A SCALING LIMIT FOR THE DEGREE DISTRIBUTION IN SUBLINEAR PREFERENTIAL ATTACHMENT SCHEMES

JIHYEOK CHOI, SUNDER SETHURAMAN,
AND SHANKAR C. VENKATARAMANI

Abstract. We consider a general class of preferential attachment schemes evolving by a reinforcement rule with respect to certain sublinear weights. In these schemes, which grow a random network, the sequence of degree distributions is an object of interest which sheds light on the evolving structures.

In this article, we use a fluid limit approach to prove a functional law of large numbers for the degree structure in this class, starting from a variety of initial conditions. The method appears robust and applies in particular to 'non-tree' evolutions where cycles may develop in the network.

A main part of the argument is to show that there is a unique nonnegative solution to an infinite system of coupled ODEs, corresponding to a rate formulation of the law of large numbers limit, through C_0-semigroup/dynamical systems methods. These results also resolve a question in Chung, Handjani and Jungreis (2003).

1. Introduction

Since the late 90's and early 2000's, much attention has been devoted to 'preferential attachment processes': Networks evolving over time by linking at each time step new nodes to vertices in the existing graph with a probability based on their connectivity. Such schemes relate to 'reinforcement' and other dynamics with a long history (cf. surveys [28], [35], [41]). It has been proposed that versions of these processes may serve as models for growing real-world networks such as the world wide internet web, and types of social structures. See Simon [42] where an early version was introduced to analyze word and science citation distributions among others.

For instance, in a 'friend network', a newcomer may be favorably disposed to link or become friends with an individual with high connectivity, or in other words, one who already has many friends. As has been observed in the literature, when the probability of selecting a vertex is proportional to its degree, the proportions of nodes with degrees $1, 2, \ldots, k, \ldots$ converge as time grows to a power-law distribution $\langle q(k) : k \geq 1\rangle$ where $0 < \lim_{k\uparrow\infty} q(k)k^{\theta} < \infty$ for some $\theta > 0$. Since the sampled empirical degree structure in many real-world networks also has such a power-law form, such preferential attachment processes, in contrast to Erdös-Rényi graphs where the degree structure decays much more rapidly, have become popular: See [1], [5], [8], [9], [13], [14], [18], [21], [30], [31], [32], and references therein.

2000 *Mathematics Subject Classification.* primary 60F17; secondary 05C80, 37H10.

Key words and phrases. preferential attachment, random graphs, degree distribution, fluid limit, law of large numbers, sublinear weights, dynamical system, semigroup, infinite, ODE, uniqueness.

At the same time, other versions of preferential attachment, where the selection probability is a nonlinear function of the connectivity have been considered, and interesting effects have been shown: See, among other works, [12], [17], [20], [24], [33], [39]. In particular, depending on the scheme and the type of nonlinearity, the degree structure asymptotically may be in the form of a 'stretched exponential' or the graph may evolve into a 'condensed' state in which a single (random) vertex may be linked with almost all the incoming nodes.

To be more specific, consider the following preferential attachment model. Suppose at time $n = 0$, the initial network G_0 is composed of two vertices with a single (undirected) edge between them. The dynamics now is that at time $n = 1$, a new vertex is attached to one of the two vertices in G_0 with probability proportional to a function of its degree to form the new network G_1. This scheme continues: More precisely, at time $n + 1$, a new node is linked to vertex $x \in G_n$ with probability proportional to $w(v_n(x))$, that is chance $w(v_n(x))/\sum_{y\in G_n} w(v_n(y))$, where $v_n(z)$ is the degree at time n of vertex z and $w : \mathbb{N} \to \mathbb{R}_+$ is a 'weight' function.

Now, for the moment, to simplify the discussion, let us assume $w(k) = k^\kappa$ for $\kappa > -\infty$. In this way, since the initial graph is a tree, all later networks G_n for $n \geq 0$ are also trees. Let now $Z_k(n)$ be the number of vertices in G_n with degree k, $Z_k(n) = \sum_{y\in G_n} 1(v_n(y) = k)$. In [24], a trichotomy of growth behaviors was observed depending on the strength of the exponent κ.

First, when w is linear, that is when $\kappa = 1$, the scheme is sometimes known as the 'Barabási-Albert' model where the degree structure satisfies, for $k \geq 1$,

$$\lim_{n\to\infty} \frac{Z_k(n)}{n} \;=\; \frac{4}{k(k+1)(k+2)} \quad \text{a.s.}$$

This power-law ($\theta = 3$), in mean-value, through an analysis of rates, was found in [19], [25]. In [7], using difference equations/concentration bounds, the limit was proved in probability. Via Pólya urns, another proof was found yielding a.s. convergence, and also central limit theorems [29]. Also, by embedding into continuous time branching processes, the same a.s. limit was proved in [39]; see also [3] where a different type of embedding was used. A form of Stein's method gives rates of convergence in total variation norm [34], [38]. A large deviation approach also obtains the limit [11].

Next, in the strict sublinear case, when $\kappa < 1$, it was shown that

$$\lim_{n\uparrow\infty} \frac{Z_k(n)}{n} \;=\; q(k) \quad \text{a.s.} \tag{1.1}$$

although q is not a power law, but in form where it decays faster than any polynomial [24], [39]: For $k \geq 1$,

$$q(k) \;=\; \frac{s^*}{k^\kappa}\prod_{j=1}^{k}\frac{j^\kappa}{s^* + j^\kappa}, \quad \text{and } s^* \text{ is determined by } \; 1 \;=\; \sum_{k=1}^{\infty}\prod_{j=1}^{k}\frac{j^\kappa}{s^*+j^\kappa}.$$

Asymptotically as $k \uparrow \infty$, when $0 < \kappa < 1$, $\log q(k) \sim -(s^*/(1-\kappa))k^{1-\kappa}$ is in 'stretched exponential' form; when $\kappa < 0$, $\log q(k) \sim \kappa k \log k$; when $\kappa = 0$, the case of uniform attachment when an old vertex is selected uniformly, $s^* = 1$ and q is geometric: $q(k) = 2^{-k}$ for $k \geq 1$.

In the superlinear case, when $\kappa > 1$, 'explosion' or a sort of 'condensation' happens in that in the limiting graph a random single vertex accumulates most of the connections. In particular, the limiting graph is shown to be a tree

$$\begin{aligned}&\text{where there is a single random vertex with an infinite number of children;}\\&\text{all other vertices have bounded degree, and of these only a finite number}\\&\text{have degree equal or larger than } \lfloor \kappa/(\kappa-1)\rfloor\end{aligned} \tag{1.2}$$

(cf. for a more precise description [33], [24]). Moreover, a corresponding law of large numbers (LLN) limit, $\lim_{n\uparrow\infty} Z_k(n)/n = q(k)$ a.s., is argued where q is degenerate in that $q(1) = 1$ but $q(k) = 0$ for $k \geq 2$ (cf. [24], [21, Chapter 4], [2]).

We now comment on the methods in the papers [39] and [33]. Both use branching process embedding techniques to establish the sublinear and superlinear degree structure results (1.1) and (1.2). More specifically, it seems a tree structure is useful in the proofs, that is the dynamics places no edges between already extant vertices to create cycles.

The purpose of this article is to show the LLN for the degree structure in a general class of 'sublinear' preferential attachment models, including the scheme discussed above, starting from various initial conditions, through a new, different 'fluid limit' approach where cycles may develop. In fact, the method taken here rigorously establishes a mean-field theory for these models. Moreover, the argument seems robust, and might be used in other combinatorial schemes (cf. Remark 2.4).

Specifically, we show (Theorem 2.3) a functional LLN for the degree counts in sublinear generalizations of the urn scheme of Chung, Handjani and Jungreis [12] and the graph model of Chung and Lu [13] (cf. Section 2 for model descriptions and assumptions). Our work answers a question in [12] to show a LLN for the associated degree structure when the weights are sublinear (cf. Remark 2.4).

The 'fluid limit' method is to analyze a more complex problem, namely that of the dynamics of paths $\{n^{-1}Z_k(\lfloor nt\rfloor) : t \in [0,1]\}$ for $k \geq 1$. The value of considering such trajectories is that they inform on the whole historical evolution of the degree structure, not only at a single time. It turns out these paths have nice properties, allowing to characterize their limit points in terms of a system of ordinary differential equations (ODEs) corresponding to a rate formulation of the degree distribution flow (cf. (2.8)). As all the counts $\{Z_k(n)\}$ are coupled together in terms of the total 'weight' of the graph $S(n) = \sum_{k\geq 1} Z_k(n)$ in the selection procedure, the ODE system derived is infinite dimensional and nonlinear, and poses nontrivial difficulties.

The ODEs derived appear natural and may be of interest in other contexts where there is exchange of proportional flow between chains of components. By a change of variables, the ODEs can be written in terms of linear 'Kolmogorov' differential equations (cf. [27]), which can be treated by C_0-semigroup/dynamical systems arguments. In particular, perhaps of interest itself, we show (Theorem 2.1) the ODEs admit a unique nonnegative solution. Therefore, addressing the original 'fluid limit' taken, all the path limit points are the same and so are uniquely characterized.

There is a large literature on fluid limits in various contexts: See [10], [16], [23], [36], [44], [46]) and references therein. Most of this previous development focuses on finite dimensional spaces. In this respect, the current article considers a nontrivial infinite dimensional fluid limit, whose analysis depends on the type of initial condition, namely 'small' versus 'large' (cf. (LIM) in Section 2), which plays a role in the results Theorems 2.1, 2.3. See also [37] for a different infinite dimensional limit in a type of Erdös-Rényi graph.

In the next Section, we detail the preferential attachment models discussed, and state results. Then, proofs of the main convergence and uniqueness results follow in succeeding Sections.

2. Models and Results

To specify the models considered, let $0 \leq p \leq 1$ be a parameter, and let $w : \{1, 2, \ldots\} \to (0, \infty)$ be a positive function, called the 'weight' function.

Graph Model. The following scheme captures the growth of a graph network:

- At time $n = 0$, the initial network G_0 is a finite, possibly disconnected graph.
- At time $n + 1 \geq 1$, form G_{n+1} as follows.
  –With probability $1 - p$, we select independently two old vertices $x, y \in G_n$ with chances $w\big(v_n(x)\big)/S(n)$ and $w\big(v_n(y)\big)/S(n)$ respectively, and then place an edge connecting x and y to form G_{n+1}.
  –However, with probability p, an edge is placed between a new vertex and an old node $x \in G_n$, chosen with probability $w\big(v_n(x)\big)/S(n)$, to form G_{n+1}.

In this model, $v_n(x) \geq 1$ is the degree of the vertex x at time n, and $S(n)$ is the total 'weight' of G_n:

$$S(n) \;=\; \sum_{k\geq 1} w(k) Z_k(n)$$

where $Z_k(n)$ is the count of vertices in G_n with degree k for $k \geq 1$.

Note that it may be possible when two vertices $x, y \in G_n$ are selected, they are the same, which means a 'loop' is added to the graph at $x = y$, and our convention here is that the degree of vertex $x = y$ is incremented by 2. In particular, at each time, the total degree of the graph increments by 2. However, as the successive independent choices in actions are random, the total number of vertices at time $n \geq 1$ is $V(G_0)$ plus a sum of n independent Bernoulli(p) variables; here, $V(G_0)$ is the initial number of vertices.

Cases of the above dynamics include the following:

- When $p = 1$ and $w(k) = k^\kappa$, the dynamics matches the preferential attachment graph scheme mentioned in the Introduction.
- When $w(k) = k$ and p is arbitrary, the scheme is discussed in [13] and the LLN result is proved: $\lim_{n\uparrow\infty} Z_k(n)/n = q(k)$ where q is in 'power law' form with $\theta = 1 + 2/(2 - p)$. Also, an associated central limit theorem for the 'leaves', nodes of degree 1, has been found in [45].
- If $p = 0$, the dynamics would always add edges and loops with respect to the initial graph G_0. We will avoid this 'degenerate' growth in what follows.

Urn Model. Consider the following urn dynamics which builds an evolving collection of urns:

- At time $n = 0$, the initial collection G_0 is a finite set of nonempty urns, each containing a finite number of balls.

- At time $n+1 \geq 1$, G_{n+1} is built as follows.
  –With probability p, a new urn with a single ball is added to the collection to form G_{n+1}.
  –However, with probability $1-p$, we select an urn x from G_n with probability $w\big(v_n(x)\big)/S(n)$, and place a new ball into it to form G_{n+1}.

Here, $v_n(x) \geq 1$ is the size or number of balls in the urn x at time n, and $S(n)$ is the total 'weight' of collection G_n:

$$S(n) \ = \ \sum_{k \geq 1} w(k) Z_k(n)$$

where $Z_k(n)$ is the number of urns in G_n with exactly k balls for $k \geq 1$.

We note that the total size or number of balls increments by 1 at each time, but as in the graph model, at time $n \geq 1$, the total number of urns is random, namely $U(G_0)$ plus the sum of n independent Bernoulli(p) variables, where $U(G_0)$ is the initial number of urns.

We now remark on some cases of the above dynamics:

- When $w(k) = k^\kappa$, the scheme is discussed in [12], and results on the evolution are given when $\kappa \geq 1$. However, when $\kappa < 1$, a LLN is stated, but the convergence of $Z_k(n)/n$ is left open (cf. Remark 2.4).
- If $p = 1$, the scheme would always add an urn with a single ball to the collection at each time. Also, if $p = 0$, no new urns are added and only the urns in the initial collection grow. Both are 'degenerate' evolutions which we will avoid in assumption (P) below.

Assumptions. We now give assumptions on p with respect to the two models, and on the weight function w under which results are stated.

