technische universität
dortmund

Generalized Becker-Döring Equations Modeling the Time Evolution of a Process of Preferential Attachment with Fitness

Flavius Guiaş

Preprint 2008-06 April 2008

Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund

tu-dortmund.de/MathPreprints

GENERALIZED BECKER-DÖRING EQUATIONS MODELING THE TIME EVOLUTION OF A PROCESS OF PREFERENTIAL ATTACHMENT WITH FITNESS

FLAVIUS GUIAŞ

ABSTRACT. We introduce an infinite system of equations modeling the time evolution of the growth process of a network. The nodes are characterized by their degree $k \in \mathbb{N}$ and a fitness parameter $f \in [0, h]$. Every new node which emerges becomes a fitness f' according to a given distribution P and attaches to an existing node with fitness f and degree k at rate fA_k, where A_k are positive coefficients, growing sublinearly in k. If the parameter f takes only one value, the dynamics of this process can be described by a variant of the *Becker-Döring* equations, where the growth of the size of clusters of size k occurs only with increment 1. In contrast to the established Becker-Döring equations, the system considered here is nonconservative, since mass (i.e. links) is continuously added. Nevertheless, it has the property of linearity, which is a natural consequence of the process which is being modeled. The purpose of this paper is to construct a solution of the system based on a stochastic approximation algorithm, which allows also a numerical simulation in order to get insight into its qualitative behaviour. In particular we show analytically and numerically the property of Bose-Einstein condensation, which was observed in the literature on random graphs.

1. INTRODUCTION

The growth process of random networks has been intensively studied in the physics literature. It intends to model phenomena like the distribution of scientific citations or the growth of the world wide web. The basic model which we are interested in is described in [7]. Every new node which appears attaches to an existing one with degree k with probability proportional to A_k. By assuming a time scaling which ensures a linear growth of the total mass of the system, a system of equations for the number densities of nodes with degree $k = 1, 2 \dots$ is derived. This turns out to be convenient especially for studying the asymptotic behaviour as $t \to \infty$. However, as it will be pointed out later, if we are interested in the correct dynamics over time, the growth will be exponential and the system of equations takes a form similar to the Becker-Döring equations for pure coagulation, see [4]. These equations decribe a cluster growth process which can occur only with increment 1, i.e. the size of a cluster can grow form k to $k+1$ due to attachment of a monomer, that is a particle of size 1. The equations considered in the mentioned reference are mass-conserving and nonlinear, since the mass of monomers varies in time. Nevertheless, it is pointed out that the original equations considered by Becker and Döring in 1935 assumed a constant concentration of monomers, which made them

2000 *Mathematics Subject Classification.* 60K40, 65C35.
Key words and phrases. network growth, preferential attachement, generalized Becker-Döring equations, stochastic particle methods.

linear. The same pattern can be observed in the growth of random networks, where links are continuously added to the system. They take the form

$$\begin{aligned} \frac{d}{dt}u_1(t) &= -A_1 u_1 + \sum_{i=1}^{\infty} A_i u_i, \\ \frac{d}{dt}u_k(t) &= -A_k u_k + A_{k-1} u_{k-1}, \;\; k = 2, 3 \dots \end{aligned} \tag{1}$$

where $u_k(t)$ denotes the concentration of nodes with degree k at time t. These equations stand for the correct time evolution of the model considered in [7].

If we think at concrete situations, this approach turns out to be unrealistic. In [2] it is introduced a model of *preferential attachement with fitness*, which is rigourously analyzed in [3]. The key assumption is that the rate of growth of a node with degree k should not only be proportional to its “degree of popularity” taken as $A_k = k$, but also to a certain “attractiveness” or “fitness’,’ which the node gets at its birth according to a given distribution. In this way it can be ensured that the degree of a node which is born later, but has higher fitness, can surpass in the long run that of the nodes with larger degree which were already present before its appearance.

The setting for this problem is the following. Consider a fitness distribution P on the compact interval $[0, h]$. We may assume either that P is discrete, taking the values $f_i,\ i = 1, \infty$ with $0 \leq f_1 < f_2 < \dots f_n < \cdots < h$ with $h = \sup_i f_i$, or that it has a density $p(f)$ with respect to the Lebesgue measure such that $p(f) > 0$ on $(0, h)$. By denoting $u_k(t, f)$ the density of nodes with degree k and fitness f at time t, one obtains the following system of equations:

$$\begin{aligned} \frac{d}{dt}u_1(t,f) &= -fA_1 u_1(t,f) + p(f)\int_0^h f' \sum_{i=1}^{\infty} A_i u_i(t,f')df', \\ \frac{d}{dt}u_k(t,f) &= -fA_k u_k(t,f) + fA_{k-1}u_{k-1}(t,f), \;\; k = 2, 3 \dots \end{aligned} \tag{2}$$

In the mentioned references the analysis of this growth process is performed within the framework of random graphs, i.e. from the node perspective. The following interesting behaviour was observed in [2] and proved in [3]: namely the *Bose-Einstein condensation*, which means a “leak” of a macroscopic fraction of links at h, the right end of the fitness interval. If $A_k = k$, this happens if the fitness distribution P has an infinite support and satisfies the condition

$$I(h) := \int_0^h \frac{f}{h-f} dP(f) < 1. \tag{3}$$

According to [3], an example in the absolutely continuous case is given by the $Beta(\mu, \nu)$-distribution on $[0, 1]$ with $\nu > \mu + 1$, while for the discrete case one can take $f_j = h - h/j$ for $j \geq 1$ and $P(f_j) = j^{-\theta}/\zeta(\theta)$, where $\zeta(\theta) = \sum_{j\geq 1} j^{-\theta}$ is the Riemann zeta function. If one takes θ sufficiently large, e.g. $\theta > 3$ (a more precise estimate can be obtained numerically: $\theta > \theta^* \approx 2.48$) then condition (3) holds true. This behaviour, which is similar to the phenomenon of *gelation* known from coagulation processes, is caused by the appearance of clusters with larger and larger fitness values. If the tail of the distribution P decays sufficiently fast when approaching h, then we will have basically very few clusters with high fitness values, which concentrate most of the links of all nodes in that range. This structure speeds up their growth process and in the long run a fraction of links "escapes" out of the system. Depending on the profile of the fitness distribution one can observe the following phenomena, which, using the terminlogy from [3] (inspired from the economic growth according to the "market-fitness" of old and young companies), can be described as follows:

- *first-mover-advantage*, which shows up for flat fitness distributions and means that the links will be mostly concentrated at those fitness values which were present at the initial moment.
- *fit-get-richer*, which means that most of the links will be concentrated in nodes with high fitness values.
- *innovation-pays-off*, i.e. the Bose-Einstein condensation described before, as an extreme case of the fit-get-richer phenomenon.

All mentioned references on this topic were concerned mainly with the long-time behaviour of the network growth process and considered only discrete time steps. The puropse of this paper is to introduce the equations (2) as a model for the time-evolution of this process, which is done in Section 2. Our next aim is to approximate them by a stochastic numerical scheme in order to get insight into the qualitative behaviour of this model. This stochastic scheme is presented in Section 3. A relative compactness property of the approximating stochastic processes is shown to hold for $A_k \leq Ck$. This yields also existence results for the deterministic system of equations in the cases $A_k = o(k)$ or $A_k = Ck$ and $I(h) > 1$, that is, in absence of Bose-Einstein condensation and away from the limit case. In this latter case the solution turns out ot be unique. In addition, we give interpret condition (3) in the framework of the equations (2), showing that in this case we have indeed a loss of mass. The results of the numerical simulations are presented in Section 4.

