is isolated converges to some value strictly less than 1.*

Theorem 2. *Assume $\beta > 0$, and let $p^{(n)}_d$ be the proportion of vertices in G_n with degree d. Then*

(a) *If $(1+\gamma)(\alpha+\gamma) < 1$ there exists a random variable X such that $p^{(n)}_d \to P(X = d)$ as $n \to \infty$, almost surely.*

(b) *Under the conditions of (a), the random variable X has a finite pth moment if $(1+\gamma)^p + (\alpha+\gamma)^p < 2$, and does not have a finite pth moment if $(1+\gamma)^p + (\alpha+\gamma)^p > 2$.*

(c) *If $(1+\gamma)(\alpha+\gamma) > 1$ then $p^{(n)}_d \to 0$ as $n \to \infty$, almost surely.*

Note that Theorem 2(b) implies that if $(1+\gamma)^p + (\alpha+\gamma)^p = 2$ the tail of the degree distribution is asymptotically close to a power law degree distribution with index given by $-(p+1)$, in the sense that qth moments exist for $q < p$ but not for $q > p$.

In [2], it is shown that the ILT model exhibits a "densification power law", which is defined to mean that, if E_n is the number of edges of G_n and V_n the number of vertices, then E_n is proportional to $(V_n)^a$ for some $a \in (1,2)$. The following result shows that our model exhibits a phase transition in this respect, with the transition occurring where $2\gamma + \alpha = 1$. Note that in our model, as in the ILT model, $V_n = 2^n V_0$ for all n.

Theorem 3. **(a)** *If $2\gamma + \alpha > 1$ then $W_n = \frac{E_n}{(1+2\gamma+\alpha)^n}$ converges to a positive limit, so that the model has a densification power law as defined by [2] with exponent $\frac{\log(1+2\gamma+\alpha)}{\log 2}$.*

(b) *If $2\gamma + \alpha < 1$ then*

$$\frac{E_n}{2^n} \to \frac{V_0 \beta}{1 - 2\gamma - \alpha},$$

almost surely, as $n \to \infty$, so that the number of edges grows at the same rate as the number of vertices

(c) *If $2\gamma + \alpha = 1$ then*

$$\frac{E_n}{2^n n} \to \frac{V_0 \beta}{2},$$

almost surely, as $n \to \infty$.

Note that the combination of Theorems 3 and 2 implies that when $2\gamma+\alpha > 1$ but $(1+\gamma)(\alpha+\gamma) < 1$ the process exhibits both a densification power law in the sense of [2] and an approximately power law limit for the degree distribution.

A further result in [2] on the ILT model concerns the spectral gap. They show that the normalised graph Laplacian $\mathcal{L}$, as defined by Chung [4], of the ILT model has a large spectral radius, defined as $\max\{|\lambda_1 - 1|, |\lambda_{n-1} - 1|\}$, where λ_1 is the second smallest eigenvalue (the smallest being $\lambda_0 = 0$ for any graph) and λ_{n-1} is the largest eigenvalue and thus that the graph has relatively poor expansion properties. The following results show that the same is also true for our model. We concentrate on the case $\beta = 1$, where the graphs are connected; otherwise λ_1 will be zero. The proofs use the Cheeger constant and its relationship to λ_1, as defined in Chapter 2 of Chung [4].

Theorem 4. *Let $\beta = 1$ and assume that G_0 is connected, so that G_n will also be connected for all n. Let $\lambda_1(G_n)$ be the smallest non-negative eigenvalue of the Laplacian of G_n. Then*

(a) *If $2\gamma + \alpha \leq 1$ then $\lambda_1(G_n) \to 0$ as $n \to \infty$.*

(b) *If $2\gamma + \alpha > 1$ then there exists a (random) Λ, with $\Lambda < 1$ almost surely, such that $\limsup_{n\to\infty} \lambda_1 = \Lambda$.*

Some of the results of Theorems 2 and 3 are illustrated by the phase diagram of the γ–α plane in Figure 1.

2 Pictures

This section shows a few examples of graphs of this type, which were generated using a script written with the igraph package, [6], in R. Figure 2 shows examples with $n = 7$, $\alpha = 0$ and $\beta = 1$, with γ varying. These show the graph changing as γ increases from a sparse tree-like graph where most vertices have low degree to a much denser graph where many vertices have high degree.

Figure 3 shows examples with $\alpha = 0$ and $\gamma = 0.366$ (this is approximately $\frac{\sqrt{3}-1}{2}$), two values of β and again $n = 7$. Those with higher β have fewer isolated vertices and have more of the vertices in the largest component.

Finally, Figure 4 shows examples with $\beta = 1$, $\gamma = 0.05$ and α varying. These show graphs with a less-tree like structure with fewer very low degree vertices as α increases.