(P) To avoid 'degeneracies', in the graph model, p is taken $0 < p \leq 1$. However, in the urn scheme, we assume that $0 < p < 1$.

(SUB) We have

$$\lim_{k \uparrow \infty} \frac{w(k)}{k} = 0. \quad \text{Hence, } \sup_{k \geq 1} \big[w(k)/k\big] \leq \mathcal{W} \text{ for some constant } \mathcal{W} < \infty.$$

This large class of weights $w(\cdot)$ includes in particular the well-studied case $w(k) = k^\kappa$ for $\kappa < 1$ discussed in the Introduction.

Markov Structure. We now derive the evolution scheme of the counts $\{Z_k(n)\}$ for $n \geq 0$. Define sigma-fields $\mathcal{F}_j = \sigma\{\{Z_k(\ell)\} : 0 \leq \ell \leq j\}$ for $j \geq 0$. Note also, in both models, as $w(\cdot) > 0$, the quantity $S(n) > 0$ for all $n \geq 0$. For each $k \geq 1$, let

$$Z_k(j+1) - Z_k(j) \ =: \ d_k(j+1). \tag{2.1}$$

Given the counts at time j, $\{Z_k(j)\}$, the distribution of the difference $d_k(j+1)$ is as follows.

With respect to the urn scheme,

$$d_1(j+1) \ = \ \begin{cases} 1 & \text{with prob. } p \\ -1 & \text{with prob. } (1-p)\frac{w(1)Z_1(j)}{S(j)} \\ 0 & \text{with prob. } (1-p)\big[1 - \frac{w(1)Z_1(j)}{S(j)}\big]. \end{cases}$$

Also, for $k \geq 2$,

$$d_k(j+1) \;=\; \begin{cases} \;\;1 & \text{with prob. } (1-p)\frac{w(k-1)Z_{k-1}(j)}{S(j)} \\ -1 & \text{with prob. } (1-p)\frac{w(k)Z_k(j)}{S(j)} \\ \;\;0 & \text{with prob. } 1-(1-p)\big[\frac{w(k-1)Z_{k-1}(j)}{S(j)} + \frac{w(k)Z_k(j)}{S(j)}\big]. \end{cases}$$

In words, at step j, an urn with one ball is added, increasing the count of singleton urns, with probabiity p. But, with probability $(1-p)Z_k(j)/S(j)$, an urn with size k is selected and a ball put into it, increasing the number of urns with $k+1$ balls and decreasing the count of urns with k balls.

In the graph scheme, $d_k(j+1)$ has a similar, but more involved expression as possible loops need to be considered. These formulas are given in the Appendix. We explain though the structure of $d_1(j+1)$ in words; a similar reasoning process underlies the other expressions in the Appendix. The difference $d_1(j+1)$ equals 1, when a new vertex arrives at step j to attach to a non-leaf vertex, with probability $p[1-w(1)Z_1(j)/S(j)]$. However, the difference is -1 when either a new edge is put between two vertices, one a leaf and the other a non-leaf, or the new edge is put as a loop connecting a single leaf vertex to itself; this first possibility occurs with probability $2(1-p)[w(1)Z_1(j)/S(j)][1-w(1)Z_1(j)/S(j)]$, and since there are $Z_1(j)$ leaves, the second possibility happens with chance $(1-p)[w(1)^2Z_1(j)/S(j)]$. Also, the difference is -2 when a new edge is put between distinct leaf vertices, with probability $(1-p)[w(1)Z_1(j)/S(j)]^2 - (1-p)[w(1)^2Z_1(j)/S(j)]$. Otherwise, the difference vanishes with the remaining probability.

These schemes may be understood as infinite dimensional Markov chains. For instance, in the urn model, let $\mathbf{Z}(n) = \langle Z_1(n), Z_2(n), \ldots\rangle$. Then, $\mathbf{Z}(n)$ has transition probability

$$P\big(\mathbf{Z}(n+1) = \mathbf{Z}(n) + e_{k+1} - e_k \big| \mathbf{Z}(n) = \mathbf{z}\big) \;=\; \begin{cases} p & k=0 \\ (1-p)\beta_k[\mathbf{z}] & k \geq 1 \end{cases}$$

Here, e_0 is the vector with all components 0, and e_k is the vector with a 1 in the kth position and 0 elsewhere for $k \geq 1$. Also, $\beta_k[\mathbf{z}] := w(k)\mathbf{z}_k/\sum_{\ell\geq 1} w(\ell)\mathbf{z}_\ell$, well defined when $\mathbf{z}$ is not the zero vector. Note that $\beta_k[\mathbf{z}]$ satisfies a 'scale-invariant' property, that is $\beta_k[\lambda\mathbf{z}] = \beta_k[\mathbf{z}]$, which will be important in the later proofs.

A similar, but more involved transition probability can be associated to the graph count evolution. Again, although the transition probability is not 'scale-invariant', as in the urn model, all its dominant terms, that is, those not associated with loops, are 'scale-invariant'.

Scaled Processes. Consider now an array of counts $\{Z_k^{(n)}(\cdot) : k \geq 1\}$ and weights $\{S^{(n)}(\cdot)\}$ for $n \geq 1$ where in the nth row the underlying process begins from initial network $G_0^{(n)}$. We do not prescribe the joint distribution of the rows as it will not matter in what follows. However, such an array allows to introduce time and space scales into the degree trajectories. Define the family of linearly interpolated processes $\{X_k^{(n)}(t) : t \geq 0, k \geq 1\}$ for $n \geq 1$, which place the proportion of counts $\{Z_k^{(n)}(j)/n\}$ into continuous time trajectories, by

$$X_k^{(n)}(t) \;:=\; \frac{1}{n}Z_k^{(n)}(\lfloor nt\rfloor) + \frac{nt - \lfloor nt\rfloor}{n}\left(Z_k^{(n)}(\lceil nt\rceil) - Z_k^{(n)}(\lfloor nt\rfloor)\right).$$

The paths $X_k^{(n)} : [0,\infty) \to \mathbb{R}_+$, in both schemes as $|d_k(j+1)| \leq 2$, belong to the space of Lipschitz functions with constant at most 2. For $t \geq 0$, let also $\mathcal{S}^{(n)}$ be the continuous interpolation of the weights $S^{(n)}$,

$$\mathcal{S}^{(n)}(t) \;:=\; \frac{1}{n}S^{(n)}(\lfloor nt\rfloor) + \frac{nt-\lfloor nt\rfloor}{n}\left(S^{(n)}(\lceil nt\rceil) - S^{(n)}(\lfloor nt\rfloor)\right). \tag{2.2}$$

With respect to constants $c_k^{(n)}, c^{(n)}, \tilde{c}^{(n)} \geq 0$ for $k \geq 1$, let

$$c_k^{(n)} \;:=\; \frac{1}{n}Z_k^{(n)}(0), \quad c^{(n)} \;:=\; \sum_{k\geq 1} c_k^{(n)}, \quad \text{and} \quad \tilde{c}^{(n)} := \sum_{k\geq 1} kc_k^{(n)}.$$

We will impose the following initial laws of large numbers:

(LIM) For constants $c_k, c, \tilde{c} \geq 0$, we have

$$c_k := \lim_{n\uparrow\infty} c_k^{(n)} \quad \text{and} \quad \tilde{c} := \sup_{n\geq 1} \tilde{c}^{(n)} < \infty.$$

Hence, $c := \lim_{n\uparrow\infty} c^{(n)} = \sum_{k\geq 1} c_k < \infty$ and $\tilde{c} \geq \sum_{k\geq 1} kc_k$.

Here, noting $\sum_{k\geq L} c_k^{(n)} \leq L^{-1}\sum_{k\geq 1} kc_k^{(n)} \leq \tilde{c}/L$, the $c^{(n)}$-limit follows. The $\tilde{c}$-inequality is Fatou's inequality. Also, as $\sum_{k\geq L} w(k)c_k^{(n)} \leq \tilde{c}[\sup_{j\geq L} w(j)/j]$, from (SUB), we may also conclude the initial weights $\sum_{k\geq 1} w(k)c_k^{(n)} \to \sum_{k\geq 1} w(k)c_k$.

We will say that the initial configuration is a 'small' configuration if $c_k \equiv 0$, and is a 'large' one if $c_k > 0$ for some $k \geq 1$. In a small configuration, the total degree/size of $G_0^{(n)}$ is $o(n)$. In particular, if the total degree/size of $G_0^{(n)}$ is uniformly bounded, for example $G_0^{(n)} \equiv G_0$ is fixed, the initial configuration is a small one. However, in a large configuration, the initial networks are already developed in that their degree/size is at least εn for some $\varepsilon > 0$. We remark similar initial conditions were used in a different context in [11].

Heuristics. First, with respect to small initial configurations, we now try to guess the long term behavior of $X_k^{(n)}(t)$ as $n \uparrow \infty$ for $t \geq 0$ and $k \geq 1$. In the sublinear weights setting, as indicated in the Introduction, all vertices should grow at the same rate, and there should be no 'explosion'. Suppose, a.s. $\lim_{n\uparrow\infty} X_k^{(n)}(t) = a_k t$ and $\lim_{n\to\infty} \mathcal{S}^{(n)}(t) = bt$ for constants $\{a_k\}$ and $b = \sum_{k\geq 1} w(k)a_k > 0$. Heuristically, for $j = \lfloor nt\rfloor$, one might approximate the increment $d_k(j+1)$, whose distribution depends on $\{Z_k(j)\}$ in a 'scale invariant' way, by an increment not dependent on n, where $Z_k(j)/nt$ and $S(j)/nt$ are set to a_k and b respectively. Then, $X_k^{(n)}(t) = Z_k(\lfloor nt\rfloor)/n$ can be approximated by the scaled position of a random walk corresponding to the approximated differences.

Let $p_0, q_0 > 0$. In the following, under (P), fix $p_0 = p$ and $q_0 = 2-p$ when considering the graph model. However, in the urn model, set the parameters $p_0 = p$ and $q_0 = 1-p$.

Then, we may calculate the rates $a_k = a_k(p_0, q_0, b)$ of growth of the number of vertices or urns with size k in terms of the inflow/outflow rates given by the structure of the count differences (cf. (2.1)):

$$\begin{aligned} a_1 &= p_0 - \frac{q_0 w(1)a_1}{b} && (2.3)\\ a_k &= q_0\frac{w(k-1)a_{k-1}}{b} - q_0\frac{w(k)a_k}{b} && \text{for } k \geq 2. \end{aligned}$$

Solving for the $\{a_k\}$, in terms of b, one obtains

$$a_1 = \frac{bp_0}{b+q_0 w(1)}, \text{ and } a_k = \frac{q_0 w(k-1)a_{k-1}}{b+q_0 w(k)} \quad \text{for } k \geq 2.$$

Therefore, for $k \geq 2$,

$$\begin{aligned} a_k &= a_1 \prod_{j=1}^{k-1} \frac{q_0 w(j)}{b+q_0 w(j+1)} \\ &= a_1 \frac{b+q_0 w(1)}{b+q_0 w(k)} \prod_{j=1}^{k-1} \frac{q_0 w(j)}{b+q_0 w(j)} \\ &= a_1 \frac{b+q_0 w(1)}{b} \Big[\prod_{j=1}^{k-1} \frac{q_0 w(j)}{b+q_0 w(j)} - \prod_{j=1}^{k} \frac{q_0 w(j)}{b+q_0 w(j)} \Big]. \end{aligned} \tag{2.4}$$

Noting $\{a_k\}$ above, under (SUB), the equation $b = \sum_{k\geq 1} w(k)a_k$ takes form

$$b = \frac{bp_0}{q_0} \sum_{k\geq 1} \prod_{j=1}^{k} \frac{q_0 w(j)}{b+q_0 w(j)}$$

or $1 = F_{p_0,q_0}(b)$ where $F_{p_0,q_0} : [0,\infty) \to \mathbb{R} \cup \{\infty\}$ is a positive function defined by

$$F_{p_0,q_0}(s) = \frac{p_0}{q_0} \sum_{k\geq 1} \prod_{j=1}^{k} \frac{q_0 w(j)}{s+q_0 w(j)}.$$

One can interpret $sF_{p_0,q_0}(s)$ as the $n \uparrow \infty$ limit of the total scaled weight given to vertices or urns with finite size, when the $n \uparrow \infty$ limit of the scaled weight of G_n or U_n is s. The equation $s = sF_{p_0,q_0}(s)$ means that none of the scaled weight of G_n or U_n escapes to an ∞-weight component.

By (SUB), $F_{p_0,q_0}(s) < \infty$ for each $s > 0$ and $\lim_{s\downarrow 0} F_{p_0,q_0}(s) = \infty$. In particular, there exists a number $s_0 = s_0(p_0, q_0) > 0$ such that

$$1 < F_{p_0,q_0}(s_0) < \infty. \tag{2.5}$$

The function F_{p_0,q_0} strictly decreases on $[s_0, \infty)$ and $\lim_{s\uparrow\infty} F_{p_0,q_0}(s) = 0$. Therefore,

$$\text{there exists a unique } s^* = s^*(p_0, q_0) > s_0 \text{ such that } F_{p_0,q_0}(s^*) = 1. \tag{2.6}$$

Hence, the parameter b above is identified from (2.6) as $b = s^*$, and $\{a_k\}$ is determined implicitly in terms of $w(\cdot)$, p_0 and q_0 through (2.3). Of course, when $w(k) = k^\kappa$ for $\kappa < 1$, $a_k = q(k)$ for $k \geq 1$ (cf. (1.1)).