2. A model for the time dynamics of preferential attachment with fitness

For a finite $h > 0$ denote $I = [0, h]$ and let $E = I \times \mathbb{N}$. The set $C_c(E)$ of continuous real functions with compact support in E consists therefore of functions of the type $\psi = (\psi_k(\cdot))_{k\in\mathbb{N}}$ with ψ_k continuous functions on I, such that $\psi_k(\cdot) \equiv 0$ for $k > n(\psi)$ (the maximal index k for which the component $\psi_k(\cdot)$ is nontrivial) .

Let further $\mathcal{M}_+(E)$ be the set of positive Radon measures on E, which can be identified with the space of positive linear forms on $C_c(E)$. For an arbitrary measure μ and a measurable function f we use the standard notation $\langle \mu, f\rangle := \int f d\mu$. Since E is separable, following [1] we define on $\mathcal{M}_+(E)$ the *vague topology* in which the convergence is defined by

$$\mu_n \xrightarrow{v} \mu \Leftrightarrow \langle \mu_n, f\rangle \to \langle \mu, f\rangle,\ \forall f \in C_c(E).$$

By considering an appropriate countable and dense (with respect to uniform convergence) set of functions $\phi^i \in C_c(E)$, $i \in \mathbb{N}$, the vague topology is induced by the metric

$$\rho(\mu, \nu) := \sum_{i=1}^{\infty} c_i(|\langle \mu, \phi^i\rangle - \langle \nu, \phi^i\rangle| \wedge 1).$$

In [1], p.240 it is considered $c_i = 2^{-i}$, but here we will take for technical reasons c_i with the property that $S_\rho = \sum c_i \cdot (n(\phi^i)\|\phi^i\|_\infty \vee 1) < \infty$. The symbols $\wedge, \vee$ stand as usual for min and max respectively and $\|\cdot\|_\infty$ for the sup-norm.

We will define next the notion of a *weak (or measure valued) solution* of the equation system modeling preferential attachment with fitness on the time interval $[0,T]$, with $T>0$ fixed. It is more relevant to have as components the link densities in nodes of degree k and fitness f given by $v_k(t,f) := ku_k(t,f)$. We then have:

$$
\begin{aligned}
\frac{d}{dt}v_1(t,f) &= -fA_1v_1(t,f) + p(f)\int_0^h f' \sum_{i=1}^{\infty} \frac{A_i}{i} v_i(t,f')df', \\
\frac{d}{dt}v_k(t,f) &= -fA_kv_k(t,f) + fA_{k-1}\frac{k}{k-1}v_{k-1}(t,f),\ \ k=2,3\ldots
\end{aligned} \tag{4}
$$

By multiplicating the integral form of the above differential equations with an arbitrary test function $\psi \in C_c(E)$ and integrating over E we note that we can group terms in the form

$$
fA_kv_k(t,f)\cdot(-\psi_k(f)+\frac{k+1}{k}\psi_{k+1}(f)) = f\frac{A_k}{k}v_k(t,f)\cdot(-k\psi_k(f)+(k+1)\psi_{k+1}(f)).
$$

Assume now that

$$
A_k \le Ck \tag{5}
$$

for all $k\in\mathbb{N}$, where C is a positive constant.

We say then that the time-dependent family of *finite* Radon measures $\mu(t) \in \mathcal{M}_+(E)$ is a solution of (4) on the time interval $[0,T]$, if for any $\psi \in C_c(E)$ we have

$$
\begin{aligned}
\langle \mu(t),\psi\rangle &= \langle \mu(0),\psi\rangle + \int_0^t \{ \sum_{k=1}^{\infty} \frac{A_k}{k}\langle \mu_k(s), -k\psi_k \mathrm{id}_I + (k+1)\psi_{k+1}\mathrm{id}_I\rangle + \\
&\qquad \langle P,\psi_1\rangle \cdot \sum_{k=1}^{\infty}\frac{A_k}{k}\langle \mu_k(s),\mathrm{id}_I\rangle \} ds,
\end{aligned} \tag{6}
$$

for all $t\in[0,T]$, where $\mu_k(A) := \mu(\{k\}\times A)$ for any Borel measurable subset A of I and $\mathrm{id}_I(f)=f$ denotes the identity function on I. Note that the first series has in fact only a finite number of terms due to the compactness of the support of ψ, while the second one can be majorized by $C\cdot h\cdot \mu(s)(E) < \infty$, since $\mu(\cdot)$ are assumed to be finite measures on E.

In a condensed form the equation can be stated as

$$
\langle \mu(t),\psi\rangle = \langle \mu(0),\psi\rangle + \int_0^t \langle \mu(s),\hat{\psi}\rangle ds, \tag{7}
$$

for all $\psi \in C_c(E)$, where $\hat{\psi}(k,f) = fA_kk^{-1}[\langle P,\psi_1\rangle - k\psi_k(f) + (k+1)\psi_{k+1}(f)]$. Note that $\hat{\psi} \in C_b(E)$, i.e. $\hat{\psi}$ is continous and bounded, but in general $\hat{\psi} \notin C_c(E)$.

3. A stochastic scheme approximating the deterministic equations

We will construct next a Markov jump process on the space of finitely supported measures on E which describes the dynamics of preferential attachment with fitness. Let $\mu = \sum_{(k,f)} m(k,f)\delta_{k,f}$ be such a measure which describes the current state of the process, where the pair (k,f) runs over the (finite) size/fitness combinations where mass is concentrated. The possible transitions $\mu \to \mu'$ of the process are defined by

$$
\mu \to \mu + \frac{1}{N}\delta_{1,f'} - \frac{k}{N}\delta_{k,f} + \frac{k+1}{N}\delta_{k+1,f} \ \text{ at rate } \ N\cdot p(f')df'\cdot f\cdot \frac{A_k}{k}m(k,f). \tag{8}
$$

The interpretation is that a new node of size 1 and fitness f' is generated according to the probability density p and attaches instantaneously to the existing nodes of type (k, f) with probability proportional to $fA_k k^{-1} m(k, f)$. The factor N gives the "resolution" of the model, N^{-1} stands for the smallest amount of density the process can "see", that is the mass of a node of size 1. By attachment of a new node, the number of nodes of size k decreases with 1 and that of nodes of size $k+1$ increases with 1. Since the values of the weights of our discrete measure μ stand for the *link density*, i.e. *size* $\times$ *number density*, the modifications after a transition step take the form given above.

By Dynkin's formula we have the martingale representation of the dynamics in terms of the infinitesimal generator Λ:

$$\Psi(\mu(t)) = \Psi(\mu(0)) + \int_0^t (\Lambda\Psi)(\mu(s))ds + M_\Psi(t)$$

for all bounded continuous functions Ψ defined on $\mathcal{M}_+(E)$, where $M_\Psi(\cdot)$ is a martingale with respect to the σ-algebra generated by the process.