3 Proofs of Theorems

We start with a lemma on the conditional expectation and variance of the number of edges in G_n, E_n. This lemma will be useful for obtaining the mean of the stationary distribution of a Markov chain which we will use to prove Theorems 1 and 2. We define $\mathcal{F}_n$ to be the σ-algebra generated by the graphs G_m for $m \leq n$, and we use the notation $Bin(n, p)$ for the binomial distribution with n trials and success probability p.

Lemma 5. *The conditional expectation and variance of E_{n+1} satisfy*

$$\begin{aligned} \mathbb{E}(E_{n+1}|\mathcal{F}_n) &= (1 + 2\gamma + \alpha)E_n + 2^n\beta V_0 \\ \mathrm{Var}(E_{n+1}|\mathcal{F}_n) &= E_n(2\gamma(1-\gamma) + \alpha(1-\alpha)) + 2^n V_0\beta(1-\beta) \end{aligned}$$

Proof. This follows from the fact that the E_{n+1} can be written $E_n + E_{n+1,1} + E_{n+1,2} + E_{n+1,3}$ where $E_{n+1,1}$, $E_{n+1,2}$ and $E_{n+1,3}$ are independent, $E_{n+1,1}$ represents the edges between parents and children of their neighbours and, conditional on $\mathcal{F}_n$ has a $Bin(2E_n, \gamma)$ distribution, $E_{n+1,2}$

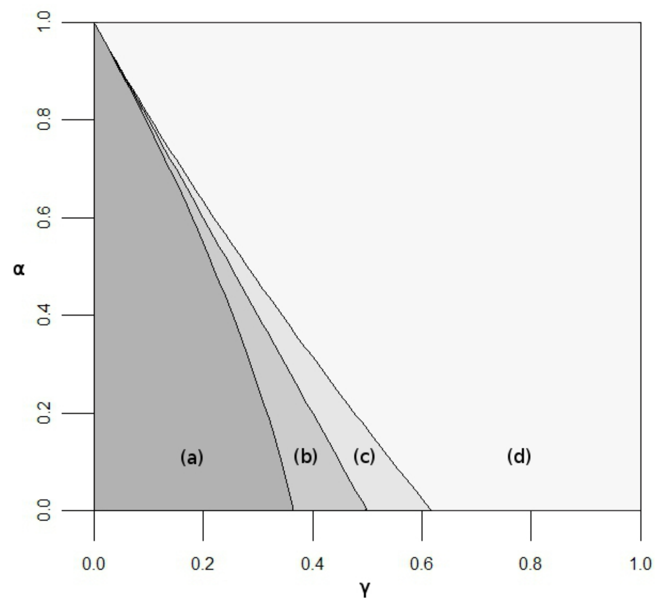

Figure 1: Phase diagram showing how the limiting behaviour of the graphs in the $\beta \neq 0$, as described by Theorems 2 and 3, case can vary with α and γ. In region (a) the degree sequence converges to a distribution with both first and second moments finite; in region (b) the limiting distribution has first moment finite but the second moment not; in region (c) the limiting distribution has infinite mean; in region (d) there is no limiting degree distribution. In regions (a) and (b) the graphs are sparse; in regions (c) and (d) they have a densification power law as described by Theorem 3. The borders between the regions intersect the γ axis at $\frac{\sqrt{3}-1}{2}$, $\frac{1}{2}$ and $\frac{\sqrt{5}-1}{2}$.

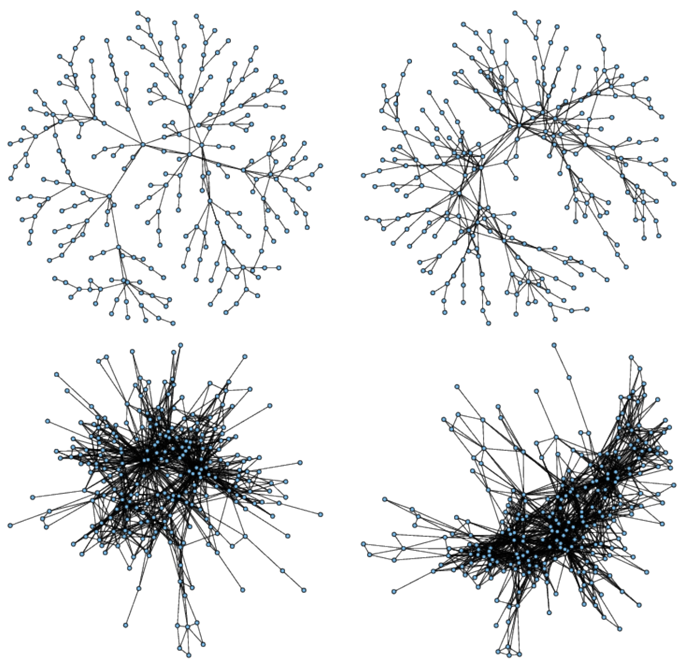

Figure 2: Example simulations with $\alpha = 0$ and $\beta = 1$. From left to right, top row first, $\gamma = 0.05, 0.2, 0.49, 0.6$.