We now observe

$$\sum_{k\geq 1} a_k = p_0 \quad \text{and} \quad \sum_{k\geq 1} ka_k = p_0 + q_0. \tag{2.7}$$

The first sum follows from a straightforward use of the last line of (2.4) and the first line of (2.3). Using $\sum_{k\geq 1} a_k = p_0$, the last line of (2.4), and the first line of

(2.3), the second sum is transformed

$$\begin{aligned}\sum_{k\geq 1} k a_k &= \sum_{k\geq 1}\sum_{\ell\geq k} a_\ell \\ &= p_0 + \sum_{k\geq 2} a_1 \frac{b+q_0w(1)}{b}\prod_{j=1}^{k-1}\frac{q_0w(j)}{b+q_0w(j)} \\ &= p_0 + p_0\sum_{k\geq 2}\prod_{j=1}^{k-1}\frac{q_0w(j)}{b+q_0w(j)}.\end{aligned}$$

Now, noting $F_{p_0,q_0}(b) = 1$, we recover the second sum formula in (2.7).

At this point, for both small and large initial configurations, one might infer an a.s. 'continuous' version of (2.3), a rate formulation for the limit of the functions $\{X_k^{(n)}\}$. These limit functions are nonlinear under large initial configurations, in contrast to when the system starts from small initial configurations. That is, suppose a.s. $X_k^{(n)} \to \varphi_k$ and $S^{(n)} \to \sum_{k\geq 1} w(k)\varphi_k$ on a subsequence as $n \uparrow \infty$. Then, these limit functions $\varphi_k(\cdot) = \varphi_k(\cdot; p_0, q_0)$ are nonnegative and satisfy the integral form of a coupled system of ODEs:

$$\begin{aligned}\dot\varphi_1(t) &= p_0 - \frac{q_0w(1)\varphi_1(t)}{\sum_{\ell\geq 1} w(\ell)\varphi_\ell(t)}, \\ \dot\varphi_k(t) &= \frac{q_0}{\sum_{\ell\geq 1} w(\ell)\varphi_\ell(t)}\big[w(k-1)\varphi_{k-1}(t) - w(k)\varphi_k(t)\big], \quad \text{for } k\geq 2.\end{aligned} \tag{2.8}$$

with initial condition $\varphi_k(0) = c_k$ for $k \geq 1$. Under small initial configurations ($c_k \equiv 0$), the ODE is singular at $t = 0$, although one can inspect that the linear functions $\{\varphi_k(t) = a_k(p_0, q_0, s^*)t : k \geq 1\}$ are a solution since $\{a_k(p_0, q_0, s^*) : k \geq 1\}$ verifies (2.3). However, under either small or large initial configurations, it does not seem easy to conclude the ODEs have a unique nonnegative solution.

But, one can think of the ODE system in the following way: First, we observe that $\sum_{k\geq 1}\dot\varphi_k(t) = p_0$. Introduce now a time-change $t = t(s)$ satisfying $\dot t = T(t)$ where $T(t) = \sum_{k\geq 1} w(k)\varphi_k(t)$ and $t(0) = 1$. Then, $\{\psi_k(s) = \varphi_k(t(s)) : k \geq 1\}$ satisfies the integral form of the following autonomous system:

$$\begin{aligned}\dot\psi_1(s) &= (p_0 - q_0)w(1)\psi_1(s) + p_0\sum_{k\geq 2} w(k)\psi_k(s), \\ \dot\psi_k(s) &= q_0\big[w(k-1)\psi_{k-1}(s) - w(k)\psi_k(s)\big], \qquad \text{for } k\geq 2.\end{aligned} \tag{2.9}$$

It will turn out that one can associate to this system a strongly continuous positive semigroup P_s whose essential growth rate is nonpositive. Such semigroups have useful decompositions and possess a 'Perron-Frobenius' eigenvector and eigenvalue. Moreover, it turns out the time-change $t(\cdot)$ can be determined in terms of P_s, allowing to characterize solutions of (2.8).

Theorem 2.1. *Suppose $p_0, q_0 > 0$ and conditions (SUB) and (LIM) hold, and recall the parameter s^* in* (2.6). *Then, under both small and large initial configurations, there is a unique nonnegative solution $\{\varphi_k(\cdot)\}$ of the integral form of ODEs* (2.8).

Moreover, for $k \geq 1$, under small initial configurations, $\varphi_k(t) = a_k(p_0, q_0, s^*)t$. Also, with respect to large initial configurations, $\lim_{t\uparrow\infty} t^{-1}\varphi_k(t) = a_k(p_0, q_0, s^*)$.

Remark 2.2. The proof of Theorem 2.1 is shorter under large initial configurations as there is no time singularity at $t = 0$. In this case, the solution is found implicitly in terms of the semigroup P_s and time-change $t = t(s)$.

However, the full machinery of 'quasicompact' semigroup asymptotics and the assumption $c = 0$ are used in the small initial configuration case. See the beginning of Section 4 for more remarks on the strategy of the proof.

Finally, we note s^* can be identified in terms of the 'Perron-Frobenius' eigenvalue alluded to earlier (cf. Propositions 4.8 and 4.9).

We now assert that the heuristic derivations (2.3) and (2.8) are correct.

Theorem 2.3. *Suppose conditions (P), (SUB), and (LIM) hold. Let* $\{\varphi_k(\cdot, p_0, q_0) : k \geq 1\}$ *be the unique nonnegative solution found in Theorem 2.1 to the integral form of ODEs* (2.8)*. Then, with respect to the graph and urn models, for* $k \geq 1$*, uniformly on compact time sets, we have*

$$\lim_{n\uparrow\infty} X_k^{(n)}(t) \;=\; \varphi_k(t; p_0, q_0) \quad \text{a.s.}$$

Remark 2.4. With respect to the urn model, when $w(k) = k^\kappa$ for $\kappa < 1$, the form of $\{a_k\}$ was derived in [12]. However, it was left open in [12] to show that the LLN, under small initial configurations, $\lim_{n\uparrow\infty} Z_k(n)/n = a_k$ holds for $k \geq 1$. In this context, a contribution of Theorem 2.3 is to give a proof of this limit.

The fluid limit argument given seems of potential use in other nonlinear preferential attachment schemes. In particular, the approach should hold for models where at each time only a finite number of vertices/edges or balls/urns are added. In this case, the differences $d_k(j + 1)$ are still uniformly bounded in k, j and the paths $X_k^{(n)}(t)$ will be Lipschitz, a primary ingredient in the proof.

In addition, although (SUB) excludes the "linear weights" case $w(k) = k + m$ for $k \geq 1$ and $m > -1$, since in this case S_n acts as an affine function of n, and the corresponding ODEs (2.8) can be uniquely integrated (cf. Corollary 1.7 in [11]), a similar fluid limit argument yields yet another proof in this situation.

Relation of (SUB) to Literature. We note, in [39], which proves, among other results, the a.s. LLN $\lim_{n\uparrow\infty} Z_k(n)/n = a_k$ for $k \geq 1$ in the graph model when $p = 1$, the only condition assumed on w is that (2.5) holds, which is more general than (SUB). For instance, a 'linear order'-type weight where $\liminf_{k\uparrow\infty} w(k)/k > 0$ and $\limsup_{k\uparrow\infty} w(k)/k < \infty$ is not allowed under (SUB), although (2.5) would be satisfied. However, as remarked in [39], (2.5) itself is only given as a sufficient condition.

A necessary condition might include the requirement, $\sum_{k\geq 1} 1/w(k) = \infty$, implied by (2.5), although this is not pursued here. In this respect, we note Theorem 1.1(ii) in [15] shows that $\sum_{k\geq 1} 1/w(k) = \infty$ is a necessary and sufficient condition for all vertices/urns to have infinite size in the limit network a.s.

In this article, the assumption (SUB) is used to enforce control on the tails of the weight sum $S(n)$, so that the limits of $S(n)/n$ and $\{Z_k(n)/n\}$ can be related (cf. Step 2, proof of Theorem 2.3). (SUB) is also useful in the proof of uniqueness of solution to the infinite dimensional ODE system derived (cf. Section 4).

3. Proof of Theorem 2.3.

We will assume Theorem 2.1, proved in the next Section, and prove Theorem 2.3 in several steps.

Step 1. Since $d_k(j)$ is uniformly bounded, $\|d_k(j)\|_{L^\infty} \leq 2$, we have, for each realization of the evolving scheme, that $X_k^{(n)}$ are Lipschitz functions with constant 2 for all $k \geq 1$ and $n \geq 1$. Since $X_k^{(n)}(0) = c_k^{(n)}$ converges to c_k, by equicontinuity and local compactness of $[0,\infty)$, we may take a diagonal subsequence n_m so that $X_k^{(n_m)}$ converges uniformly for t in compact subsets of $[0,\infty)$ to a Lipschitz function φ_k with constant 2, for each $k \geq 1$, which may depend on the realization: For $N > 0$,

$$\lim_{m\uparrow\infty} \sup_{t\in[0,N]} \left|X_k^{(n_m)}(t) - \varphi_k(t)\right| \;=\; 0.$$

Step 2. With respect to the graph model, as the total degree increments by 2 at each time, we have $\sum_{k\geq 1} kZ_k^{(n)}(n) = n\tilde{c}^{(n)} + 2n$. On the other hand, in the urn scheme, the total number of balls increases by 1 at each time, and so the total size $\sum_{k\geq 1} kZ_k(n) = n\tilde{c}^{(n)} + n$. Hence, in both models, given $\sum_{k\geq 1} kX_k^{(n)}(t) \leq \tilde{c}^{(n)} + 2t$, we have, for each $L \geq 1$, that

$$\sum_{k\geq L} w(k)X_k^{(n_m)}(t) \;\leq\; \big[\sup_{k\geq L} w(k)/k\big] \sum_{k\geq 1} kX_k^{(n_m)}(t) \;\leq\; \big[\sup_{k\geq L} w(k)/k\big](\tilde{c}^{(n)} + 2t). \tag{3.1}$$

Therefore,

$$\sum_{k\leq L} w(k)X_k^{(n_m)}(t) \;\leq\; \mathcal{S}^{(n_m)}(t) \;\leq\; \sum_{k\leq L} w(k)X_k^{(n_m)}(t) + (\tilde{c}^{(n)} + 2t)[\sup_{k>L} w(k)/k]$$

and also, for $N > 0$, noting (LIM),

$$\lim_{n_m\uparrow\infty} \sup_{t\in[0,N]} \Big|\mathcal{S}^{(n_m)}(t) - \sum_{k\leq L} w(k)\varphi_k(t)\Big| \;\leq\; (\tilde{c} + 2N)[\sup_{k>L} w(k)/k].$$

In addition, by Fatou's lemma, from (3.1), we obtain $\sum k\varphi_k(t) \leq \tilde{c} + 2t$. Then,

$$\sup_{t\in[0,N]} \sum_{k>L} w(k)\varphi_k(t) \;\leq\; \big[\sup_{k>L} w(k)/k\big] \sup_{t\in[0,N]} \sum_{k\geq 1} k\varphi_k(t) \;\leq\; (\tilde{c}+2N)\big[\sup_{k>L} w(k)/k\big].$$

Putting together these estimates, we have for each $L \geq 1$ that

$$\lim_{n_m\uparrow\infty} \sup_{t\in[0,N]} \Big|\mathcal{S}^{(n_m)}(t) - \sum_{k\geq 1} w(k)\varphi_k(t)\Big| \;\leq\; 2(\tilde{c} + 2N)[\sup_{k>L} w(k)/k].$$

Therefore, by assumption (SUB), taking $L \uparrow \infty$, we have

$$\mathcal{S} := \lim_{m\uparrow\infty} \mathcal{S}^{(n_m)} \;=\; \sum_{k\geq 1} w(k)\varphi_k$$

converges uniformly for $t \in [0, N]$. Since $\{\mathcal{S}^{(n_m)}\}$ are continuous functions, we see also that $\mathcal{S}$ is a continuous function.

Step 3. We now derive bounds for the limit function $\mathcal{S}$. Under (SUB) and (LIM), in both models, given the bound $(\tilde{c} + 2)n$ on the total degree/size of the network at time n, we have

$$\mathcal{S}^{(n)}(t) \;=\; \sum_{k\geq 1} w(k)X_k^{(n)}(t) \;\leq\; \mathcal{W}\sum_{k\geq 1} kX_k^{(n)}(t) \;\leq\; (\tilde{c} + 2t)\mathcal{W}.$$

Also, in both models, for $L \geq 1$, we have

$$\begin{aligned}\mathcal{S}^{(n)}(t) &\geq \big(\inf_{k\leq L} w(k)\big)\sum_{k\leq L} X_k^{(n)}(t)\\ &\geq \big(\inf_{k\leq L} w(k)\big)\Big[\sum_{k\geq 1} X_k^{(n)}(t) - \frac{1}{L+1}\sum_{k\geq 1} kX_k^{(n)}(t)\Big]\\ &\geq \big(\inf_{k\leq L} w(k)\big)\Big[\sum_{k\geq 1} X_k^{(n)}(t) - \frac{\tilde{c}+2t}{L+1}\Big].\end{aligned}$$

Since, in both models, at time $n \geq 1$, the number of vertices/urns at time n equals $nc^{(n)}$ plus the sum of n independent Bernoulli(p) variables, we have

$$\lim_{n_m\uparrow\infty}\sum_{k\geq 1} X_k^{(n_m)}(t) \ = \ c + pt \quad \text{a.s.}$$

Therefore, with $\hat{L}$ such that $\tilde{c}/(\hat{L}+1) \leq c/2$ and $2/(\hat{L}+1) \leq p/2$, we conclude from the above estimates that

$$2^{-1}(c+pt)\big(\inf_{k\leq \hat{L}} w(k)\big) \ \leq \ \mathcal{S}(t) \ \leq \ (\tilde{c}+2t)\mathcal{W}.$$