By taking $\Psi(\mu) = \langle \mu, \psi \rangle$ for test functions $\psi \in C_c(E)$ we have

$$\langle \mu(t), \psi \rangle = \langle \mu(0), \psi \rangle + \int_0^t (\Lambda\Psi)(\mu(s))ds + M_\psi(t). \tag{9}$$

The infinitesimal generator can be computed in this case by

$$\begin{aligned}
(\Lambda\Psi)(\mu(t)) &= \lambda(\mu) \int_{\mathcal{M}_+(E)} (\Psi(\mu') - \Psi(\mu)) r(\mu, d\mu') \\
&= \int_0^h \sum_{(k,f)} [\psi_1(f') - k\psi_k(f) + (k+1)\psi_{k+1}(f)] \cdot f \cdot \frac{A_k}{k} \cdot m(k, f) \cdot p(f')df' \\
&= \langle P, \psi_1 \rangle \cdot \sum_k \frac{A_k}{k} \langle \mu_k(t), \mathrm{id}_I \rangle + \sum_k \frac{A_k}{k} \langle \mu_k(t), -k\psi_k \mathrm{id}_I + (k+1)\psi_{k+1}\mathrm{id}_I \rangle \\
&= \langle \mu(t), \hat{\psi} \rangle,
\end{aligned} \tag{10}$$

where $r(\mu, d\mu') = \lambda^{-1}(\mu) \cdot N \cdot \sum_{(k,f)} fA_k k^{-1} m(k, f) p(f')df'$ is the transition kernel corresponding to (8) and $\lambda(\mu) = N \cdot \sum_{(k,f)} fA_k k^{-1} m(k, f)$ is the total rate. In order to ensure that the jumps do not accumulate, the total rate function has to be bounded. We achieve this by fixing a maximal time $T > 0$, considering a sufficiently large constant a and introducing the stopping time $\tau_a := \inf\{t : \mu(t)(E) \geq a\} \wedge T$. We will consider therefore the stopped process $\mu(t \wedge \tau_a)$ which belongs to the space $\mathcal{M}_+^a(E) := \{\mu \in \mathcal{M}_+(E) : \mu(E) \leq a\}$.

For simplicity of the notation we have suppressed the dependence on N and $\tau_a = \tau_a(N)$, but the dynamics above define a family of stochastic processes indexed by N, with values in $D([0, \infty), \mathcal{M}_+^a(E))$, that is the set of right continuous with left limits, $\mathcal{M}_+^a(E)$-valued functions defined on $[0, \infty)$. The appropriate convergence in this space is defined by the *Skorokhod topology*, which is weaker than the topology of uniform convergence and takes in account also the distance in time between the jumps. For details see [5], p. 117.

We point out that $\mathcal{M}_+^a(E)$ is separable, being a subspace of the separable metric space $(\mathcal{M}_+(E), \rho)$. Furthermore, it is also complete: cf. [1], Lemma 30.3, if $\mu_n \xrightarrow{v} \mu$ for finite measures μ_n, we then have $\mu(E) \leq \liminf \mu_n(E)$, which is $\leq a$ for $\mu_n \in \mathcal{M}_+^a(E)$.

We will show next a relative compactness property of this family, i.e. the probability distributions $\mathcal{L}(\mu_{N_k}(\cdot))$ on the space $D([0,\infty), \mathcal{M}_+^a(E))$ of a subsequence $\mu_{N_k}(\cdot)$ converge weakly to the law of a limit process as $N \to \infty$. Conditions under which the weak limits solve equation (6) will be also given.

We first need some preliminaries. Define an upper bound $\bar{M}(t)$ for the total mass process $\mu(t)(E)$ by

$$\bar{M}_N(t) \to \bar{M}_N(t) + \frac{2}{N} \text{ at rate } NCh \cdot \bar{M}_N(t). \tag{11}$$

Note that after a step in (8) the mass $\mu_N(t)(E)$ increases by $2/N$. We define a higher total rate for this transition step by considering only the maximal value for the fitness, using the bound (5) of the growth coefficients and majorizing the sum over the terms $m(k,f)$ by the total mass $\bar{M}_N$ itself. The process $\bar{M}_N(t)$ delivers therefore an upper bound for the increase of the total mass $\mu_N(t)(E)$. By standard results of convergence towards solution of ODE's ([5], Chapter 11) we have that $E[\bar{M}_N(t)] \to \bar{M}_0 \cdot \exp(2Cht)$ as $N \to \infty$, provided the convergence of the initial condition to $\bar{M}_0$.

Proposition 1. *Assume that the sequence $\mu_N(0)(E)$ of initial masses has a limit for $N \to \infty$. Then for every $T \geq 0$ we have $\inf_N \tau_a(N) \to T$ a.s. for $a \to \infty$.*

Proof. Using the majorant property of $\bar{M}_N$ and the Markov inequality we have:

$$P(\tau_a(N) = T) = P(\mu_N(T)(E) \leq a) \geq P(\bar{M}_N(T) \leq a) \geq 1 - \frac{E[\bar{M}_N(T)]}{a}$$
$$= 1 - \frac{\bar{M}_0 \cdot \exp(2ChT) + \varepsilon(N)}{a} \geq 1 - \eta_a$$

for all N (since $\varepsilon(N) \to 0$), with $\eta_a \to 0$ as $a \to \infty$. This proves the statement of the proposition.

□

Lemma 1. *For a given $T > 0$ there exists a family of random variables $\gamma_N(\theta)$, $\theta \in [0,1]$, with $\sup_N E[\gamma(\theta)] \to 0$ as $\theta \to 0$ such that*

$$E[\rho(\mu_N(t+s), \mu_N(t)) \,|\mathcal{F}_t^N] \leq E[\gamma_N(\theta) \,|\mathcal{F}_t^N]$$

for all $0 \leq t \leq T$ and $0 \leq s \leq \theta$, where $\mathcal{F}_t^N$ is the σ-algebra generated by the process $\mu_N(t)$.

Proof. For the sake of simplicity we suppress the index N. We then have:

$$E[\rho(\mu(t+s), \mu(t)) \,|\mathcal{F}_t] = E[\sum_i c_i(|\langle \mu(t+s), \phi^i\rangle - \langle \mu(t), \phi^i\rangle| \wedge 1) \,|\mathcal{F}_t]$$
$$\leq E[\sum_i c_i(\left|\int_t^{t+s} (\Lambda\Phi^i)(\mu(\tau))d\tau\right| \wedge 1 + |M_{\phi^i}(t+s) - M_{\phi^i}(t)| \wedge 1) \,|\mathcal{F}_t].$$

Taking into account the expression (10) of the infinitesimal generator and the stopping property of the process we estimate further:

$$E[\rho(\mu(t+s), \mu(t)) \,|\mathcal{F}_t] \leq$$
$$\leq E[s \cdot 3Cha \sum_i c_i \cdot n(\phi^i)\|\phi^i\|_\infty + \sum_i c_i(|M_{\phi^i}(t+s) - M_{\phi^i}(t)| \wedge 1) \,|\mathcal{F}_t]$$
$$\leq \quad E[s \cdot 3Cha \cdot S_\rho + \sum_i c_i(|M_{\phi^i}(t+s) - M_{\phi^i}(t)| \wedge 1) \,|\mathcal{F}_t] \leq E[\gamma(\theta) \,|\mathcal{F}_t]$$

for

$$\gamma(\theta) = \theta \cdot 3Cha \cdot S_\rho + \sup_{0\le s\le\theta} \sum_i c_i |M_{\phi^i}(t+s) - M_{\phi^i}(t)| \wedge 1 =: \gamma_1(\theta) + \gamma_2(\theta).$$

We only have to analyze the convergence of the second term. We have:

$$\begin{aligned} E[\gamma_2(\theta)] &\le E[\sum_i c_i \sup_{0\le s\le\theta} |M_{\phi^i}(t+s) - M_{\phi^i}(t)|] \\ &= \sum_i c_i E[\sup_{0\le s\le\theta} |M_{\phi^i}(t+s) - M_{\phi^i}(t)|] \\ &\le \sum_i 2c_i \left(E[|M_{\phi^i}(t+\theta) - M_{\phi^i}(t)|^2]\right)^{1/2}. \end{aligned} \tag{12}$$

where in the last step we used Doob's inequality for the martingale $M_{\phi^i}(t+s) - M_{\phi^i}(t)$ with respect to the variable s. The interchange of the expectation with the infinite summation is possible due to [6], Theorem 9.2..