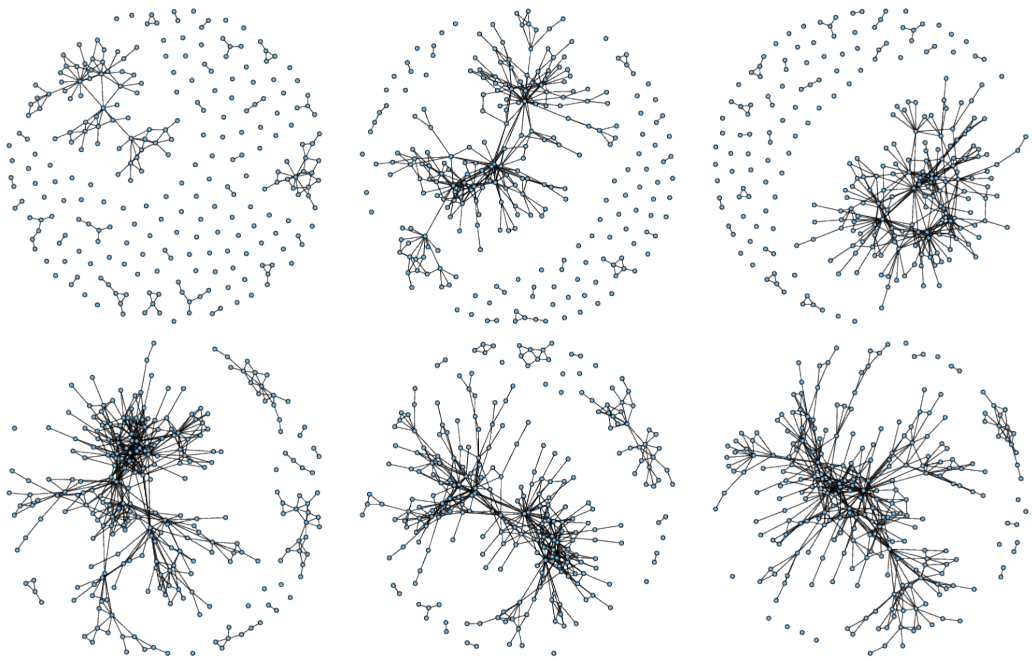

Figure 3: Example simulations with $\alpha = 0$ and $\gamma = 0.366$. The top row have $\beta = 0.4$ and the bottom row $\beta = 0.8$.

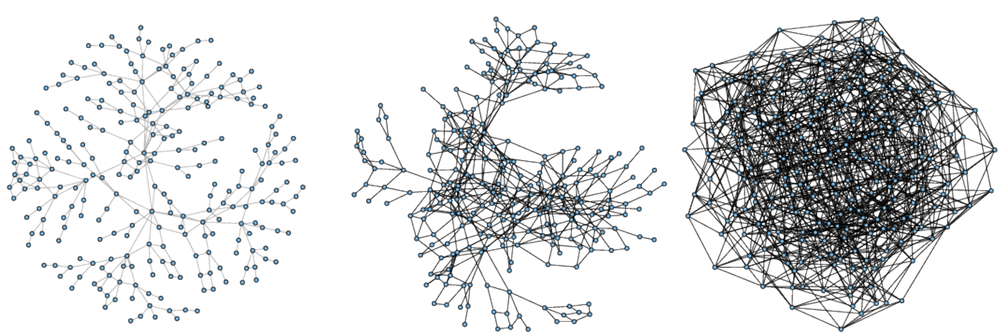

Figure 4: Example simulations with $\beta = 1$ and $\gamma = 0.05$. From left to right, $\alpha = 0.2, 0.5, 0.9$.

represents the edges between children of neighbouring vertices and, conditional on $\mathcal{F}_n$ has a $Bin(E_n, \alpha)$ distribution, and $E_{n+1,3}$ represents the edges between parents and their children and, conditional on $\mathcal{F}_n$ has a $Bin(V_n, \beta)$ distribution. As $V_n = 2^n V_0$ the result follows. □

3.1 Proof of Theorem 3

We start with (a), the case where $2\gamma + \alpha > 1$. The following approach is based on that in [1] for multitype branching processes, the idea being that the vertices and edges in the graph G_n can be thought of as the two types in a population undergoing branching. However the resulting multitype branching process is not irreducible, so the results in [1] cannot be used directly.

Given an edge in G_m between vertices u and v, there will be an edge in G_{m+1} between $u1$ and v_1, and in addition there will be edges between $u1$ and $v0$ and $v1$ and $u0$ each with probability γ and an edge between $u0$ and $v0$ with probability α. We can consider these edges as offspring of the edge between u and v, and thus consider the set of edges in G_n (for $n > m$) which are descendants of the edge between u and v in G_m as a generation in a Galton-Watson branching process with offspring mean $1 + 2\gamma + \alpha$, and where the extinction probability is zero and the number of offspring bounded. Treating the descendants of a given edge in G_m as a subset of the edge set of G_n, this shows that $\liminf \frac{E_n}{(1+2\gamma+\alpha)^n}$ is a positive random variable.