Step 4. From the decomposition (2.1), we have, for $k \geq 1$, that

$$X_k^{(n)}(t) - X_k^{(n)}(0) = M_k^{(n)}(\lfloor nt\rfloor) + \frac{1}{n}\sum_{j=0}^{\lfloor nt\rfloor -1} E\big[d_k^{(n)}(j+1)\big|\mathcal{F}_j\big] + \frac{nt-\lfloor nt\rfloor}{n} d_k(\lfloor nt\rfloor+1)$$

where

$$M_k^{(n)}(\ell) \ = \ \frac{1}{n}\sum_{j=0}^{\ell-1}\Big(d_k^{(n)}(j+1) - E\big[d_k^{(n)}(j+1)\big|\mathcal{F}_j\big]\Big)$$

is a martingale with respect to $\{\mathcal{F}_\ell : \ell \geq 0\}$, and, for the urn scheme,

$$E\big[d_k^{(n)}(j+1)\big|\mathcal{F}_j\big] \ = \ \begin{cases} p - (1-p)\frac{w(1)X_1^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} & \text{for } k=1\\ \frac{(1-p)w(k-1)X_{k-1}^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} - \frac{(1-p)w(k)X_k^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} & \text{for } k\geq 2\end{cases}$$

and, for the graph model, after calculating with $\{d_k^{(n)}\}$ in the Appendix,

$$\begin{aligned}&E\big[d_k^{(n)}(j+1)\big|\mathcal{F}_j\big]\\ &= \begin{cases} p-(2-p)\frac{w(1)X_1^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} \\ \qquad + \frac{1-p}{n}\Big\{\frac{[w(1)]^2X_1^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2}\Big\} & \text{for } k=1\\[2ex] \frac{(2-p)w(1)X_1^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} - \frac{(2-p)w(2)X_2^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} \\ \qquad + \frac{1-p}{n}\Big\{\frac{-2[w(1)]^2X_1^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2} + \frac{[w(2)]^2X_2^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2}\Big\} & \text{for } k=2\\[2ex] \frac{(2-p)w(k-1)X_{k-1}^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} - \frac{(2-p)w(k)X_k^{(n)}(j/n)}{\mathcal{S}^{(n)}(j/n)} \\ \qquad + \frac{1-p}{n}\Big\{\frac{[w(k-2)]^2X_{k-2}^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2} + \frac{-2[w(k-1)]^2X_{k-1}^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2} & \text{for } k\geq 3,\\ \qquad\qquad + \frac{[w(k)]^2X_k^{(n)}(j/n)}{[\mathcal{S}^{(n)}(j/n)]^2}\Big\} \end{cases}\end{aligned}$$

Step 5. Let $\langle M_k^{(n)}(j)\rangle$ be the quadratic variation of $M_k^{(n)}(j)$. In our context, noting $|d_k^{(n)}(j)| \leq 2$ is bounded, we have

$$|\langle M_k^{(n)}(\lfloor nt\rfloor)\rangle| \;=\; \frac{1}{n^2}\sum_{j=0}^{\lfloor nt\rfloor-1}\left(d_k^{(n)}(j+1) - E[d_k^{(n)}(j+1)|\mathcal{F}_j]\right)^2 \;\leq\; Ctn^{-1}.$$

Therefore, for $\epsilon > 0$, by Burkholder-Davis-Gundy inequalities, we have

$$\begin{aligned} P\Big(\sup_{s\in[0,N]}|M_k^{(n)}(\lfloor ns\rfloor)| > \epsilon\Big) \;&\leq\; \frac{1}{\epsilon^4}E\Big[\max_{0\leq j\leq \lfloor nN\rfloor}|M_k^{(n)}(j)|^4\Big] \\ &\leq\; CE\Big[\langle M_k^{(n)}(\lfloor nN\rfloor)\rangle^2\Big] \;\leq CN^2n^{-2}. \end{aligned}$$

Then, by Borel-Cantelli lemma, $\lim_{n\uparrow\infty}\sup_{t\in[0,N]}|M_k^{(n)}(\lfloor nt\rfloor)| = 0$ a.s.

Step 6. To obtain an integral equation, from the development in Step 4, since $X_k^{(n)}(0) = c_k^{(n)} \to c_k$ by (LIM) and $n^{-1}d_k(\lfloor nt\rfloor + 1)$, $M_k^{(n)}(\cdot)$ vanish uniformly a.s., we need only evaluate the limit of

$$\frac{1}{n_m}\sum_{j=0}^{\lfloor n_m t\rfloor-1}E\big[d_k^{(n_m)}(j+1)\big|\mathcal{F}_j\big] \;=\; \int_0^t E\big[d_k^{(n_m)}(\lceil n_m s\rceil)\big|\mathcal{F}_{\lfloor n_m s\rfloor}\big]ds.$$

By Steps 2 and 3, given uniformly $X_k^{(n_m)}(s) \to \varphi_k(s)$ and $\mathcal{S}^{(n_m)}(s) \to \mathcal{S}(s)$ for $s \in [0, N]$, and positivity of $\mathcal{S}^{(n_m)}(s)$ and $\mathcal{S}(s)$ for $s > 0$, in both models, we have for $0 < s \leq N$,

$$\begin{aligned} &\lim_{m\uparrow\infty}E\big[d_k^{(n_m)}(\lceil n_m s\rceil)\big|\mathcal{F}_{\lfloor n_m s\rfloor}\big] \\ &\quad= \begin{cases} p_0 - q_0\frac{w(1)\varphi_1(s)}{\mathcal{S}(s)} & \text{for } k=1 \\ \frac{q_0}{\mathcal{S}(s)}\big[w(k-1)\varphi_{k-1}(s) - w(k)\varphi_k(s)\big] & \text{for } k\geq 2. \end{cases} \end{aligned}$$

Given the pointwise bound $|d_k^{(n_m)}| \leq 2$, by dominated convergence, as $n_m \uparrow \infty$, we conclude $\{\varphi_k : k \geq 1\}$ satisfies the integral equation corresponding to (2.8) with initial condition $\varphi_k(0) = c_k$ for $k \geq 1$.

Finally, as $\{\varphi_k(t)\}$ is nonnegative by construction, by Theorem 2.1, we conclude it is the unique solution to the ODEs (2.8). Hence, for each realization in a full probability set, this solution $\{\varphi_k\}$ is the unique limit family for the evolutions $\{X_k^{(n)}\}$. Therefore, we conclude a.s., as $n \uparrow \infty$, that $X_k^{(n)}$ converges uniformly, on compact time sets, to φ_k, for $k \geq 1$. □

4. Proof of Theorem 2.1.

As noted, in Chapter VI in [22], 'semigroups are everywhere', and are useful in many applications. In this vein, our strategy is to exploit the properties of a semigroup associated to the transformed ODEs (2.9). The following outline summarizes the main steps. Throughout the section, the condition (SUB) will be assumed.

Outline. The first step is to give useful estimates (Proposition 4.1) on the scaled numbers of vertices/urns, weight, and degree/size of the system. These estimates allow to define the time-scale $\dot{t}(s) = T(t)$ and the system $\psi_k(s) = \varphi_k(t(s))$ for $k \geq 1$ given in (2.9).

Next, we construct a strongly continuous semigroup P_t corresponding to (2.9), and estimate its growth (Proposition 4.2). Although the transformed ODEs (2.9) do not fit into the general theoretical framework of linear 'Kolmogorov' differential systems, recently considered in certain host patch/parasite models [26], [27], [6], nevertheless one can use this framework to prove Proposition 4.2.

The third step is to show positivity of P_t (Proposition 4.3), which will allow the use of certain 'Perron-Frobenius' operator results.

The fourth step is to show that $e^{-\varepsilon t}P_t$, for $\varepsilon > 0$, is 'quasicompact' (Propositions 4.4 and 4.5). Such a result is a statement about the 'essential' growth rate of P_t. Importantly, it allows to approximate the semigroup by a 'finite-dimensional' evolution operator (see [4] for a discussion of the use of 'quasicompactness' with respect to ergodic theory). Perhaps, of interest, we show that P_t is not compact, nor 'eventually' compact in the Appendix (Proposition 6.1), both properties which have been useful in the study of population evolutions (cf. Section VI.1 in [22] and references therein).

In the next step, we further identify this finite-dimensional evolution through some computations and a Perron-Frobenius theorem (Propositions 4.6, 4.7, 4.8 and 4.9). At this point, one can complete the proof of Theorem 2.1 with respect to large initial conditions.

However, under small initial configurations, because of the time singularity at $t = 0$, more work is required. The sixth step identifies, under small initial conditions, the global trajectory $\psi(s) = \langle \psi_k(s) : k \geq 1\rangle$ as a dilation of a Perron-Frobenius eigenvector (Proposition 4.11). As a consequence, the time-scale and other quantities may be uniquely identified (Proposition 4.12). The method used here may be of interest in itself. By a time-reversal argument, we show the only part of the finite-dimensional evolution, consistent with nonnegativity of the solution and small initial conditions, corresponds to motion in terms of a dominant Perron-Frobenius eigenvalue-eigenvector pair.

Last, at the end of the section, combining previous steps, the proof of Theorem 2.1 is given.

Step 1: With respect to a nonnegative solution of the integral form of the nonlinear ODEs (2.8), define for $t \geq 0$ the nonnegative functions

$$V(t) = \sum_{k\geq 1} \varphi_k(t), \quad T(t) = \sum_{k\geq 1} w(k)\varphi_k(t), \quad \text{and} \quad D(t) = \sum_{k\geq 1} k\varphi_k(t).$$

We now derive properties of the functions V, T and D, representing the scaled vertices/urns, weight, and degree/size of the system respectively. Recall $p_0, q_0 > 0$.

Proposition 4.1. *Suppose (SUB) holds. For $t \geq 0$, $T(\cdot)$ is continuous. Also, $V(t) = c + p_0 t$, $D(t) \leq \tilde{c} + (p_0 + q_0)t$, and in addition there is a constant $C_0 > 0$ such that*

$$C_0^{-1}[c + p_0 t] \ \leq \ T(t) \ \leq \ C_0[\tilde{c} + (p_0 + q_0)t].$$

Proof. We first consider $D(t)$. From the ODEs (2.8), write for $L \geq 1$ that

$$\begin{aligned} \sum_{k=1}^{L} k\varphi_k(t) \ &= \ \sum_{k=1}^{L} kc_k + \sum_{k=1}^{L} k\big(\varphi_k(t) - \varphi_k(0)\big) \qquad (4.1)\\ &= \ \sum_{k=1}^{L} kc_k + \ p_0 t + q_0 \int_0^t \sum_{k=1}^{L-1} \frac{w(k)\varphi_k(u)}{T(u)} du - q_0 L \int_0^t \frac{w(L)\varphi_L(u)}{T(u)} du. \end{aligned}$$

Hence, by (LIM), dropping the last negative term, $D(t) \leq \tilde{c} + (p_0 + q_0)t$ as desired.

Considering $V(t)$ now, write

$$\sum_{k=1}^{L} \varphi_k(t) \; = \; \sum_{k=1}^{L} c_k + \; p_0 t - q_0 \int_0^t \frac{w(L)\varphi_L(u)}{T(u)} du.$$

Since $\int_0^t \sum_{k\geq 1} w(k)\varphi_k(u)/T(u)du = t < \infty$, the last integral above vanishes as $L \uparrow \infty$. Hence, by (LIM), $V(t) = c + p_0 t$.

Also, the lower and upper bounds on $T(t)$ follow from the argument as given in Step 3 of the proof of Theorem 2.3 in Section 3. The constant C_0 can be taken as $C_0 = \max\{[(1/2)\inf_{1\leq k\leq \hat{L}} w(k)]^{-1}, \mathcal{W}\}$ where $\hat{L}$ is the smallest integer satisfying $\tilde{c}/(\hat{L}+1) \leq c/2$ and $2/(\hat{L}+1) \leq p_0/2$.

To show T is continuous, write

$$T(t) - T(s) \; = \; \sum_{k=1}^{L} w(k)\big(\varphi_k(t) - \varphi_k(s)\big) + \sum_{k>L} w(k)\varphi_k(t) - \sum_{k>L} w(k)\varphi_k(s).$$

The last two terms, for large L, are small by the inequality $\sum_{k\geq L} w(k)\varphi_k(u) \leq [\sup_{j>L} w(j)/j]D(u)$ and (SUB). Now, as $\{\varphi_k\}$ satisfies the integral form of the ODEs (2.8), they are continuous functions. Hence, one sees T is also continuous. $\square$

We now analyze more carefully the time scale $t = t(s)$ and associated system $\{\psi_k(s) = \varphi_k(t(s))\}$ mentioned above the statement of Theorem 2.1. Recall

$$\dot{t}(s) \; = \; T(t) \quad \text{and} \quad t(0) \; = \; 1.$$

Since T is continuous by Proposition 4.1, a solution $t = t(s)$ exists. Also, given the bounds on T in Proposition 4.1, by comparison estimates, we have for $s \geq 0$ that

$$\begin{aligned} & t(0)e^{C_0^{-1}p_0 s} + \frac{c}{p_0}\big[e^{C_0^{-1}p_0 s} - 1\big] \qquad (4.2)\\ & \quad \leq \; t(s) \; \leq \; t(0)e^{C_0(p_0+q_0)s} + \frac{\tilde{c}}{(p_0+q_0)}\big[e^{C_0(p_0+q_0)s} - 1\big],\\ & t(0)e^{-C_0(p_0+q_0)s} + \frac{\tilde{c}}{(p_0+q_0)}\big[e^{-C_0(p_0+q_0)s} - 1\big]\\ & \quad \leq \; t(-s) \; \leq \; t(0)e^{-C_0^{-1}p_0 s} + \frac{c}{p_0}\big[e^{-C_0^{-1}p_0 s} - 1\big]. \qquad (4.3)\end{aligned}$$

As $T(t) > 0$, for $t > 0$, we have $t(s)$ is a strictly increasing, invertible function of s. Then, under small initial configurations $c_k \equiv 0$, as $s \downarrow -\infty$, we have $t(s) \downarrow 0$. Under large initial conditions $c > 0$, there is an $-\infty < s_0 < 0$ where $t(s_0) = 0$.