Using the martingale property we compute further:

$$\begin{aligned} &E[|M_{\phi^i}(t+\theta) - M_{\phi^i}(t)|^2 \,|\mathcal{F}_t] = E[M^2_{\phi^i}(t+\theta) - 2M_{\phi^i}(t+\theta)M_{\phi^i}(t) + M^2_{\phi^i}(t) \,|\mathcal{F}_t] \\ &= E[M^2_{\phi^i}(t+\theta) \,|\mathcal{F}_t] - 2M_{\phi^i}(t)E[M_{\phi^i}(t+\theta) \,|\mathcal{F}_t] + M^2_{\phi^i}(t) \\ &= E[M^2_{\phi^i}(t+\theta) \,|\mathcal{F}_t] - M^2_{\phi^i}(t), \end{aligned}$$

and therefore:

$$\begin{aligned} &E[|M_{\phi^i}(t+\theta) - M_{\phi^i}(t)|^2] = E[E[|M_{\phi^i}(t+\theta) - M_{\phi^i}(t)|^2 \,|\mathcal{F}_t]] \\ &= E[E[M^2_{\phi^i}(t+\theta) \,|\mathcal{F}_t] - M^2_{\phi^i}(t)] = E[M^2_{\phi^i}(t+\theta)] - E[M^2_{\phi^i}(t)]. \end{aligned} \tag{13}$$

By standard techniques we have:

$$\begin{aligned} E[M_{\phi^i}(t)^2] &= E[\int_0^t [\Lambda(\Phi^i)^2 - 2\Phi^i\Lambda\Phi^i](\mu(s))ds] \\ &= E[\int_0^t \lambda(\mu(s)) \int (\Phi^i(\mu') - \Phi^i(\mu))^2 r(\mu(s), d\mu')ds] \\ &= \frac{1}{N} E[\int_0^t \int_0^h \sum_{(k,l)} (\phi^i_1(f') - k\phi^i_k(f) + (k+1)\phi^i_{k+1}(f))^2 \cdot \\ &\quad \cdot f \cdot \frac{A_k}{k} \cdot m(k,f)p(f')df'ds], \end{aligned} \tag{14}$$

while a similar formula can be computed for the second moment at $t+\theta$. This yields:

$$\begin{aligned} &E[M^2_{\phi^i}(t+\theta)] - E[M^2_{\phi^i}(t)] = \\ &= \frac{1}{N} E[\int_t^{t+\theta} \int_0^h \sum_{(k,l)} (\phi^i_1(f') - k\phi^i_k(f) + (k+1)\phi^i_{k+1}(f))^2 \cdot \\ &\cdot f \cdot \frac{A_k}{k} \cdot m(k,f)p(f')df'ds] \le \frac{1}{N}\theta \cdot 3Cha \cdot n^2(\phi^i) \cdot \|\phi^i\|^2_\infty. \end{aligned}$$

Inserting in (12) we obtain:

$$E[\gamma_2(\theta)] \le \sum_i 2c_i \sqrt{\frac{3Cha\cdot\theta}{N}} \cdot n(\phi^i) \cdot \|\phi^i\|_\infty \le 2\sqrt{\frac{3Cha\cdot\theta}{N}} \cdot S_\rho \to 0$$

as $\theta \to 0$ uniformly in N.

□

We can state now the convergence result of the constructed stochastic scheme.

Proposition 2. *Assume that condition (5) holds and that the initial distributions* $\mu_N(0)$ *of the family of stochastic processes given by the transitions (8) converge in distribution to a finite measure* $\tilde{\mu}_0$ *on* E. *Then the family* $(\mu_N(\cdot \wedge \tau_a))_{N=1}^{\infty}$ *is relatively compact in* $D([0,\infty), \mathcal{M}_+^a(E))$. *Moreover, any weak limit* $\tilde{\mu}(\cdot)$ *has continuous paths a.s..*

Proof. According to [1], Corollary 31.3., the set $\mathcal{M}_+^a(E)$ is vaguely compact. Together with Lemma 1, this enables us to use Theorem 8.6. from [5], p.137 in order to infer the relative compactness property. The transitions (8) imply that the distance between two consecutive jumps can be estimated by

$$\rho(\mu, \mu') = \frac{1}{N} \sum_i c_i(|\langle \phi^i(f') - k\phi^i(f) + (k+1)\phi^i(f)| \wedge 1) \le \frac{S_\rho}{N}.$$

We then have

$$\int_0^\infty e^{-t}[\sup_{s \le t} \rho(\mu(s), \mu(s-)) \wedge 1]dt \le \frac{S_\rho}{N} \int_0^\infty e^{-t} dt \to 0$$

as $N \to \infty$. By using Theorem 10.2. from [5], p.148, we obtain that the paths of the limit process, denoted by $\tilde{\mu}$, are a.s. continuous. □

We will analyze next the behaviour of the limit processes.

Theorem 1. *Assume that* $A_k = o(k)$ *and convergence of initial conditions. Then any weak limit* $\tilde{\mu}(\cdot)$ *of the family of stochastic processes given by (8) solves equation (6) on* $[0, \tau_a]$, *where* $P(\tau_a = T) \to 1$ *as* $a \to \infty$.

Proof. Denote the weakly convergent subsequence again with $\mu_N(\cdot)$. This means that we have $E[F(\mu_N)] \to E[F(\tilde{\mu})]$ for all bounded and continuous functions F on $D([0,\infty), \mathcal{M}_+^a(E))$. For $\psi \in C_c(E)$ and $t \in [0,T]$ consider in particular

$$F_\psi(\mu) = \langle \mu(t), \psi \rangle - \langle \mu(0), \psi \rangle - \int_0^t \langle \mu(s), \hat{\psi} \rangle ds,$$

where $\hat{\psi}(k,f) = fA_k k^{-1}[\langle P, \psi_1 \rangle - k\psi_k(f) + (k+1)\psi_{k+1}(f)]$.

Taking into account that $\mu \in \mathcal{M}_+^a(E)$, it is easy to see that F_ψ is bounded. According to [5], Proposition 5.2. on p.118, convergence in the Skorokhod space implies pointwise convergence in all continuity points of the limit function. Since our limit process has continuous paths a.s., we will replace on $D([0,\infty), \mathcal{M}_+^a(E))$ the Skorokhod topology with the topology associated to the pointwise convergence and have to show that F_ψ is continuous with respect to it. The continuity of the part corresponding to the first two terms in the expression of F_ψ is clear, since this means nothing else than to say that if $\mu_n(\cdot) \to \mu(\cdot)$ pointwise in $D([0,\infty), \mathcal{M}_+^a(E))$, then $\mu_n(t) \xrightarrow{v} \mu(t)$ for all t, i.e. $\langle \mu_n(t) - \mu(t), \psi \rangle \to 0$ for all t.