Now define

$$W_n = \frac{V_n + \frac{2\gamma+\alpha-1}{\beta} E_n}{(1 + 2\gamma + \alpha)^n}.$$

Then $\mathbb{E}(W_{n+1}|\mathcal{F}_n) = W_n$, so $(W_n)_{n\in\mathbb{N}}$ is a non-negative martingale, and thus almost surely has a non-negative limit W. The above conclusion on $\liminf \frac{E_n}{(1+2\gamma+\alpha)^n}$ shows that $P(W = 0) = 0$, giving the result.

For (b), the case where $2\gamma + \alpha < 1$, Lemma 5 shows that $\mathbb{E}(E_n) = \frac{V_0\beta}{1-2\gamma-\alpha} 2^n + o(2^n)$ and

$$\operatorname{Var}(E_n) = 2^{n-1} \frac{V_0\beta}{1 - 2\gamma - \alpha} (2\gamma(1-\gamma) + o(2^n) + \alpha(1-\alpha)) + 2^n V_0 \beta (1-\beta) + (1+2\gamma+\alpha)^2 \operatorname{Var}(E_{n-1}),$$

which shows that

$$\operatorname{Var}\left(\frac{E_n}{2^n}\right) = O\left(\left(\max\left(\frac{1}{2}, \frac{(1+2\gamma+\alpha)^2}{4}\right)\right)^n\right),$$

which implies the result via Chebyshev's inequality and the Borel-Cantelli Lemmas.

For (c), the case where $2\gamma+\alpha = 1$, an iterative use of Lemma 5 shows that $\mathbb{E}(E_n) = 2^n\left(E_0 + \frac{\beta V_0}{2}n\right)$ and $\mathrm{Var}(E_{n+1}) = 2^{n-1}n\beta V_0(2\gamma(1-\gamma)+\alpha(1-\alpha)+4)+O(2^n)$. Then

$$\mathrm{Var}\left(\frac{E_n}{2^n n}\right) = O\left(\frac{1}{n^2 2^n}\right),$$

allowing the Chebyshev/Borel-Cantelli argument again.

3.2 Proof of Theorem 4

We consider some small m, and find the Cheeger constant of G_m. By the definition in [4], this will be $e(S_m, \bar{S_m})/\operatorname{vol}(S_m)$ for some $S_m \subseteq V(G_m)$, where for two subsets of the vertex set S and S' $e(S, S')$ is the number of edges between a vertex in S and one in S', and $\operatorname{vol}(S)$ is the sum of the degrees of vertices in S. (Note that $\operatorname{vol}(S) = 2e(S,S) + e(S, \bar{S}.)$ Now consider the descendants of S_m in G_n as a subset $S_n \subseteq V(G_n)$. Then the same arguments as in the proof of Theorem 3, applied to the subgraphs descending from S_m and $\bar{S_m}$, show that if $2\gamma+\alpha < 1$ then the $e(S_n, S_n)$ and $e(\bar{S_n}, \bar{S_n})$ both grow at rate 2^n, in the sense that $\frac{e(S_n,S_n)}{2^n}$ and $\frac{e(\bar{S_n},\bar{S_n})}{2^n}$ converge almost surely to positive constants as $n \to \infty$, and similarly if $2\gamma+\alpha = 1$ $e(S_n, S_n)$ and $e(\bar{S_n}, \bar{S_n})$ both grow at rate $2^n n$, and if $2\gamma+\alpha > 1$ $e(S_n, S_n)$ and $e(\bar{S_n}, \bar{S_n})$ both grow at rate $(1+2\gamma+\alpha)^n$.

Next, again as in the proof of Theorem 3, $(e(S_n, \bar{S_n}))_{n\in\mathbb{N}}$ forms a Galton-Watson branching process with mean of the offspring distribution $1+2\gamma+\alpha$, and extinction probability zero, so $e(S_n, \bar{S_n})$ will grow at rate $(1+2\gamma+\alpha)^n$. So for $2\gamma+\alpha > 1$ $\frac{e(S_n,\bar{S_n})}{\min(\operatorname{vol}(S_n),\operatorname{vol}(\bar{S_n}))}$, which by the definition in [4] is greater than the Cheeger constant of G_n, converges to a constant (less than 1, as $\operatorname{vol}(S_n) = 2e(S_n, S_n) + e(S_n, \bar{S_n})$) and this constant bounds the lim sup of the Cheeger constant of G_n above.