The system $\psi_k(s) = \varphi_k(t(s))$ for $k \geq 1$ obeys the integral form of ODEs (2.9), with boundary conditions, under small initial configurations, $\lim_{s\downarrow -\infty} \psi_k(s) = 0$ and, under large initial configurations, $\psi_k(s_0) = c_k$, for $k \geq 1$. Also, given $\varphi_k(\cdot) \geq 0$, of course, in the corresponding time-ranges $\psi_k(\cdot) \geq 0$ for $k \geq 1$.

Step 2: In terms of $\Psi = \langle \psi_k : k \geq 1\rangle$,

$$\dot{\Psi} \; = \; A\Psi \qquad (4.4)$$

where

$$A = \begin{pmatrix} (p_0 - q_0)w(1) & p_0 w(2) & p_0 w(3) & \cdots \\ q_0 w(1) & -q_0 w(2) & 0 & \cdots \\ 0 & q_0 w(2) & -q_0 w(3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

It will be convenient to write $A = B + K$ where

$$B = \begin{pmatrix} -q_0 w(1) & 0 & 0 & \cdots \\ q_0 w(1) & -q_0 w(2) & 0 & \cdots \\ 0 & q_0 w(2) & -q_0 w(3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$K = \begin{pmatrix} p_0 w(1) & p_0 w(2) & p_0 w(3) & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

We remark, it will be useful in later analysis, that B is the transpose of a 'generator' matrix.

For a vector $x = \langle x_k : k \geq 1 \rangle$ where $x_k \in \mathbb{C}$, define the norm $\|x\| = \sum_{k\geq 1} k|x_k|$ and the Banach lattice (cf. Section VI.1b in [22])

$$\Omega = \big\{ x = \langle x_k : k \geq 1 \rangle : \|x\| < \infty \big\}.$$

Let ℓ_c be the space of compactly supported vectors, that is vectors with finitely many nonzero entries, and note $\ell_c \subset \Omega$. The operators A, B and K are well-defined on ℓ_c. Moreover, $A\ell_c \subset \ell_c$, $B\ell_c \subset \ell_c$ and $K\ell_c \subset \ell_c$, and hence A, B, K are densely defined on Ω.

We observe that $K : \Omega \to \Omega$ is a bounded operator: As $w(k) \leq \mathcal{W}k$ for $k \geq 1$ (SUB), the bound $\|K\| \leq p_0\mathcal{W}$ may be computed. Also, since K is a bounded rank 1 operator on Ω, it follows K is compact.

Now, we note the ODEs associated to B, $\dot{\zeta} = B\zeta$, fall into the framework of the 'Kolmogorov' differential equations considered in [27]. Indeed, given $\sup_k w(k)/k \leq \mathcal{W}$, in the notation of [27], with $\alpha_{k,k} = -q_0 w(k+1)$, $\alpha_{k+1,k} = q_0 w(k+1)$ for $k \geq 0$, $\alpha_{j,k} = 0$ otherwise, $\alpha^\diamond = \sup_k \sum_{j=0}^\infty \alpha_{j,k} = 0$, $c_0 = 2q_0\mathcal{W}$, $c_1 = 2q_0\mathcal{W}$, $\epsilon = 1$ and $\omega = c_1 \vee (\alpha^\diamond + c_0) = c_0$, one inspects $\sum_{j=1}^\infty j\alpha_{j,k} = q_0 w(k+1)$, and the ODEs associated to B satisfy Assumptions 1, 2 in [27]. We remark the full statement of (SUB) is not used in this verification or in the proof of the following proposition, only that $\sup_k w(k)/k < \infty$.

Proposition 4.2. *Both A and B generate strongly continuous semigroups P_t, P_t^B : $\Omega \to \Omega$ with bounds $\|P_t\| \leq 2e^{(2q_0\mathcal{W}+2\|K\|)t}$ and $\|P_t^B\| \leq 2e^{2q_0\mathcal{W}t}$ for $t \geq 0$ respectively.*

Proof. By Theorem 2 in [27], there is a strongly continuous semigroup $P_t^B : \Omega \to \Omega$, generated by B restricted to domain $D(B) \cap \{x : x_k \in \mathbb{R}\}$ where $D(B) := \{x \in \Omega : \sum_{k\geq 1} w(k)|x_k| < \infty, Bx \in \Omega\}$, with bound $\|P_t^B\| \leq e^{\omega t}$. Extending to complex $x \in D(B)$, where operators act linearly on the real and imaginary parts of x, we have $\|P_t^B\| \leq 2e^{\omega t}$. Since $\sup_k w(k)/k < \infty$, we have $D(B) = \{x \in \Omega : Bx \in \Omega\}$.

Moreover, by the perturbation Theorem III.1.3 in [22], as K is bounded, we have $A = B + K$ with domain $D(A) = D(B)$ generates a strongly continuous semigroup $P_t : \Omega \to \Omega$ with bound $\|P_t\| \le e^{(\omega+\|K\|)t}$. □

Step 3: Recall that a strongly continuous semigroup $V_t : \Omega \to \Omega$ is positive if $(V_t x)_k \ge 0$ for $k \ge 1$ when $x \in \Omega \cap \{x : x_k \in \mathbb{R}\}$ and $x_k \ge 0$ for $k \ge 1$ (cf. Section VI.1b in [22]).

Below, in Proposition 4.3, we show the semigroups generated by B and A are positive. In passing, we remark in fact P_t, although not P_t^B, is irreducible, that is $[(\lambda I - A)^{-1}x]_k > 0$ for $k \ge 1$ when $x_k \ge 0$ for all $k \ge 1$ but $x \ne 0$ (cf. Section VI.1b in [22]). Indeed, from the ODEs (4.4) and a calculation left to the reader, $(P_s x)_k = \psi_k(s) > 0$ for $k \ge 1$ and $s > 0$. Then, $(\lambda I - A)^{-1}x = \int_0^\infty e^{-\lambda s} P_s x ds$ is composed of positive entries. We will not need this stronger result in the following.

Proposition 4.3. *The semigroups P_t and P_t^B are both positive.*

Proof. We show P_t is positive; the same argument also proves P_t^B is positve. Since $\ell_c \subset D(A)$, we can calculate $P_t x$ for $x \in \ell_c$ such that $x_k \ge 0$ for $k \ge 1$ by the equation $(d/dt)P_t x = AP_t x$, in other words ODEs (4.4), and initial condition $P_0 x = x$ (cf. Lemma II.1.3 in [22]). As $q_0 > 0$, all off-diagonal entries in A are nonnegative. Hence, inspection of these ODEs reveals that $P_t x \ge 0$. Now, for $x \in \Omega$ such that $x_k \ge 0$ for $k \ge 1$, take $x^{(n)} \in \ell_c$ so that $x_k^{(n)} \ge 0$ for $k \ge 1$ and $x^{(n)} \to x$ in Ω. Since, for fixed $t \ge 0$, P_t is bounded (cf. Proposition 4.2), $P_t x^{(n)} \to P_t x$ in Ω. Hence, $P_t x \ge 0$. □

Step 4: The growth rate $w_0(E)$ of a semigroup P_t^E with generator E is $w_0(E) = \lim_{t\uparrow\infty} t^{-1} \log \|P_t^E\|$ (cf. Proposition IV.2.2 in [22]). Also, the essential growth rate $w_{\text{ess}}(E)$ of P_t^E is $w_{\text{ess}}(E) = \lim_{t\uparrow\infty} t^{-1} \log \|P_t^E\|_{\text{ess}}$ where $\|P_t^E\|_{\text{ess}} = \inf\{\|P_t^E - M\| : M \text{ compact}\}$ (cf. Proposition IV.2.10 in [22]). In particular, inputting $M \equiv 0$, we obtain

$$w_{\text{ess}}(E) \ \le \ w_0(E).$$

Proposition 4.4. *We have that $w_0(B) \le 0$.*

Proof. Again, since the ODEs associated with B satisfy Assumptions 1,2 in [27], by Theorem 4 in [27], we infer that

$$w_0(B) \ \le \ \alpha^\diamond \vee \limsup_{k\to\infty} k^{-1} \sum_{j=1}^\infty j\alpha_{j,k} \ = \ 0,$$

since $\alpha^\diamond = 0$ and by (SUB) $\limsup_{k\uparrow\infty} k^{-1} \sum_{j=1}^\infty j\alpha_{j,k} = \lim_{k\uparrow\infty} q_0 \frac{w(k+1)}{k} = 0$. □

We now show for all small $\varepsilon > 0$ that $e^{-\varepsilon t} P_t$ is a quasicompact semigroup, that is the essential growth rate $w_{\text{ess}}(A - \varepsilon I) < 0$. This is one characterization of being 'quasicompact' (cf. Proposition V.3.5 in [22]). Such semigroups have nice representations which we will leverage later on.

Proposition 4.5. *For all small $\varepsilon > 0$, the semigroup $e^{-\varepsilon t} P_t$ is quasicompact.*

Proof. We will show that $e^{-\varepsilon t} P_t^B$, the semigroup generated by $B - \varepsilon I$, is quasicompact. Then, by the perturbation result Proposition V.3.6 in [22], as K is a compact operator, $e^{-\varepsilon t} P_t$ the semigroup generated by $A - \varepsilon I = B + K - \varepsilon I$ is also quasicompact.

As stated in Proposition V.3.5 in [22], for a strongly continuous semigroup, quasicompactness is equivalent to the essential growth rate of the semigroup being strictly negative. We will apply this characterization to $B - \varepsilon I$. Since $w_0(B) \le 0$ by Proposition 4.4, we have $w_{\rm ess}(B - \varepsilon I) \le w_0(B - \varepsilon I) = w_0(B) - \varepsilon < 0$. □

Now, by the quasicompact semigroup representation Theorem V.3.7 in [22] applied to $e^{-\varepsilon t}P_t$, for $\varepsilon > 0$, there are only a finite number m of spectral values z in the right half-plane, if any, of $A - \varepsilon$, and each of these is a pole of the resolvent $R(\cdot, A - \varepsilon)$ with finite algebraic multiplicity. Moreover, when $m \ge 1$, we may write for $t \ge 0$ that

$$e^{-\varepsilon t}P_t \;=\; \sum_{r=1}^{m} U_r(t) + R(t). \tag{4.5}$$

Here, with respect to the rth pole λ_r with multiplicity k_r and spectral projection Q_r (cf. Proposition IV.1.16 in [22]),

$$U_r(t) \;=\; e^{\lambda_r t}\sum_{j=0}^{k_r-1}\frac{t^j}{j!}(A - (\varepsilon + \lambda_r)I)^j Q_r.$$

Also, Theorem V.3.7 in [22] states the remainder satisfies $\|R(t)\| \le Me^{-\beta t}$ for some $\beta > 0$ and $M \ge 1$.

In effect, $e^{-\varepsilon t}P_t$ acts as a finite-dimensional operator on Range(Q_r) and leaves it invariant for $1 \le r \le m$. In particular, $e^{-\varepsilon t}P_t$ and $\{Q_r\}$ commute and

$$e^{-\varepsilon t}P_tQ_r = U_r(t) \;\text{ and }\; R_t = e^{-\varepsilon t}P_t\Big[I - \sum_{r=1}^{m} Q_r\Big]. \tag{4.6}$$

Step 5: Let now $\sigma(E)$ be the spectrum of a generator E on Ω. The largest real part of the spectrum is denoted $s(E) = \sup\{\mathrm{Re}(\lambda) : \lambda \in \sigma(E)\}$.

To make use of the representation of $e^{-\varepsilon t}P_t$ above, we now examine the spectrum of A in a right half-plane. A goal in the next propositions is to show that $s(A)$ is positive and a simple eigenvalue. Also, we derive the form of its eigenvector.

Proposition 4.6. *The generator A has only one real eigenvalue in the strict right half-plane $\{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$. This eigenvalue corresponds to a one-dimensional eigenspace with an eigenvector with all positive entries. As a consequence, $s(A) > 0$.*

Proof. We solve $Ax = \lambda x$ for $\lambda > 0$. We have

$$\begin{aligned}\lambda x_1 &= (p_0 - q_0)x_1 + p_0\sum_{k\ge 2} w(k)x_k\\ \lambda x_k &= q_0\big\{w(k-1)x_{k-1} - w(k)x_k\big\} \quad \text{for } k \ge 2.\end{aligned}$$

This gives, for $k \ge 2$,

$$x_k \;=\; x_1\prod_{r=2}^{k}\frac{q_0w(r-1)}{\lambda + q_0w(r)}, \tag{4.7}$$

the same equations for $a_k(p_0, q_0, \lambda)$ (cf. (2.4)).