In a similar way we obtain pointwise convergence $\langle \mu_n(s), \hat{\psi}_2 \rangle \to \langle \mu(s), \hat{\psi}_2 \rangle$ for $\hat{\psi}_2(k,f) = fA_k k^{-1}[-k\psi_k(f) + (k+1)\psi_{k+1}(f)]$, since ψ has compact support. For $\hat{\psi}_1(k,f) = fA_k k^{-1}\langle P, \psi_1 \rangle$, due to $A_k = o(k)$ we have $\hat{\psi}_1 \in C_0(E)$, which is the space of continuous functions on E which vanish at infinity (defined as the closure of $C_c(E)$ in the sup-norm). Since $\mu_n(s)(E) \le a$ and $\mu_n(s) \xrightarrow{v} \mu(s)$, according to [1], Theorem 30.6., we have $\langle \mu_n(s), \phi\rangle \to \langle \mu(s), \phi\rangle$ for all $\phi \in C_0(E)$. The continuity of the integral term is obtained then by using the dominated convergence theorem.

Returning to our processes, we obtain therefore that $E[|F_\psi(\mu_N)|] \to E[|F_\psi(\tilde{\mu})|]$ or, equivalently, $E[|M_\psi^N(t)|] \to E[|F_\psi(\tilde{\mu})|]$, with M_ψ^N the martingale given in (9). By using the same steps as in the estimates for γ_2 in the proof of Lemma 1, we obtain that $E[\sup_{t\le T}|M_\psi^N(t)|] \to 0$ as $N \to \infty$, which implies that $F_\psi(\tilde{\mu})$=0 a.s., for every $\psi \in C_c(E)$.

In order to finish the proof, it remains to show that we can commute in the above statement, that is, that we have $F_\psi(\tilde{\mu})$=0 for every $\psi \in C_c(E)$ a.s..

Note first that we can state that $F_{\phi^i}(\tilde{\mu})$=0 for every $i \in \mathbb{N}$ a.s., for the functions $(\phi^i)_{i\in\mathbb{N}}$, which are dense in $C_c(E)$ and were used for the construction of the vague distance on $\mathcal{M}_+(E)$. Let now $\psi \in C_c(E)$ be arbitrary. Due to the construction of the functions ϕ^i (see [1], p. 240), we can find a sequence (ϕ^{i_n}) with $n(\phi^{i_n}) \le n_{max}$ and $\|\phi^{i_n} - \psi\|_\infty \to 0$. The index n_{max} depends only on the function ψ. We have now to show that $F_{\phi^{i_n}}(\tilde{\mu}) \to F_\psi(\tilde{\mu})$. Using the approximation property of (ϕ^{i_n}) we obtain immediately $\langle \tilde{\mu}(\cdot), \phi^{i_n}\rangle \to \langle \tilde{\mu}(\cdot), \psi\rangle$.

We have next

$$\sum_k A_k k^{-1}\langle \tilde{\mu}_k(s), \mathrm{id}_I\rangle \cdot |\langle P, \psi_1\rangle - \langle P, \phi_1^{i_n}\rangle| \le \|\phi^{i_n} - \psi\|_\infty \cdot Ch \cdot \tilde{\mu}(s)(E) \to 0$$

and

$$\begin{aligned}&\sum_k A_k \langle \tilde{\mu}_k(s), |\psi_k(\cdot) - \phi_k^{i_n}(\cdot)|\rangle \le Cn_{max} \cdot \|\phi^{i_n} - \psi\|_\infty \cdot \tilde{\mu}(s)(E)\\ &\le Cn_{max} \cdot \|\phi^{i_n} - \psi\|_\infty \cdot a \to 0.\end{aligned}$$

Applying now the dominated convergence theorem we obtain the desired statement.

Using Proposition 1 for estimating the stopping time τ_a we obtain the conclusion of the theorem. □

The above proof can not be extended in a straightforward way to the case where we have $A_k = O(k)$. In order to obtain convergence of the terms $\langle \mu_n, \hat{\psi}_1\rangle$ we would need a stronger property, namely a convergence of the type $\langle \mu_n(t), \psi\rangle \to \langle \mu(t), \psi\rangle$ for all $\psi \in C_b(E)$ (weak convergence). According to [1], Theorem 3.8., this is implied by $\mu_n(t) \xrightarrow{v} \mu(t)$ and $\mu_n(t)(E) \to \mu(t)(E)$. That is, we need in addition a mass conservation property which is not automatically ensured, since by the vague convergence in the limit we may have a loss of mass, see [1], p.221.

In order to deal with this problem we will consider next only the case $A_k = Ck$. Without loss of generality we may assume $C = 1$.

Let us make some formal considerations about equations (4) in this special case. We will modify slightly the interval where the fitness values are defined. Namely, if

$$I(h) = \int_0^h \frac{f}{h-f}\,dP(f) \ge 1,$$

we obtain then by monotonicity arguments the existence of a $h^* \ge h$ with $I(h^*) = 1$. Note that if P has a finite support this is always the case. We extend then the probability density to be 0 on $(h, h^*]$. If $I(h) < 1$ then we take $h^* = h$.

Consider $h' \ge h$. We make the observation that for $f \in [0, h')$ we have

$$\frac{f}{h'-f} = \sum_{k=1}^\infty \left(\frac{f}{h'}\right)^k.$$

Denote the total mass corresponding to the fitness f at time t by

$$m(t,f) = \sum_{k=1}^{\infty} v_k(t,f)$$

and

$$M_k(t) := \int_0^{h'} \left(\frac{f}{h'}\right)^k m(t,f)df.$$

By summing up equations (4) for $A_k = k$ we obtain then

$$\frac{dm(t,f)}{dt} = f \cdot m(t,f) + h'p(f)M_1(t). \tag{15}$$

The equations for M_k are obtained by multiplying with $(f/h')^k$ and integrating with respect to f:

$$\frac{dM_k(t)}{dt} = h'[M_{k+1}(t) + M_1(t)\int_0^{h'} \left(\frac{f}{h'}\right)^k p(f)df]. \tag{16}$$

By summing up all these equations we obtain for

$$M(t) = \sum_{k=1}^{\infty} M_k(t) \tag{17}$$

the following equation:

$$\begin{aligned} \frac{dM(t)}{dt} &= h'[M(t) - M_1(t) + M_1(t)\int_0^{h'} \sum_{k=1}^{\infty} \left(\frac{f}{h'}\right)^k p(f)df] \\ &= h'[M(t) - M_1(t) + M_1(t)\int_0^{h'} \frac{f}{h'-f} p(f)df] \\ &= h'[M(t) + (I(h') - 1)M_1(t)]. \end{aligned} \tag{18}$$

If $I(h^*) = 1$, then we have

$$\frac{dM(t)}{dt} = h^* M(t), \tag{19}$$

while if $I(h^*) < 1$ then we can registrate the "loss of mass" due to the Bose-Einstein condensation, since the series of normalized moments $M(t)$ grows slower that the expected exponential growth.

Note that if $I(h) > 1$ then we have a unique solution. Equation (18) holds in fact for all $h' \geq h$, therefore also for $h' = h$. Inserting (19) in (18) for $h' = h$ we obtain

$$h^* M(t) = h[M(t) + (I(h) - 1)M_1(t)]$$

and thus

$$M_1(t) = \frac{(h^* - h)M(t)}{I(h) - 1}.$$

Note that the quantities $m(t,f), v_k(t,f)$ depend in fact only on h, since $p \equiv 0$ on $(h, h']$ implies that we do not have mass concentrated at these fitness values. $h'M_1(t) = \int_0^{h'} fm(t,f)df$ is therefore independent on $h' > h$, depending only on h. Inserting now $hM_1(t)$ in the first equation of (4), we obtain $v_1(t,f)$ in a unique way. We insert then succesively the unique value of v_k in the equation for v_{k+1} and obtain that (4) has a unique solution.

The next theorem states the properties of the limit process in the case $A_k = Ck$.