In the case where $2\gamma+\alpha < 1$

$$\frac{e(S_n, \bar{S_n})}{\max(\operatorname{vol}(S_n), \operatorname{vol}(\bar{S_n}))} = O\left(\left(\frac{1+2\gamma+\alpha}{2}\right)^n\right) \to 0$$

as $n \to \infty$, and hence so is the Cheeger constant. Similarly if $2\gamma+\alpha = 1$

$$\frac{e(S_n, \bar{S_n})}{\max(\operatorname{vol}(S_n), \operatorname{vol}(\bar{S_n}))} = O\left(\frac{1}{n}\right),$$

and again so is the Cheeger constant.

Hence by the Cheeger inequality (Lemma 2.1 and Theorem 2.2 of [4]), $\limsup_{n\to\infty} \lambda_1(G_n) < 1$ almost surely in the case $2\gamma + \alpha > 1$, and $\lambda_1(G_n)$ tends to zero in the case $2\gamma + \alpha \leq 1$.

3.3 Proofs of Theorems 1 and 2

The proofs of Theorems 1 and 2 will rely on defining a certain Markov chain whose value X_n represents the degree of a random vertex in the graph G_n. We construct this by letting v_0 be a vertex of G_0 chosen uniformly at random, and then, using the binary string notation for the vertices described above, for $n \geq 1$ let $v_n = v_{n-1}1$ with probability 1/2 and letting $v_n = v_{n-1}0$ with probability 1/2. We then let X_n be the degree of v_n in G_n.

Then

$$X_{n+1} = \xi_{n+1} X_n + (1 - \xi_{n+1}) W_{n+1} + Y_{n+1} + Z_{n+1}, \tag{1}$$

where, conditional on G_n, $Y_{n+1} \sim Bin(X_n, \gamma)$, $W_{n+1} \sim Bin(X_n, \alpha)$, $Z_{n+1} \sim Bin(1, \beta)$ and $\xi_{n+1} \sim Bin(1, \frac{1}{2})$, with all these variables being conditionally independent given G_n.

Here, $\xi_{n+1} = 1$ if our vertex in G_{n+1} is a parent and 0 if it is a child, W_{n+1} represents child-child connections (so does not appear if $\xi_{n+1} = 1$), Y_{n+1} represents connections between a child and its parents' neighbours, and Z_{n+1} represents the connection between the child and its parent.

As defined above, $(X_n)_{n\in\mathbb{N}}$ is a discrete time Markov chain on the natural numbers (including zero if $\beta < 1$). It is irreducible and aperiodic if $\beta > 0$, $\alpha < 1$ and $\gamma < 1$. (If $\beta = 0$ then zero is an absorbing state, and if either α or γ is 1 then X_n is increasing in n and so the chain is certainly not irreducible, but otherwise $P(X_{n+1} = 1|\mathcal{F}_n)$ is always positive.)

Proposition 6. *If $2\gamma + \alpha < 1$ the distribution of X_n converges in the Wasserstein-1 metric to a unique fixed point with finite mean $\frac{2\beta}{1-2\gamma-\alpha}$.*

Proof. Note that if we have another random variable $\hat{X}_n$ with a different distribution on $\mathbb{N}_0$, we can apply (1) to it by defining, conditional on $\hat{X}_n$, $\hat{Y}_{n+1} \sim Bin(\hat{X}_n, \gamma)$ and $\hat{W}_{n+1} \sim Bin(\hat{X}_n, \alpha)$ using the same set of Bernoulli trials as for Y_{n+1} and W_{n+1} respectively, and letting

$$\hat{X}_{n+1} = \xi_{n+1} \hat{X}_n + (1 - \xi_{n+1}) \hat{W}_{n+1} + \hat{Y}_{n+1} + Z_{n+1}.$$

Then, conditional on $\mathcal{F}_n$, if $\hat{X}_n > X_n$ we have

$$\hat{X}_{n+1} - X_{n+1} = \xi_{n+1}(\hat{X}_n - X_n) + (1 - \xi_{n+1})(\hat{W}_{n+1} - W_{n+1}) + \hat{Y}_{n+1} - Y_{n+1},$$

where $\hat{W}_{n+1} - W_{n+1} \sim Bin(\hat{X}_n - X_n, \alpha)$ and $\hat{Y}_{n+1} - Y_{n+1} \sim Bin(\hat{X}_n - X_n, \gamma)$, and similarly, if $\hat{X}_n < X_n$ we have

$$X_{n+1} - \hat{X}_{n+1} = \xi_{n+1}(X_n - \hat{X}_n) + (1 - \xi_{n+1})(W_{n+1} - \hat{W}_{n+1}) + Y_{n+1} - \hat{Y}_{n+1},$$

where $W_{n+1} - \hat{W}_{n+1} \sim Bin(X_n - \hat{X}_n, \alpha)$ and $Y_{n+1} - \hat{Y}_{n+1} \sim Bin(X_n - \hat{X}_n, \gamma)$. Combining these, we can see that

$$\mathbb{E}(|X_{n+1} - \hat{X}_{n+1}||\mathcal{F}_n) = \frac{1}{2}(2\gamma + \alpha + 1)|X_n - \hat{X}_n|,$$

so that we have a contraction in the Wasserstein metric if $2\gamma + \alpha < 1$. Hence in this case there is convergence in the Wasserstein-1 metric of the degree distributions to a unique fixed point with finite mean.