In particular, by (SUB), a calculation shows that $x \in \Omega$, and

$$\sum_{k\ge 2} w(k)x_k \;=\; x_1w(1)\sum_{k\ge 2}\prod_{r=2}^{k}\frac{q_0w(r)}{\lambda + q_0w(r)}$$

converges for $\lambda > 0$. Hence, plugging into the equation involving x_1 above,

$$\lambda \;=\; p_0 - q_0 + p_0 w(1) \sum_{k\geq 2} \prod_{r=2}^{k} \frac{q_0 w(r)}{\lambda + q_0 w(r)}. \tag{4.8}$$

Since the right-side of the equation (4.8) strictly decreases as λ grows and also diverges to infinity as $\lambda \downarrow 0$, we conclude there is exactly one $\lambda > 0$ which satisfies (4.8). This λ is the desired unique real eigenvalue, with positive eigenvector x when $x_1 > 0$. □

Proposition 4.7. *For $0 \leq \varepsilon < s(A)$, $s(A - \varepsilon I) > 0$ is the only real eigenvalue of $A - \varepsilon I$ in the strict right-half plane $\{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$. All other eigenvalues λ of $A - \varepsilon I$, if they exist, satisfy $\mathrm{Re}(\lambda) < s(A - \varepsilon I)$.*

Proof. First, for $\varepsilon > 0$, as $e^{-\varepsilon t} P_t$ is quasicompact (Proposition 4.5), as noted above, there are only a finite number of spectral values of $A - \varepsilon I$ in the right half-plane $\{z \in \mathbb{C} : \mathrm{Re}(z) \geq 0\}$, and these are all eigenvalues. In particular, there are only a finite number of spectral values/eigenvalues of A in the half-plane $\{z \in \mathbb{C} : \mathrm{Re}(z) \geq \varepsilon\}$.

Then, as $s(A) > 0$ (Proposition 4.6), and by positivity of P_t (Proposition 4.3) and the 'Perron-Frobenius' type Theorem VI.1.10 in [22], $s(A)$ is an eigenvalue of A and, by Proposition 4.6, the only real one in the strict right-half plane. Moreover, with $\varepsilon = s(A)/2$, as there are only a finite number of eigenvalues z of A with real part $\mathrm{Re}(z) \geq s(A)/2$, by another 'Perron-Frobenius' type Theorem VI.1.12(i) in [22], the boundary spectrum of A must be a singleton. Hence, any other eigenvalue z of A satisfies $\mathrm{Re}(z) < s(A)$.

Then, for all $0 \leq \varepsilon < s(A)$, the value $s(A - \varepsilon I) = s(A) - \varepsilon$ is the only real eigenvalue of $A - \varepsilon I$ in the strict right half-plane, and all other eigenvalues have real part strictly less than $s(A - \varepsilon I)$. □

Consider now the dual space

$$\Omega' \;=\; \big\{ z : \text{ There exists } C \text{ such that } |z_k| \leq Ck \text{ for all } k \geq 1 \big\}$$

and $\|z\|_{\Omega'}$ is the smallest such constant C. It will be helpful now to find an eigenvector of

$$A^* \;=\; \begin{pmatrix} (p_0 - q_0)w(1) & q_0 w(1) & 0 & 0 & \cdots \\ p_0 w(2) & -q_0 w(2) & q_0 w(2) & 0 & \cdots \\ p_0 w(3) & 0 & -q_0 w(3) & q_0 w(3) & \cdots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}$$

with positive entries.

Proposition 4.8. *There exists an eigenvector $x^* \in \Omega'$ of A^*, with all entries positive, corresponding to a real eigenvalue $\lambda^* > 0$. Moreover, λ^* can be taken as $\lambda^* = s^*$, where we recall that s^* solves $1 = F_{p_0,q_0}(s^*)$ (cf. (2.6)).*

Proof. For a possible eigenpair x^*, λ^*, we obtain equations

$$\begin{aligned} x_1^* &= \frac{q_0 w(1) x_2^*}{\lambda^* + q_0 w(1)} + \frac{p_0 w(1) x_1^*}{\lambda^* + q_0 w(1)} \\ x_k^* &= \frac{q_0 w(k) x_{k+1}^*}{\lambda^* + q_0 w(k)} + \frac{p_0 w(k) x_1^*}{\lambda^* + q_0 w(k)} \quad \text{for } k \geq 2. \end{aligned} \tag{4.9}$$

Note, by (SUB), the sum

$$\sum_{k\geq 2}\frac{p_0w(k)}{\lambda^*+q_0w(k)}\prod_{r=1}^{k-1}\frac{q_0w(r)}{\lambda^*+q_0w(r)}$$

converges for each $\lambda^*>0$. Also, consider the equation

$$1 \;=\; \sum_{k\geq 2}\frac{p_0w(k)}{\lambda^*+q_0w(k)}\prod_{r=1}^{k-1}\frac{q_0w(r)}{\lambda^*+q_0w(r)}+\frac{p_0w(1)}{\lambda^*+q_0w(1)}, \tag{4.10}$$

which is the same as $1 = F_{p_0,q_0}(\lambda^*)$ and identifies, as concluded in (2.6), that $\lambda^* = s^*$.

Iterating (4.9), we may solve

$$\begin{aligned} x_1^* \;&=\; \lim_{N\uparrow\infty} x^*_{N+1}\prod_{r=1}^{(n)}\frac{q_0w(r)}{\lambda^*+q_0w(r)} \\ &\qquad +x_1^*\sum_{k\geq 2}\frac{p_0w(k)}{\lambda^*+q_0w(k)}\prod_{r=1}^{k-1}\frac{q_0w(r)}{\lambda^*+q_0w(r)}+x_1^*\frac{p_0w(1)}{\lambda^*+q_0w(1)}. \end{aligned}$$

Since $\lambda^*=s^*$, noting (4.10), necessarily $\lim_{N\uparrow\infty} x^*_{N+1}\prod_{r=1}^{(n)}\frac{q_0w(r)}{\lambda^*+q_0w(r)}=0$.

In this case, for $j\geq 2$, with convention $\prod_{r=j+1}^{j}\cdot = 1$,

$$\begin{aligned} x_j^* \;&=\; x_1^*\sum_{k\geq j}\frac{p_0w(k+1)}{\lambda^*+q_0w(k+1)}\prod_{r=j}^{k}\frac{q_0w(r)}{\lambda^*+q_0w(r)}+x_1^*\frac{p_0w(j)}{\lambda^*+q_0w(j)} \\ &=\; \frac{x_1^*w(j)}{\lambda^*+q_0w(j)}\Big[\sum_{k\geq j}\frac{q_0p_0w(k+1)}{\lambda^*+q_0w(k+1)}\prod_{r=j+1}^{k}\frac{q_0w(r)}{\lambda^*+q_0w(r)}+p_0\Big]. \end{aligned}$$

Again, by (SUB), one sees that $|x_j^*|\leq Cj$ for a uniform constant C for $j\geq 1$. In particular, the eigenvector $x^*\in\Omega'$ and if $x_1^*>0$, then $x_j^*>0$ for $j\geq 2$. $\square$

Proposition 4.9. *The eigenvalue $s(A)$ of A is simple and moreover $\lambda^*=s(A)$. Also, $x^*\perp w$ for any (generalized) eigenvector w of A other than the one with eigenvalue $s(A)$.*

Proof. Consider the eigenvector x^* with eigenvalue λ^* in Proposition 4.8 consisting of all positive entries when say $x_1^*=1$. Then, with respect to the positive eigenvector x with eigenvalue $s(A)$ of A in Proposition 4.7, we note $\langle x^*, Ax\rangle = s(A)\langle x^*,x\rangle = \lambda^*\langle x^*,x\rangle$. Since $\langle x^*,x\rangle>0$, we have $\lambda^*=s(A)$.

Moreover, suppose there exists a generalized eigenvector w where $(A-s(A)I)w = zx$ for some $z\neq 0$. Then, $\lambda^*\langle x^*,w\rangle = \langle x^*,Aw\rangle = \langle x^*, s(A)w\rangle + z\langle x^*,x\rangle$. Since $\lambda^*=s(A)$ and $\langle x^*,x\rangle>0$, we must have $z=0$ which is a contradiction. Hence, $s(A)$ is a simple eigenvalue of A.

Finally, for any eigenvector w of A with eigenvalue $\lambda_w\neq s(A)$, $s(A)\langle x^*,w\rangle = \langle x^*,Aw\rangle = \lambda_w\langle x^*,w\rangle$. Since $\lambda_w\neq s(A)$, we have $x^*\perp w$. If w' is a generalized eigenvector corresponding to eigenvalue λ_w, we have $(A-\lambda_w)^k w' = zw$ for some power k and constant z. Then, $(s(A)-\lambda_w)^k\langle x^*,w'\rangle = \langle (A^*-\lambda_w)^k x^*, w'\rangle = z\langle x^*,w\rangle = 0$. Again, as $\lambda_w\neq s(A)$, we have $x^*\perp w'$. $\square$

At this point, we conclude the preliminary work with respect to starting from large configurations, and one may now skip to the proof of Theorem 2.1, at the end of the section, for these configurations. More development is needed however for small initial configurations.

Step 6: Consider now, under small initial configurations, the global trajectory $\psi(s) = \langle \psi_k(s) : k \geq 1\rangle$ satisfying the ODEs (4.4) such that $\psi(0) = \phi(1)$ and $\lim_{s\downarrow -\infty} \psi(s) = 0$. To characterize $\psi(s)$, we will need the following estimate.

Lemma 4.10. *Under small initial configurations, for $s < 0$,*

$$Y(s) \ := \ \|\psi(s)\| \ \leq \ (p_0 + q_0)e^{C_0^{-1}p_0 s}$$

where C_0 is the constant in Lemma 4.1.

Proof. From (4.3), applied to small initial configurations ($c = 0$) and negative times, noting $t(0) = 1$, we have $t(s) \leq e^{C_0^{-1}p_0 s}$. Also, note by Proposition 4.1 that $D(u) \leq (p_0 + q_0)u$ for $u \geq 0$. Then, $Y(s) = \sum_{k\geq 1} k\psi_k(s) = \sum_{k\geq 1} k\varphi_k(t(s)) = D(t(s)) \leq (p_0 + q_0)t(s) \leq (p_0 + q_0)e^{C_0^{-1}p_0 s}$. $\square$

Proposition 4.11. *With respect to small initial configurations, we identify $\psi(s) = e^{s(A)s}\psi(0)$, for all $s \in \mathbb{R}$, where $\psi(0)$ is an eigenvector of A, with eigenvalue $s(A)$, satisfying relations* (4.7)*.*

Proof. The argument proceeds in stages.

Step I. Let $0 < \varepsilon < C_0^{-1}p_0/2$ where C_0 is the constant in Proposition 4.1. Recall the quasicompact representation of $e^{-\varepsilon t}P_t$ in (4.5). For $u \in \mathbb{R}$, define $\xi(u)$ by the equation

$$e^{-\varepsilon u}\psi(u) \ = \ e^{-\varepsilon u}\sum_{r=1}^m Q_r\psi(u) + \xi(u). \tag{4.11}$$

Then, for $t \geq 0$ and $s \in \mathbb{R}$, on the one hand,

$$e^{-\varepsilon(s+t)}\psi(s+t) \ = \ e^{-\varepsilon(s+t)}\sum_{r=1}^m Q_r\psi(s+t) + \xi(s+t). \tag{4.12}$$

On the other hand, as $R_t = e^{-\varepsilon t}P_t[I - \sum_{r=1}^m Q_r]$ (cf. (4.6)),

$$\begin{aligned} e^{-\varepsilon(s+t)}\psi(s+t) \ &= \ e^{-\varepsilon t}P_t\big(e^{-\varepsilon s}\psi(s)\big) \\ &= \ e^{-\varepsilon t}P_t\big[\sum_{r=1}^m Q_r\big]\big(e^{-\varepsilon s}\psi(s)\big) + R_t\big(e^{-\varepsilon s}\psi(s)\big). \end{aligned}$$

Since $R_t\big(e^{-\varepsilon s}\psi(s)\big) = R_t\xi(s)$, and $e^{-\varepsilon t}P_t$ and $\{Q_r\}$ commute,

$$e^{-\varepsilon(s+t)}\psi(s+t) \ = \ e^{-\varepsilon(s+t)}\sum_{r=1}^m Q_r\psi(s+t) + R_t\xi(s). \tag{4.13}$$

Hence, combining (4.12) and (4.13), we have $\xi(s+t) = R_t\xi(s)$, and with the bound on R_t after (4.5),

$$\|\xi(s+t)\| \ \leq \ \|R(t)\|\|\xi(s)\| \ \leq \ Me^{-\beta t}\|\xi(s)\|.$$

Step II. We now argue that $\xi(u) = 0$ for all $u \in \mathbb{R}$. First, for $s < 0$, from Lemma 4.10 and $\varepsilon < C_0^{-1}p_0/2$,

$$\|e^{-\varepsilon s}\psi(s)\| \ = \ e^{-\varepsilon s}Y(s) \ \leq \ (p_0 + q_0)e^{(C_0^{-1}p_0 - \varepsilon)s} \ \leq \ (p_0 + q_0)e^{(C_0^{-1}p_0/2)s}. \tag{4.14}$$

Second, from its finite-dimensional form, the operator $e^{-\varepsilon t}P_t|_{\mathrm{Range}(Q_r)}$ is invertible for $t \geq 0$. Denote the inverse on the range of Q_r as

$$e^{\varepsilon t}P_{-t}|_{\mathrm{Range}(Q_r)} \;=\; e^{-\lambda_r t}\sum_{j=0}^{k_r-1}((-t)^j/j!)(A-(\varepsilon+\lambda_r)I)^j Q_r \;=:\; U_r(-t)$$

where U_r is extended to $\mathbb{R}_-$. Then, for $s<0$, we have

$$e^{\varepsilon s}P_{-s}Q_r\big(e^{-\varepsilon s}\psi(s)\big) \;=\; Q_r\big(P_{-s}\psi(s)\big) \;=\; Q_r\psi(0).$$

Hence, after inverting,

$$e^{-\varepsilon s}\sum_{r=1}^m Q_r\psi(s) \;=\; \sum_{r=1}^m U_r(s)\psi(0). \tag{4.15}$$

Third, with respect to a constant $C = C(\{\lambda_r\},\{k_r\},\varphi(1))$, for $s<0$, from (4.11) and (4.15), and bound (4.14) and $\lambda_r \geq 0$ for $1\leq r\leq m$,

$$\begin{aligned}\|\xi(s)\| \;&\leq\; \|e^{-\varepsilon s}\psi(s)\| + \Big\|\sum_{r=1}^m U_r(s)\psi(0)\Big\| \\ &\leq\; (p_0+q_0) + C|s|^{\max_{1\leq r\leq m}k_r-1}\end{aligned}$$

As a consequence, for fixed $u = s+t$ where $s<0$ and $t>0$, as $t\uparrow\infty$, we have

$$\|\xi(u)\| \;=\; \|\xi(s+t)\| \;\leq\; Me^{-\beta t}\big[(p_0+q_0)+C|u-t|^{\max k_r-1}\big] \;\to\; 0.$$

Therefore, in equation (4.11), for all $u\in\mathbb{R}$,

$$e^{-\varepsilon s}\psi(u) = \sum_{r=1}^m U_r(u)\psi(0). \tag{4.16}$$

Step III. Recall $\psi(\cdot)$ is assumed nonnegative. We now show that $\psi(u) = e^{s(A)u}\psi(0)$ for all $u\in\mathbb{R}$ where $\lambda = s(A)$ is the simple eigenvalue of A with largest real part. We will also conclude $\psi(0)$ is an eigenvector corresponding to $s(A)$.