Theorem 2. *Assume that $A_k = Ck$, convergence of initial conditions and $I(h) > 1$. Then any weak limit $\tilde{\mu}(\cdot)$ of the family of stochastic processes given by (8) solves equation (6) on $[0, \tau_a]$, where $P(\tau_a = T) \to 1$ as $a \to \infty$.*

Proof. The lines of the proof are similar to Theorem 1. According to the previous observations, in order to obtain the weak convergence property, we need to show convergence of the total mass.

Define now the quantities $\tilde{M}_N(t)$ corresponding to the approximating jump processes in a similar way to (17). Assume also that we are working directly with the weakly convergent subsequence. We have then the transitions

$$\tilde{M}_N \to \tilde{M}_N + N^{-1}\left(\frac{f}{h^* - f} + \frac{f'}{h^* - f'}\right) \text{ at rate } Np(f')df' \cdot f m_N(t, f),$$

where f' is the fitness of the new generated node, f runs over all (finite) fitness values which are already present and $m_N(t, f)$ stands fot the total mass at fitness f.

By Dynkin's formula we obtain

$$\begin{aligned} \tilde{M}_N(t) &= \tilde{M}_N(0) + \int_0^t \int_0^{h^*} \int_0^{h^*} \left(\frac{f}{h^* - f} + \frac{f'}{h^* - f'}\right) p(f')df' \cdot f m_N(t, f) df ds + \\ &\quad + \mathbf{M}_N(t) \\ &= \tilde{M}_N(0) + \int_0^t h^*[\tilde{M}_N(s) + (I(h^*) - 1)\tilde{M}_1(s)]ds + \mathbf{M}_N(t) \\ &= \tilde{M}_N(0) + \int_0^t h^* \tilde{M}_N(s) ds + \mathbf{M}_N(t), \end{aligned} \tag{20}$$

where $\mathbf{M}_N(t)$ is the corresponding martingale for which we have

$$E[\mathbf{M}_N(t)^2] = \frac{1}{N} E[\int_0^t \int_0^{h^*} \int_0^{h^*} \left(\frac{f}{h^* - f} + \frac{f'}{h^* - f'}\right)^2 p(f')df' \cdot f m_N(t, f) df ds].$$

Now if $I(h) > 1$, then $h^* > h$ and we can bound the terms of the type $f/(h^* - f)$ by a constant independent on f, since f takes values actually only in $[0, h]$. This implies that $E[\mathbf{M}_N(t)^2] \to 0$ with the order $N^{-1/2}$. We obtain then immediately that $\sup_{t \le T} |\tilde{M}_N(t) - M(t)| \to 0$ in mean square, where M is the solution of (19). This means that in the limit all moments M_k, $k \ge 1$ are conserved, therefore also the total mass

$$M_0(t) = \int_0^{h^*} m(t, f) df$$

which satisfies the equation

$$\frac{dM_0(t)}{dt} = 2h^* M_1(t),$$

obtained also formally by summing up all equations in (4) and integrating w.r.t f.

This conservation property enables us to infer weak convergence and argue similarly as in Theorem 1 in order to get the desired result.

□

4. Numerical simulations

The basic setting which we consider is the following: define on $[0, h]$ the fitness values $f_j = h - h/j$ for $j = 1, M$ and the probability distribution $p_j = P(f_j) = j^{-\theta}/\zeta_M(\theta)$, where $\zeta_M(\theta) = \sum_{j=1}^{M} j^{-\theta}$. Consider only the case $A_k = k$, since this is the most interesting one. As stated in the introduction, one can compute numerically a critical value $\theta^* \approx 2.48$ such that $I(h) < 1$ for $\theta > \theta^*$.

We compare basically the numerical simulations for two values $\theta_1 < \theta^* < \theta_2$ which imply different regimes of our problem. Consider also $h = 10$ and $M = 100$.

We are interested to compute quantities like:

- $m(t, f_j)$: the total mass concetrated at fitness f_j
- $v_k(t, f_j)$: the mass of nodes of degree k and fitness f_j
- structure of the mass spectrum at a given fitness value.

Note that in the case $A_k = Ck$ we can in fact decouple the equations in a very convenient way, which considerably speeds up the simulations. Namely, in the equation for $v_1(t, f)$ for a given f are entering only the total masses $m(t, f_j)$ for all j. Therefore we can allow to compute the dynamics of the total mass cf. (15) at several f_j's, without destroying the full dynamics at the other fitness values. We basically do the following thing: if the support size of the mass spectrum corresponding to a given fitness becomes larger than a given bound, then we continue computing only the total mass corresponding to that fitness value. In this way we still track the complete mass spectrum at fitness values where we typically have one single large cluster and a few other smaller clusters. This pattern is also a numerical indication of the Bose-Einstein condensation, where a positive fraction of the total mass is concentrated in a single large cluster.

Depending on the situation, we will plot the quantities either with respect to the fitness index $j = 1 \dots M$, or with respect to the actual fitness value $f_j \in [0, h)$. Note that in the distribution we considered we have in fact $f_1 = 0$, so the nodes with fitness f_1 to not gather additional links, but they only attach to nodes with larger fitness values. The values corresponding to f_1 will be omitted from our plots. Depending on the situation, we will consider the values $N = 10^4, 10^5, 10^6, 10^7$.

We will compare first the convergence properties of the numerical scheme for $\theta_1 = 1.5$ and $\theta_2 = 4$ for the initial condition $v_1(0, f_2) = 0.01$ and 0 otherwise.

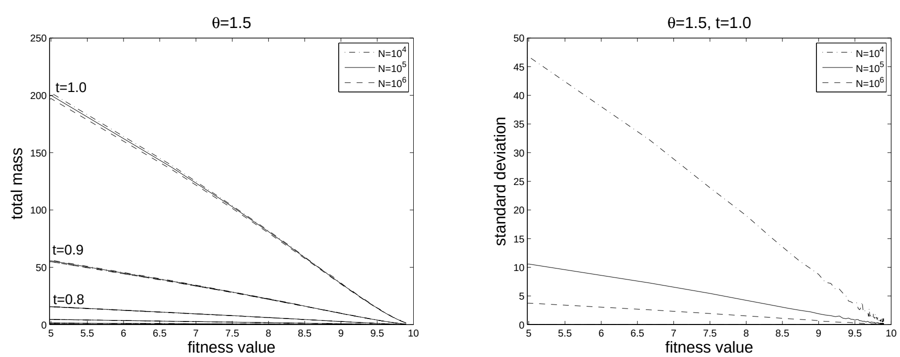

FIGURE 1. $m(t, f)$: mean value and standard deviation for $\theta = 1.5$, $t \leq 1$

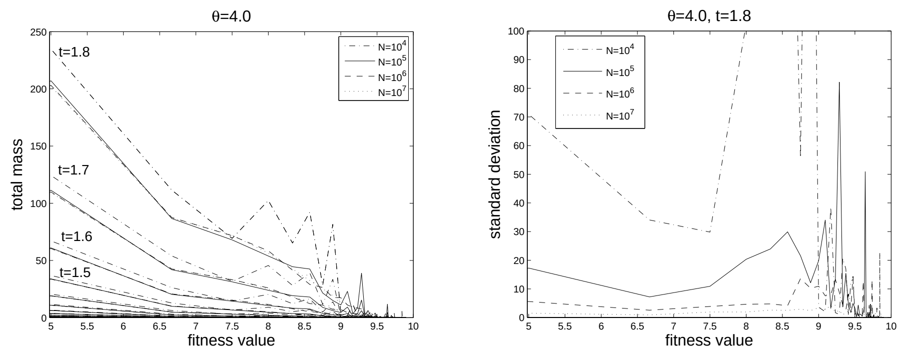

FIGURE 2. $m(t,f)$: mean value and standard deviation for $\theta = 4$, $t \leq 1.8$