We can calculate the mean of this distribution by using Lemma 5: letting $m = 2$ we get

$$\mathbb{E}(k_2^{(n+1)}) = (2\gamma + \alpha + 1)k_2^{(n)} + 2^n \beta v_0,$$

where v_0 is the number of vertices in the initial graph, and solving this we find that the expected number of edges in G_n is

$$\frac{\beta v_0(2^n - (1 + 2\gamma + \alpha)^n)}{1 - 2\gamma - \alpha},$$

so (as the number of vertices in G_n is $2^n v_0$) the expected average degree is

$$\frac{\beta(2^n - (1 + 2\gamma + \alpha)^n)}{2^{n-1}(1 - 2\gamma - \alpha)},$$

which converges to $\frac{2\beta}{1-2\gamma-\alpha}$ as $n \to \infty$. $\square$

To go further than this we use Foster-Lyapunov techniques, as described in Meyn and Tweedie [8] in the more general case of an uncountable state space. The following lemma on the conditional moments of X_{n+1} (including negative and fractional moments) will be useful.

Lemma 7. *Let $p \in \mathbb{R}$. Then as $x \to \infty$,*

$$\mathbb{E}\left(\left(\frac{1 + x + Y_{n+1} + Z_{n+1}}{1 + x}\right)^p |X_n = x\right) \to (1 + \gamma)^p,$$

and

$$\mathbb{E}\left(\left(\frac{1 + W_{n+1} + Y_{n+1} + Z_{n+1}}{1 + x}\right)^p |X_n = x\right) \to (\alpha + \gamma)^p.$$

Proof. In the case where $p < 0$ this is a special case of Theorem 2.1 of García and Palacios in [7]. When $p > 0$, the result follows from convergence in distribution of the conditional distributions of $\frac{1+x+Y_{n+1}+Z_{n+1}}{1+x}$ and $\frac{1+W_{n+1}+Y_{n+1}+Z_{n+1}}{1+x}$ given $X_n = x$ to $1+\gamma$ and $\alpha+\gamma$ respectively as $x \to \infty$, together with the fact that they are positive and bounded above by 2. □

Proposition 8. *If $(1 + \gamma)(\alpha + \gamma) < 1$, the Markov chain is positive recurrent, and thus the distribution of X_n converges to a stationary distribution.*

Proof. This uses Theorem 11.0.1 of [8].

We choose $p \in (0, 1)$ such that $(1+\gamma)^p + (\alpha+\gamma)^p < 2$. Because $\frac{\mathrm{d}}{\mathrm{d}p}((1+\gamma)^p+(\alpha+\gamma)^p)$ is negative at $p = 0$ if $\log(1 + \gamma) + \log(\alpha + \gamma) < 0$, it will be possible to find such a p if $(1 + \gamma)(\alpha + \gamma) < 1$.

We now let $V(x) = x^p$. In [8], the drift $\Delta V(x)$ is defined as

$$\Delta V(x) = \mathbb{E}(V(X_{n+1}) - V(X_n)|X_n = x),$$

and by Theorem 11.0.1 of [8] the chain will be positive recurrent if (for some V) $\Delta V(x) \leq -1$ for x large enough. Now

$$\begin{aligned}\mathbb{E}(X_{n+1}^p|X_n = x) &= \frac{x^p}{2}\left(\mathbb{E}\left(\left(1+\frac{Y_{n+1}}{x}+\frac{Z_{n+1}}{x}\right)^p|G_n\right)+\mathbb{E}\left(\left(\frac{W_{n+1}}{x}+\frac{Y_{n+1}}{x}+\frac{Z_{n+1}}{x}\right)^p|G_n\right)\right)\\ &\leq \frac{x^p}{2}\left((1+\gamma)^p+(\alpha+\gamma)^p\right)+o(x^p)\\ &\quad \text{(by Lemma 7)},\end{aligned}$$

so

$$\Delta V(x) \leq x^p\left(\frac{(1+\gamma)^p+(\alpha+\gamma)^p}{2}-1\right)+o(x^p),$$

which will be less than -1 for x large enough, giving the result.

□

We now investigate the tail behaviour of the stationary distribution, in the case where Proposition 8 shows one exists.

Proposition 9. *Let $p > 0$. If $(1 + \gamma)^p + (\alpha + \gamma)^p < 2$, then a random variable X with the stationary distribution of the chain has finite pth moment $\mathbb{E}(X^p)$, and we have convergence of pth moments, $\mathbb{E}(X_n^p) \to \mathbb{E}(X^p)$ as $n \to \infty$.*

Proof. Again this uses a Foster-Lyapunov type technique, in this case Theorem 14.0.1 of [8] which states that if, for a given function $f \geq 1$, we can find V such that $\Delta V(x) < -f(x)$ for x large enough then f has a finite integral with respect to the stationary distribution and that $\mathbb{E}(f(X_n))$ converges to this integral. We will set $f(x) = x^p + 1$.