Indeed, the eigenvalue λ_r with largest real part is of form $s(A)-\varepsilon>0$ with a corresponding eigenvector x with all positive entries (cf. Proposition 4.6). Recall x^* is the positive eigenvector of A^* with eigenvalue $\lambda^* = s(A)$ and that all (generalized) eigenvectors w of $A-\varepsilon I$ corresponding to $\lambda_r \neq s(A)-\varepsilon$ are orthogonal to x^* (cf. Propositions 4.7, 4.8, 4.9).

Let $\{\lambda_r : r\in I_\alpha\}$ be those eigenvalues with the same real part $\mathrm{Re}(\lambda_r)=\alpha$, and let I_α be the associated index set. For $0\leq j\leq \max_{1\leq r\leq m}k_r-1$, consider the sum

$$\begin{aligned}A(\alpha,j,s) \;&:=\; \sum_{r\in I_\alpha} e^{\lambda_r s}(s^j/j!)(A-(\varepsilon+\lambda_r)I)^j Q_r\psi(0) \\ &=\; e^{s\alpha}\frac{s^j}{j!}\sum_{r\in I_\alpha} e^{is\mathrm{Im}(\lambda_r)}((A-(\varepsilon+\lambda_r)I)^j Q_r\psi(0).\end{aligned}$$

There are a finite number of nontrivial sums indexed by α, j. Let $\bar\alpha$ be the minimum real part of the eigenvalues $\{\lambda_r\}$ and suppose $\bar\alpha\neq s(A)$, the largest real part. Let $\hat{j}$ be the maximum of k_r-1 among the eigenvalues λ_r with real part $\bar\alpha$.

Suppose $A(\bar\alpha,\hat{j},s)\neq 0$ for some $s\in\mathbb{R}$. We claim that we can find integers $\{k_{r,\ell} : r\in I_\alpha\}$ and $n_\ell\geq 1$ where $\lim_{\ell\uparrow\infty}n_\ell=\infty$ and $\max_{r\in I_\alpha}|n_\ell\mathrm{Im}(\lambda_r)/(2\pi) - k_{r,\ell}| \leq n_\ell^{-1/|I_\alpha|}$ for $\ell\geq 1$: Indeed, if $\{\mathrm{Im}(\lambda_r)/(2\pi) : r\in I_\alpha\}$ are all rational, this is the case; if one of $\{\mathrm{Im}(\lambda_r)/(2\pi) : r \in I_\alpha\}$ is irrational, then Dirichlet's simultaneous Diophantine approximation theorem, Corollary II.1B in [40], implies the claim.

Then, at times $u_\ell = s - n_\ell$ for large $\ell \geq 1$, the sum $(e^{u_\ell \bar{\alpha}} u_\ell^{\hat{j}}/\hat{j}!)^{-1} A(\bar{\alpha}, \hat{j}, u_\ell)$ well approximates $(e^{s\bar{\alpha}} s^{\hat{j}} \hat{j}!)^{-1} A(\bar{\alpha}, \hat{j}, s)$ in Ω, and the corresponding amplitude $|e^{u_\ell \bar{\alpha}} u_\ell^{\hat{j}}/\hat{j}!|$ dominates those of all the other sums $A(\alpha, j, u_\ell)$ for $(\alpha, j) \neq (\bar{\alpha}, \hat{j})$ as $\ell \uparrow \infty$.

Suppose $|u_\ell|$ is large. Then,

$$e^{\varepsilon u_\ell} \psi_k(u_\ell) \;=\; \Big[\sum_{r=1}^{m} U_r(u_\ell)\psi(0)\Big]_k \;\sim\; A(\bar{\alpha}, \hat{j}, u_\ell)_k$$

with respect to components k of $A(\bar{\alpha}, \hat{j}, u_\ell)$ which are nonzero. Given $e^{\varepsilon u_\ell}\psi(u_\ell)$ is real, the components of $A(\bar{\alpha}, \hat{j}, u_\ell)$ are real. Since $x^* \perp w$ for any generalized eigenvector w of λ_r, $r \in I_{\bar{\alpha}}$, there must be a component of $A(\bar{\alpha}, \hat{j}, u_\ell)$ which is strictly negative. This contradicts the nonnegativity of $e^{\varepsilon u_\ell}\psi(u_\ell)$. Therefore, $A(\bar{\alpha}, \hat{j}, s) = 0$ for $s \in \mathbb{R}$.

Similarly, considering the remaining finite number of sums $A(\alpha, j, u)$, strictly ordered according to their growth as $u \downarrow -\infty$, we conclude $A(\alpha, j, s) = 0$ when $\alpha < s(A)$ for $s \in \mathbb{R}$.

In particular, for $u \in \mathbb{R}$, we have $\sum_{r:\lambda_r \neq s(A)} U_r(u)\psi(0) = 0$. Noting (4.16), $\psi(s) = e^{s(A)u} Q\psi(0)$, where Q is projection onto the eigenvector x of $\lambda = s(A)$ given by relations (4.7). Finally, $\psi(0)$ is also a corresponding eigenvector since $\psi(0) = Q\psi(0)$. □

We now identify, under small initial configurations, the 'time-change' $t = t(s)$', and other quantities, given in the beginning of this Section.

Proposition 4.12. *With respect to small initial configurations, we have $t(u) = e^{s(A)u}$ for $u \in \mathbb{R}$ and $T(t) = s(A)t$ for $t \geq 0$. Also, $\psi_k(0) = a_k(p_0, q_0, s^*)$ for $k \geq 1$.*

Proof. Since $\psi(u) = e^{s(A)u}\psi(0)$ from Proposition 4.11, and $\psi(u) = \varphi(t(u))$, we have from Lemma 4.1 that $p_0 t(u) = \sum_{k\geq 1} \varphi_k(t(u)) = e^{s(A)u} \sum_{k\geq 1} \psi_k(0)$.

Since $t(0) = 1$, we have $\sum_{k\geq 1} \psi_k(0) = p_0$. This shows $t(u) = e^{s(A)u}$ for $u \in \mathbb{R}$. Next, as $s(A)t(u) = \dot{t}(u) = T(t(u))$, and $t = t(u)$ is onto $\mathbb{R}$, we have $T(t) = s(A)t$ for $t \geq 0$.

Finally, we recall that $s(A) = s^*$, by Propositions 4.8 and 4.9. Also, by Proposition 4.11, $\psi(0)$ is an eigenvector of A with eigenvalue $s(A) = s^*$ satisfying (4.7), the same relations $\{a_k(p_0, q_0, s^*)\}$ satisfy in (2.4). Therefore, there exists a constant $z \neq 0$ where $\psi_k(0) = z a_k(p_0, q_0, s^*)$ for $k \geq 1$. However, $p_0 = \sum_{k\geq 1} \psi_k(0) = z \sum_{k\geq 1} a_k(p_0, q_0, s^*) = z p_0$, noting (2.7). Hence, $z = 1$ to finish the proof. □

Proof of Theorem 2.1. First, consider large initial configurations and recall the time s_0 defined after (4.3) so that $t(s_0) = 0$, and $\psi(s_0)_k = c_k$ for $k \geq 1$. Then, $\psi(s + s_0) = P_s \psi(s_0)$ for $s \geq 0$ is the unique solution to the integral form of ODEs (2.9) (cf. Proposition II.6.4 in [22]). Now, $\sum_{k\geq 1} \psi_k(s + s_0) = \sum_{k\geq 1} \varphi(t(s + s_0)) = p_0 t(s + s_0) + c$. In particular, $t(\cdot)$ is uniquely specified in terms of $\{\psi_k(\cdot)\}$ and c. Hence, $\{\varphi_k(u) = \psi_k(t^{-1}(u)) : u \geq 0\}$ is uniquely determined.

Moreover, for $\varepsilon > 0$ small and $s > 0$, as $e^{-\varepsilon s} P_s$ satisfies representation (4.5), and the dominant eigenvalue $s(A) - \varepsilon > 0$ is simple (Proposition 4.9), we have

$e^{-s(A)s}\psi(s) = e^{-(s(A)-\varepsilon)s}e^{-\varepsilon s}P_s\psi(0)$ converges in Ω to an eigenvector v with eigenvalue $s(A)$ of A as $s \uparrow \infty$ (cf. discussion before Corollary V.3.3 in [22]). Write,

$$p_0 + \frac{c}{t(s)} \;=\; \frac{1}{t(s)}\sum_{k\geq 1}\varphi_k(t(s)) \;=\; \frac{e^{s(A)s}}{t(s)} \cdot e^{-s(A)s}\sum_{k\geq 1}\psi_k(s).$$

Taking $s \uparrow \infty$, as $t(s) \uparrow \infty$ from (4.2), we conclude $p_0 = z\sum_{k\geq 1} v_k$ where $z = \lim_{s\uparrow\infty} e^{s(A)s}/t(s)$, which necessarily converges. By the eigenvector formula (4.7) which $\{a_k(p_0, q_0, s^*)\}$ satisfies, fact $s(A) = s^*$ (Propositions 4.8 and 4.9), and equality $\sum_{k\geq 1} a_k(p_0, q_0, s^*) = p_0$ (cf. (2.7)), we identify $zv_k = a_k(p_0, q_0, s^*)$ for $k \geq 1$. Hence, for $k \geq 1$,

$$\lim_{s\uparrow\infty}\frac{\varphi_k(s)}{s} \;=\; \lim_{s\uparrow\infty}\frac{e^{s(A)t^{-1}(s)}}{s}[e^{-s(A)t^{-1}(s)}\psi_k(t^{-1}(s))] \;\to\; zv_k \;=\; a_k(p_0, q_0, s^*).$$

Now, consider small initial configurations. By Propositions 4.11 and 4.12, $\psi(s) = \langle t(s)a_k(p_0, q_0, s^*) : k \geq 1\rangle$, and $t(s) = e^{s(A)s}$ are identified. Therefore, for $u \geq 0$ and $k \geq 1$, $\varphi_k(u) = \psi_k(t^{-1}(u)) = a_k(p_0, q_0, s^*)u$ is uniquely determined. $\square$

5. Appendix: $d_k(j+1)$ in the graph model

As mentioned, formation of loops need to be considered. For $k \geq 3$,

$$d_k(j+1) = \begin{cases} 2 & \text{with prob. } (1-p)\big[\frac{w(k-1)Z_{k-1}(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(k-1)]^2 Z_{k-1}(j)}{[S(j)]^2} \\ 1 & \text{with prob. } p\frac{w(k-1)Z_{k-1}(j)}{S(j)} \\ & \quad +(1-p)\frac{[w(k-2)]^2 Z_{k-2}(j)}{[S(j)]^2} \\ & \quad +2(1-p)\frac{w(k-1)Z_{k-1}(j)}{S(j)}\big[1-\frac{w(k-1)Z_{k-1}(j)}{S(j)}-\frac{w(k)Z_k(j)}{S(j)}\big] \\ 0 & \text{with prob. } p\big[1-\frac{w(k-1)Z_{k-1}(j)}{S(j)}-\frac{w(k)Z_k(j)}{S(j)}\big] \\ & \quad +(1-p)\frac{[w(k-1)]^2 Z_{k-1}(j)}{[S(j)]^2} \\ & \quad +2(1-p)\frac{w(k-1)Z_{k-1}(j)}{S(j)}\frac{w(k)Z_k(j)}{S(j)} \\ & \quad +(1-p)\big[1-\frac{w(k-1)Z_{k-1}(j)}{S(j)}-\frac{w(k)Z_k(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(k-2)]^2 Z_{k-2}(j)}{[S(j)]^2} \\ -1 & \text{with prob. } p\frac{w(k)Z_k(j)}{S(j)} \\ & \quad +(1-p)\frac{[w(k)]^2 Z_k(j)}{[S(j)]^2} \\ & \quad +2(1-p)\frac{w(k)Z_k(j)}{S(j)}\big[1-\frac{w(k-1)Z_{k-1}(j)}{S(j)}-\frac{w(k)Z_k(j)}{S(j)}\big] \\ -2 & \text{with prob. } (1-p)\big[\frac{w(k)Z_k(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(k)]^2 Z_k(j)}{[S(j)]^2}. \end{cases}$$