The numerical values of the total mass (mean and standard deviation) obtained over 20 independent realizations are plotted in Figures 1 and 2. The time intervals were chosen such that the mass at the fitness value f_2 (where the initial condition was concentrated) should be approximatively equal in both cases, in order to have the same basis of comparison. For having a better picture, the standard deviation is plotted only at the maximal time. We note the clear convergence of the method for $\theta_1 = 1.5$, while for $\theta_2 = 4$ the convergence is significantly slower and less uniform. Only if we take $N = 10^7$ we can observe a convergence behaviour also at large fitness values. Note that the relative compactness property proved in Proposition 2 holds for $A_k \leq Ck$, but in this case, for $A_k = k$ and $I(h) < 1$, we have no information about the connection between the weak limits and the deterministic equation (4). We do not know if the limit is deterministic at all and the convergence which shows up only at a very high precision ($N = 10^7$) may be also due to the fact that from the numerical point of view we can work only with finitely supported distributions, for which we always have $I(h) \geq 1$.

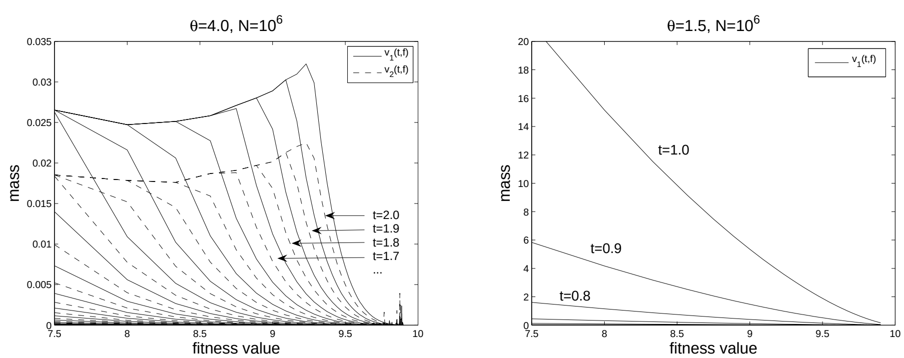

FIGURE 3. Equilibrium of the components for $\theta = 4$ (left picture) versus continous increase for $\theta = 1.5$ (right picture)

By plotting the mass concentrated at certain sizes we observe a different qualitative behaviour for the two values of θ. For $\theta = 4$ we observe that the quantities $v_k(t,f)$ increase up to a certain time t_f and then remain constant. This is shown in the left picture from Figure 3 for v_1 and v_2. For $\theta = 1.5$ we observe a completely different behaviour, namely a continuous increase of v_1 (right picture). Note also the significant discrepancy between the values of the component v_1 itself, remembering that both simulations started with the same initial condition and were stopped at times such that the total mass at f_2 had approximatively the same value (about 200). By taking $N = 10^4$ we were able to compute the same dynamics for $\theta = 1.5$ up to the time $t = 1.4$, but even in this case the pattern did not change. Theoretically it is possible that equilibrium shows up at later times, which are beyond our available computational possibilities, but the present numerical simulations show nevertheless a different tendency.

Arrived at this point, let us make some comments about the computational limits of the problem. Since the total rate is proportional to the total mass, which has essentially an exponential growth, if we arrive at very large values the time steps between the jumps become smaller and smaller. Therefore, giving a time increment for saving the results of 0.1 as in our case, we always end up at a certain problem time t^* from which we never reach the next time step $t^* + 0.1$ in an affordable physical time, since the time distance between two consecutive jumps decreases dramatically and the jumps tend to accumulate. This is the point T where we have to stop the simulations. Note that this has nothing to do with the stopping time τ_a introduced for technical reasons. a is a theoretical constant which we can take arbitrarily large. We "stop" the process in the stochastic sense only if the value of the total mass becomes dramatically larger than the expected value given by the deterministic dynamics. Due to the light-tailed property of exponential distribution of the waiting times between the jumps, such events are extremeley rare and in fact do not show up for a normal number of simulations.

Our next numerical experiment has a different nature. Consider $\theta_1 = 2.1 < \theta^* < \theta_2 = 2.6$ and as initial condition $v_1(0, f_2) = 10^{-1}, v_1(0, f_{90}) = 10^{-5}$ and 0 otherwise. We take $N = 10^5$. That is, we put a small amount of mass, equivalent to only one numerical particle, at one of the large fitness values and a much larger amount of mass at f_2. The behaviour for the two parameters is shown in Figure 4.

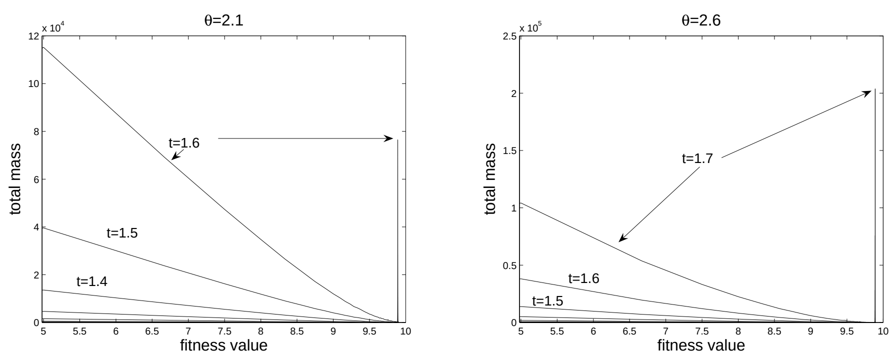

FIGURE 4. "First mover advantage" scenario for $\theta = 2.1$ (left picture) versus "fit-get-rich" scenario for $\theta = 2.6$ (right picture)

We note that for $\theta < \theta^*$ the properties of the fitness distribution do not allow that the mass remains concentrated at high fitness values and in fact the small fitness values will prevail. For $\theta > \theta^*$, even close to it like in our example, we notice an opposite tendency, namely that the mass concentrated at higher fitness values will prevail. For larger θ this phenomenon is certainly even more accentuated, as we will see further.

Consider $\theta = 4$ and take for the moment the same initial condition as in the first numerical example. That is we take $v_1(0, f_2) = 0.01$, all other components being 0. Instead of running a larger number of independent simulations, we will present only two runs, but computed up to a larger time, close to the computational limit described before. Here we take again $N = 10^5$.