Let $V(x) = kx^p$, where k is chosen so that

$$k\left(\frac{(1+\gamma)^p + (\alpha+\gamma)^p}{2} - 1\right) < -1.$$

Then, by Lemma 7,

$$\Delta V(x) \leq kx^p\left(\frac{(1+\gamma)^p + (\alpha+\gamma)^p}{2} - 1\right) + o(x^p),$$

and so $\Delta V(x) \leq -f(x)$ for x large enough, giving the result. □

Proposition 10. *Let $p > 0$. If $(1+\gamma)^p + (\alpha+\gamma)^p > 2$, then a random variable X with the stationary distribution of the chain does not have finite pth moment $\mathbb{E}(X^p)$.*

Proof. As $Z_{n+1} \geq 0$, we have

$$\mathbb{E}(X_{n+1}^p | X_n = x) \geq \frac{x^p}{2}\left(\mathbb{E}\left(\left(1 + \frac{Y_{n+1}}{x}\right)^p | X_n = x\right) + \mathbb{E}\left(\left(\frac{W_{n+1}}{x} + \frac{Y_{n+1}}{x}\right)^p | X_n = x\right)\right),$$

so by Lemma 7

$$\mathbb{E}(X_{n+1}^p | X_n = x) \geq \frac{x^p}{2}\left((1+\gamma)^p + (\alpha+\gamma)^p\right) + o(x^p).$$

Hence the pth moment of X_n tends to infinity as $n \to \infty$, so by Theorem 14.0.1 of [8], again applied to $f(x) = x^p + 1$, the stationary distribution cannot have a finite pth moment. □

Proposition 11. *If $\beta > 0$ and $(1+\gamma)(\alpha+\gamma) > 1$ the Markov chain is transient.*

Proof. By (1) and Lemma 7,

$$\frac{E\left((1+X_{n+1})^p | X_n = x\right)}{(1+x)^p} \to \frac{(1+\gamma)^p + (\alpha+\gamma)^p}{2}, \tag{2}$$

so we apply Theorem 8.0.2 (i) of [8] with $V(x) = 1 - (1+x)^p$ for some $p < 0$ such that $(1+\gamma)^p + (\alpha+\gamma)^p < 2$. With this choice of V, (2) shows that $\Delta V(x) > 0$ for x large enough, and as V is bounded and positive on the natural numbers Theorem 8.0.2 (i) of [8] gives the result.

□

The case where $\beta = 0$ is something of a special case as the chain is not irreducible. However we can show that when $(1+\gamma)(\alpha+\gamma) \leq 1$ the probability that a randomly chosen vertex is isolated tends to 1, while there is positive probability that a randomly chosen vertex is not isolated when $(1+\gamma)(\alpha+\gamma) > 1$.

Proposition 12. *If $\beta = 0$, then*

1. *if $(1+\gamma)(\alpha+\gamma) \leq 1$ then almost surely $X_n = 0$ for n sufficiently large, and the proportion of isolated vertices in G_n tends to 1 almost surely as $n \to \infty$;*

2. *if $(1+\gamma)(\alpha+\gamma) > 1$ then there is $q > 0$ such that the probability that $X_n \to \infty$ as $n \to \infty$ is q and the probability that $X_n \to 0$ as $n \to \infty$ is $1-q$.*

Proof. We note that (X_n) follows a Smith-Wilkinson branching process in random environment, [9]. The environmental variables which determine the random environment are the random variables ξ_n, with the offspring distribution of the branching process at time n having mean $1+\gamma$ if $\xi_{n+1} = 1$ and $\alpha+\gamma$ if $\xi_{n+1} = 0$. Hence, by Theorem 3.1 of [9], the branching process dies out with probability 1 if $\frac{1}{2}\log(1+\gamma) + \frac{1}{2}\log(\alpha+\gamma) \leq 0$, i.e. if $(1+\gamma)(\alpha+\gamma) \leq 1$, and the branching process dies out with probability strictly less than 1 otherwise, hence there is positive probability that $X_n \to \infty$ as $n \to \infty$. To see that the proportion of isolated vertices tends to 1 almost surely when $(1+\gamma)(\alpha+\gamma) \leq 1$, note that the proportion of isolated vertices is increasing (as if a vertex v is isolated in G_n both $v0$ and $v1$ are isolated in G_{n+1}) and therefore must converge to some value, which cannot be less than 1 as the degree of a random vertex converges to zero almost surely. □