$$d_1(j+1) = \begin{cases} 1 & \text{with prob. } p\big[1-\frac{w(1)Z_1(j)}{S(j)}\big] \\ 0 & \text{with prob. } p\frac{w(1)Z_1(j)}{S(j)} + (1-p)\big[1-\frac{w(1)Z_1(j)}{S(j)}\big]^2 \\ -1 & \text{with prob. } 2(1-p)\frac{w(1)Z_1(j)}{S(j)}\big[1-\frac{w(1)Z_1(j)}{S(j)}\big] \\ & \quad +(1-p)\frac{[w(1)]^2 Z_1(j)}{[S(j)]^2} \\ -2 & \text{with prob. } (1-p)\big[\frac{w(1)Z_1(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(1)]^2 Z_1(j)}{[S(j)]^2}. \end{cases}$$

$$d_2(j+1) = \begin{cases} 2 & \text{with prob. } (1-p)\big[\frac{w(1)Z_1(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(1)]^2 Z_1(j)}{[S(j)]^2} \\ 1 & \text{with prob. } p\frac{w(1)Z_1(j)}{S(j)} \\ & \quad +2(1-p)\frac{w(1)Z_1(j)}{S(j)}\big[1-\frac{w(1)Z_1(j)}{S(j)}-\frac{w(2)Z_2(j)}{S(j)}\big] \\ 0 & \text{with prob. } p\big[1-\frac{w(1)Z_1(j)}{S(j)}-\frac{w(2)Z_2(j)}{S(j)}\big] \\ & \quad +(1-p)\frac{[w(1)]^2 Z_1(j)}{[S(j)]^2} \\ & \quad +2(1-p)\frac{w(1)Z_1(j)}{S(j)}\frac{w(2)Z_2(j)}{S(j)} \\ & \quad +(1-p)\big[1-\frac{w(1)Z_1(j)}{S(j)}-\frac{w(2)Z_2(j)}{S(j)}\big]^2 \\ -1 & \text{with prob. } p\frac{w(2)Z_2(j)}{S(j)} \\ & \quad +(1-p)\frac{[w(2)]^2 Z_2(j)}{[S(j)]^2} \\ & \quad +2(1-p)\frac{w(2)Z_2(j)}{S(j)}\big[1-\frac{w(1)Z_1(j)}{S(j)}-\frac{w(2)Z_2(j)}{S(j)}\big] \\ -2 & \text{with prob. } (1-p)\big[\frac{w(2)Z_2(j)}{S(j)}\big]^2 \\ & \quad -(1-p)\frac{[w(2)]^2 Z_2(j)}{[S(j)]^2}. \end{cases}$$

6. Appendix: Non-compactness of semigroups

We know from Proposition 4.2 that, for each $t \geq 0$, P_t is a bounded operator. If for some $t_0 \geq 0$, we also knew that P_{t_0} were compact, then we would conclude by the semigroup property that P_{u+t_0} is compact for $u \geq 0$, in other words that the semigroup P_t is 'eventually' compact. However, we show in the following that this is not the case for P_t and also P_t^B.

Proposition 6.1. *For each $t \geq 0$, the operators P_t and P_t^B are not compact.*

Proof. For $x \in \Omega$, let $\zeta(t;x) = P_t^B x$. From the form of B (cf. after (2.9)), we observe that

$$\begin{aligned}\sum_{k=1}^{L} k\zeta_k(t;x) &= \sum_{k=1}^{L} k\zeta_k(0;x) + q_0 \int_0^t \sum_{k=1}^{L-1} w(k)\zeta_k(s;x)ds \\ &\qquad - q_0 L w(L) \int_0^t \zeta_L(s;x)ds \end{aligned} \tag{6.1}$$

when $x \in \ell_c$ (cf. proof of Proposition 4.1).

Fix now $x \in \ell_c$ positive. Since P_t^B is positive (Proposition 4.3), we have $\zeta_k(\cdot;x) \geq 0$ for $k \geq 1$ and $\sum_{k=1}^{L} k\zeta_k(t;x) \leq \sum_{k=1}^{L} k\zeta_k(0;x) + \mathcal{W} q_0 \int_0^t \sum_{k=1}^{L} k\zeta_k(s;x)ds$. Therefore, the upper bound $\sum_{k\geq 1} k\zeta_k(t;x) \leq e^{\mathcal{W} q_0 t} \sum_{k \geq 1} k\zeta_k(0;x)$.

We now derive a lower bound. In (6.1), by the upper bound, limits of all terms as $L \uparrow \infty$ converge. In particular, by positivity, $\sum_{k\geq 1} w(k) \int_0^t \zeta_k(s;x)ds = \int_0^t \sum_{k \geq 1} w(k)\zeta_k(s;x)ds < \infty$, and so the limit $\lim_L Lw(L)\int_0^t \zeta_L(s;x)ds = 0$. Therefore, from (6.1) and positivity, we get $\sum_{k\geq 1} k\zeta_k(t;x) \geq \sum_{k\geq 1} k\zeta_k(0;x) = \|x\|$.

Then, starting from $n = 1$, let $x^{(1)} \in \ell_c$ where $x_1^{(1)} = 1$ and $x_k^{(1)} = 0$ for $k > 1$. Let also $L^{(1)}$ be an index so that $\sum_{k > L^{(1)}} k\zeta_k(t;x^{(1)}) \leq 1/2$. For $n \geq 1$, define $x^{(n+1)} \in \ell_c$ where $x_{L^{(n)}}^{(n+1)} = (L^{(n)})^{-1}$ and $x_k^{(n+1)} = 0$ for $k \neq L^{(n)}$. Define also $L^{(n+1)} > L^{(n)}$ as an index where $\sum_{k > L^{(n+1)}} k\zeta_k(t;x^{(n+1)}) \leq 1/2$.

Note that the sequence $\{x^{(n)}\}$ is bounded in Ω: $\|x^{(n)}\| = \sum_{k\geq 1} k|x_k^{(n)}| = 1$.

We now show $\|P_t^B x^{(n)} - P_t^B x^{(m)}\| \geq 1/2$ for all $1 \leq m < n$. By the form of B, there is no flow 'backwards', that is $\zeta_k(t;x^{(n)}) \equiv 0$ for $k < L^{(n-1)}$. Write

$$\begin{aligned}\sum_{k\geq 1} k|\zeta_k(t;x^{(m)}) - \zeta_k(t;x^{(n)})| &\geq \sum_{k \leq L^{(m)}} k|\zeta_k(t;x^{(m)}) - \zeta_k(t;x^{(n)})| \\ &= \sum_{k\leq L^{(m)}} k\zeta_k(t;x^{(m)}) \geq \|x^{(m)}\| - 1/2 = 1/2.\end{aligned}$$

Hence, P_t^B cannot be a compact operator for any $t \geq 0$.

Similarly, P_t cannot be compact for any $t \geq 0$: Suppose P_{t_0} is compact. Then, as P_u is bounded for each $u \geq 0$, P_{t_0+u} is compact for $u \geq 0$. Because K is compact and $B = A - K$, by say the perturbation result Theorem III.1.14(i) in [22], $P_{t_1}^B$ for some $t_1 \geq 0$ would also be compact, a contradiction. □

Acknowledgement. We would like to thank the referees for their useful and constructive comments which helped improve the exposition of the paper.

S.S. was partially supported by ARO W911NF-14-1-0179.

S.C.V. was partially supported by NSF 0807501.

References

[1] Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* **74,** 47–97.

[2] Athreya, K. B. (2007). Preferential attachment random graphs with general weight function. *Internet Math.* **4,** 401–418.

[3] Athreya, K. B., Ghosh, A. P. and Sethuraman S. (2008). Growth of preferential attachment random graphs via continuous-time branching processes. *Proc. Indian Acad. Sci. Math. Sci.* **118,** 473–494.

[4] Baladi, V. (2000). *Positive Transfer Operators and Decay of Correlations.* Adv. Ser. in Nonlinear Dynamics **16** World Scientific, Singapore.

[5] Barabási, Albert-László (2009). Scale-free networks: a decade and beyond. *Science* **325,** 412–413.

[6] Barbour, A.D., Pugliese, A. (2005). Asymptotic behavior of a metapopulation model. *Ann. Appl. Probab.* **15** 1306–1338.

[7] Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18,** 279–290.

[8] Bornholdt, S. and Schuster, H. G.(eds) (2003). *Handbook of Graphs and Networks: From the Genome to the Internet.* Wiley-VCH, Weinheim, Germany.

[9] Caldarelli, G. (2007). *Scale-Free Networks: Complex Webs in Nature and Technology.* Oxford University Press, USA.

[10] Chen, H., Yao, D.D. (2001) *Fundamentals of Queuing Networks: Performance, Asymptotics, and Optimization.* Applications of Mathematics; Stochastic Modelling and Applied Probability **46**. Springer-Verlag, New York.

[11] Choi, J. and Sethuraman, S. (2011) Large deviations for the degree structure in preferential attachment schemes. *Ann. Appl. Probab.* **23,** 722–763.

[12] Chung F., Handjani, S. and Jungreis, D. (2003). Generalizations of Pólya's urn problem. *Annals of Combinatorics* **7,** 141–153.

[13] Chung, F. and Lu, L. (2006). *Complex Graphs and Networks* vol. 107 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences, Washington, DC.

[14] Cohen R. and Havlin S. (2010). *Complex Networks: Structure Robustness and Function.* Cambridge University Press.

[15] Collevechio, A., Cotar, C. and LiCalzi, M. (2013). On a preferential attachment and generalized Pólya's urn model. *Ann. Appl. Probab.* **23,** 1219–1253.

[16] Darling, R.W.R and Norris, J.R. (2008). Differential equation approximations for Markov chains. *Probability Surveys* **5**, 37–79.

[17] Dereich, S. and Mörters, P. (2009). Random networks with sublinear preferential attachment: Degree evolutions. *Electron. J. Probab.* **14,** 1222–1267.

[18] Dorogovtsev, S. N. and Mendes, J. F. F. (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW.* Oxford University Press, USA

[19] Dorogovtsev, S.N., Mendes, J.F.F., and Samukhin, A.N. (2000). Structure of Growing Networks with Preferential Linking. *Phys. Rev. Lett.* **85,** 4633-4636.

[20] Drinea, E., Frieze, A. and Mitzenmacher, M. (2002). Balls and Bins models with feedback. *Proc. of the 11th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 308–315.

[21] Durrett, R. (2007). *Random Graph Dynamics.* Cambridge U. Press.

[22] Engle, K-J., Nagel, R. (2000). *One-Parameter Semigroups for Linear Evolution Equations* Springer-Verlag, New York.

[23] Foss, S. Konstantopoulos, T. (2004). An overview of some stochastic stability methods. *J. Oper. Research* **47** 275–303.

[24] Krapivsky, P. and Redner, S. (2001). Organization of growing random networks. *Phys. Rev. E* **63,** 066123-1 – 066123-14.

[25] Krapivsky, P.L., Redner, S., and Leyvraz, F. (2000). Connectivity of Growing Random Networks *Phys. Rev. Lett.* **85,** 4629-4632.

[26] Martcheva, M., Thieme, H. R. (2008). Infinite ODE systems modeling size-structured metapopulations, macroparasitic diseases, and prion proliferation. *Structured Population Models in Biology and Epidemiology* Lecture Notes in Math **1936** 51–113, Springer, Berlin.

[27] Martcheva, M., Thieme, H. R., Dhirasakdanon, T. (2006) Kolmogorov's differential equations and positive semigroups on first moment sequence spaces. *J. Math. Biol.* **53** 642-671.
[28] Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Math.* **1,** 226–251.
[29] Móri, T. F. (2002). On random trees. *Studia Sci. Math. Hungar.* **39,** 143–155.
[30] Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review* **45,** 167–256 (electronic).
[31] Newman, M. E. J. (2010). *Networks: An Introduction.* Oxford University Press, USA.
[32] Newman, M. E. J. and Watts, D. J. (2006). *The Structure and Dynamics of Networks.* Princeton University Press, USA.
[33] Oliveira, R. and Spencer, J. (2005). Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.* **2,** 121–163.
[34] Peköz, E.A., Röllin, A., Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23,** 1188–1218.
[35] Pemantle, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4,** 1–79.
[36] Robert, P. (2003). *Stochastic Networks and Queues.* Applications of Mathematics; Stochastic Modelling and Applied Probability **52**. Springer-Verlag, Berlin.
[37] Rath, B., Toth, B. (2009). Erdos-Renyi random graphs + forest fires = self-organized criticality. *Elec. J. Probab.* **14** 1290–1327.
[38] Ross, N. (2012). Power laws in preferential attachment graphs and Stein's method for the negative binomial distribution. Arxiv preprint arXiv 1208.1558.
[39] Rudas, A., Tóth, B. and Valkó, B. (2007). Random trees and general branching processes. *Random Struct. Algorithms* **31**, 186–202.
[40] Schmidt, W.M. (1980). *Diophantine Approximation* Lecture Notes in Mathematics **785** Springer-Verlag, Berlin.
[41] Simkin, M. V. and Roychowdhury, V. P. (2011). Re-inventing Willis. *Phys. Rep.* **502**, 1–35.
[42] Simon, H. (1955). On a class of skew distribution functions. *Biometrika* **42**, 425–440.
[43] Webb, G. (1985). *Theory of Nonlinear Age Dependent Population Dynamics.* Marcel Dekker, New York.
[44] Whitt, W. (2002). *Stochastic-Process Limits.* Springer Series in Operations Research. Springer, New York.
[45] Wimmer, M. (2006) A law of large numbers and central limit theorem for the leaves in a random graph model. MS thesis, Department of Mathematics, Iowa State University.
[46] Wormald, N. C. (1995). Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5,** 1217–1235.

Department of Mathematics, Syracuse University, Syracuse, NY 13244
e-mail: `jchoi46@syr.edu`

Department of Mathematics, University of Arizona, Tucson, AZ 85721
e-mail: `sethuram@math.arizona.edu`

Department of Mathematics, University of Arizona, Tucson, AZ 85721
e-mail: `shankar@math.arizona.edu`