Proposition 13. **(a)** *If $\beta > 0$ and $(1+\gamma)(\alpha+\gamma) < 1$, the degree distribution of the graph converges to the stationary distribution of the Markov chain in the sense that if we let $p_d^{(n)}$ be the proportion of vertices in G_n with degree d, and let X be a random variable with the stationary distribution of the Markov chain, then $p_d^{(n)} \to P(X = d)$ as $n \to \infty$, almost surely, for all $d \in \mathbb{N}_0$.*

(b) *If $\beta > 0$ and $(1+\gamma)(\alpha+\gamma) > 1$ then $p_d^{(n)} \to 0$ as $n \to \infty$, almost surely, for all $d \in \mathbb{N}_0$.*

Proof. The graph at stage r contains $2^r v_0$ vertices. We then consider the edges of G_{r+s} in two sets: those which are between descendants of the same vertex in G_r, and those which are between descendants of different vertices in G_r. For the former, the appearance of edges between descendants of one given vertex is independent of what happens to the descendants of the other vertices, so we can model these edges of G_{r+s} as consisting of $2^r v_0$ independent copies of $\tilde{G}_s$, where $(\tilde{G}_n)_{n\in\mathbb{N}}$ represents the process as evolved from a single vertex with no edges, $\tilde{G}_0$. Then, by Chebyshev's inequality and a Borel-Cantelli argument, as $r \to \infty$ proportions of vertices in G_{r+s} with degree d excluding connections to descendants of different vertices in G_r converge, almost surely, to $P(X_s = d)$. In the $(1+\gamma)(\alpha+\gamma) > 1$ case we know that $P(X_s = d) \to 0$ as $s \to \infty$ for all d, and the actual degree of a vertex is bounded below by the degree excluding some connections, so this is enough to prove (b).

To complete the proof of (a), we need to consider edges between vertices which are descendants of different vertices in G_r. We couple the process, starting from G_r, with a process with $\beta = 0$ by removing all edges between a vertex and its offspring, and all edges descended from such edges. The edges thus removed from G_{r+s} will all be between vertices descended from the same vertex in G_r, so all edges between vertices descended from different vertices in G_r are present in the $\beta = 0$ version. But by Proposition 12 the proportion of vertices in G_{r+s} which have non-zero degree in the $\beta = 0$ version tends to zero as $s \to \infty$, and so this also applies to the proportion of vertices in G_{r+s} which have edges connecting them to vertices with a different ancestor in G_r.

Hence as both r and $s \to \infty$ the proportion which have degree d converges to $P(X_s = d)$. □

Finally we can put the Propositions above together to deduce Theorems 1 and 2.

Proof of Theorem 1. Theorem 1 follows from Proposition 12; in the supercritical case where $(1+\gamma)(\alpha+\gamma) > 1$ the probability that a randomly chosen vertex in the graph is isolated tends to $1 - q < 1$.

Proof of Theorem 2. Theorem 2 follows immediately from Propositions 8, 9, 10 and 13.

4 Acknowledgements

The author would like to acknowledge the support of the EPSRC funded Amorphous Computing, Random Graphs and Complex Biological Systems group (grant reference EP/D003105/1), and would also like to thank John Biggins and Andrew Wade for some useful discussions while the paper was being written and two anonymous referees for suggesting some improvements to the paper.

References

[1] K. E. Athreya and P. E. Ney. *Branching Processes.* Dover Publications, Mineola, New York, 1972.

[2] A. Bonato, N. Hadi, P. Horn, P. Prałat, and C. Wang. Models of on-line social networks. *Internet Mathematics*, 6:285–313, 2011.

[3] F. Chung, L. Lu, T. Dewey, and D. Gales. Duplication models for biological networks. *Journal of Computational Biology*, 10:677–687, 2003.

[4] F. R. K. Chung. *Spectral Graph Theory.* Number 92 in CBMS Regional Conference Series. AMS, Providence, Rhode Island, 1997.

[5] N. Cohen, J. Jordan, and M. Voliotis. Preferential duplication graphs. *Journal of Applied Probability*, 2010.

[6] G. Csárdi and T. Nepusz. The igraph software package for complex network research. *InterJournal Complex Systems.*

[7] Nancy Lopes García and José Luis Palacios. On inverse moments of nonnegative random variables. *Statist. Probab. Lett.*, 53(3):235–239, 2001.

[8] S. Meyn and R. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993.

[9] W. Smith and W. Wilkinson. On branching processes in random environments. *Annals of Mathematical Statistics*, 40:814–827, 1969.

[10] R. Southwell and C. Cannings. Games on graphs that grow deterministically. In *Proc. International Conference on Game Theory for Networks GameNets '09*, pages 347–356, 13–15 May 2009.

[11] R. Southwell and C. Cannings. Some models of reproducing graphs. 1 Pure reproduction. *Applied Mathematics*, 1:137–145, 2010.

[12] R. Southwell and C. Cannings. Some models of reproducing graphs. 2 Age capped vertices. *Applied Mathematics*, 1:251–259, 2010.