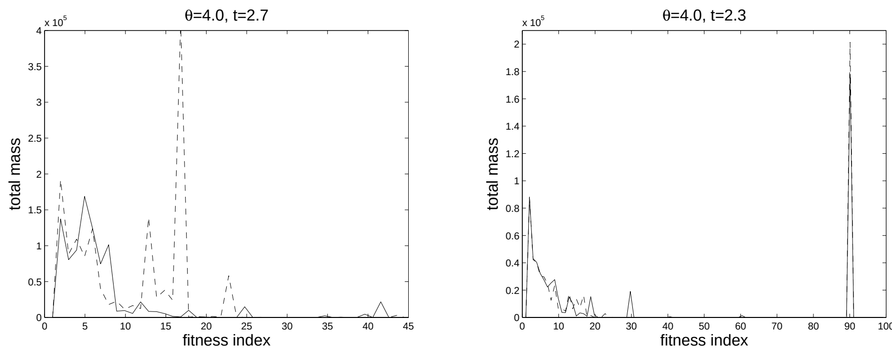

FIGURE 5. Two computations of the total mass for $\theta = 4$ and different initial conditions: appearance of mass concentrations at large fitness values (left picture) versus prevailance of mass at large fitness values (right picture)

Even if initially we had mass only at f_2, in the left picture in Figure 5 at later times we notice the appearance of significant mass concentrations at larger fitness values. Note that the fitness index 10 corresponds to the actual fitness value of 9, in the interval $[0, 10]$. Since the values of the mass have become too large, we cannot compute further. But, due to the linearity of the problem, we may "rescale" the whole setting and consider an initial condition similar to that corresponding to the simulations in Figure 4: a larger mass at small fitness values and a very small mass at a large fitness value. The absolute values are not important. The results are plotted in the right picture of Figure 5 and show the prevailance of the mass concentrated at larger fitness values, where the initial "seed" has been placed. This right plot is the equivalent to Figure 4, but for $\theta = 4$. We see clearly that the growth of the total mass at this large fitness value is even more accentuated.

A similar pattern can be observed in Figure 6, where we take as initial condition $v_1(0, f_i) = (10 \cdot (i-1))^{-2}$ for $i = 2, 4, 6$, $v_1(0, f_{90}) = 10^{-5}$ and 0 otherwise. We now put some additional mass also at the larger fitness values f_4 and f_6, which is again significantly larger than the mass placed initially at f_{90}.

We note the four peaks corresponding to the fitness values where we have placed the initial mass "seeds". Their height is almost in an inverse relation to the initial values of the mass: the peak at f_{90} is higher than those at f_2 and f_4 and on the same level with that at f_6, despite the initial proportions of the mass values.

We will show next a numerical indication of the Bose-Einstein condensation, where a macroscopic amount of mass is concentrated in one single huge cluster. Figure 7 shows the mass spectrum at the fitness values where the second moment

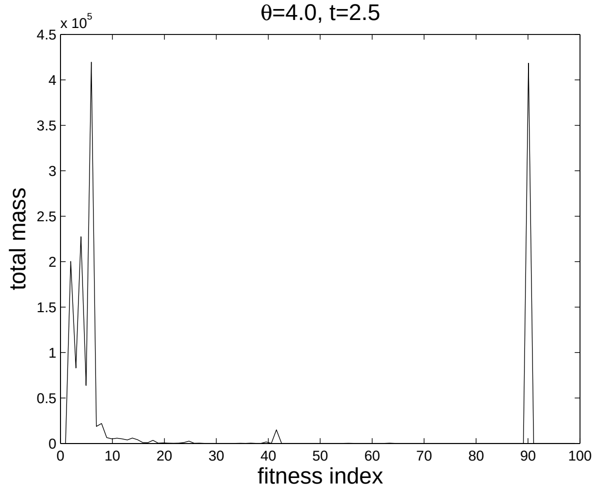

FIGURE 6. The growth speed of the total mass is increasing with the fitness

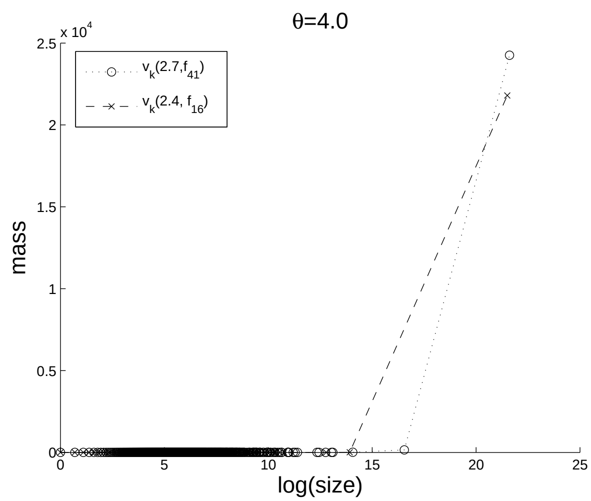

FIGURE 7. Numerical evidence for the "innovation-pays-off" scenario

in the left picture of Figure 5 is maximal. They correspond to the highest peak in the picture (realization corresponding to the dotted line, f_{16} at $t = 2.4$), respectively the peak at the right end of the picture (realization corresponding to the continuous line, f_{41} at $t = 2.7$). In both cases almost all mass at the corresponding fitness values is concentrated in a single huge cluster. The observed phenomenon is due to the small probability of appearance of clusters with high fitness values. But, once they appear, they become dominant in time. The "loss" of mass which can be shown only theoretically occurs due to a continouos shift towards the right end of the interval of peaks corresponding to large clusters like those presented above, which in the long run always dominate the previous ones.

5. Conclusions

In this paper we introduced an infinite system of equations of Becker-Döring type which models the time dynamics of network growth in the context of preferential attachment with fitness. We construct a stochastic scheme and give conditions under which we obtain convergence to solutions of the system. In a special case, which is the mostly treated in the literature on random graphs, we show uniqueness as well as presence of Bose-Einstein condensation, which is consistent with the existing results on asymptotic behaviour which disregard the exact time evolution. The equations are linear, which on the one hand is an advantage for handling them. On the other hand, in contrast to the usual Becker-Döring or general coagulation-fragmentation equations, they are nonconservative, i.e. the total mass grows continuously. Exactly this is the phenomenon which leads to the main difficulty, especially when trying to prove the limit dynamics for the general case $A_k = O(k)$. This is also the main challenge for their numerical simulations, since a very large value of the total mass implies that the time steps between two consecutive jumps become so small, that in fact the jumps tend to accumulate and we can not afford to wait an arbitrarily long (physical) time in order to perform computations on an arbitrary time interval $[0, T]$ of the problem. The maximal possible T is practically prescribed by the problem parameters and the computational resources. The mentioned difficulties are therefore a challenge for future research in this direction.

References

[1] Bauer, H., *Maß- und Integrationstheorie*, (2.Auflage), de Gruyter, 1992
[2] Bianconi, G., Barábasi, A.-L., *Bose-Einstein condensation in complex networks*, Phys.Rev.Lett., 86(24) (2001), pp.5632-5635
[3] Borgs, C., Chayes, J.T., Daskalakis, C., Roch, S., *First to market is not everything: an analysis of preferential attachment with fitness*, Proceedings of ACM STOC 2007, pp. 135-144
[4] Ball, J.M., Carr, J., Penrose, O., *The Becker-Döring cluster equations: basic properties and asympotic behaviour of solutions*, Commun. Math. Phys. 104 (1986), pp.657-692
[5] Ethier, S., Kurtz, T.G., *Markov Processes, Characterization and Convergence*, Wiley, New York, 1986.
[6] Jacod, J., Protter, P., *Probability Essentials* (second edition), Springer, 2003
[7] Krapivsky, P.L., Redner, S., *Organization of growing random networks*, Phys.Rev. E, (63)(6) (2001), pp. 066123-1-066123-14

Department of Mathematics, Dortmund University of Technology, Vogelpothsweg 87, 44221 Dortmund, Germany
E-mail address: flavius.guias@math.uni-dortmund.de

Preprints ab 2008

2008-01 **Henryk Zähle**
Weak approximation of SDEs by discrete-time processes

2008-02 **Benjamin Fine, Gerhard Rosenberger**
An Epic Drama: The Development of the Prime Number Theorem

2008-03 **Benjamin Fine, Miriam Hahn, Alexander Hulpke, Volkmar große Rebel, Gerhard Rosenberger, Martin Scheer**
All Finite Generalized Tetrahedron Groups

2008-04 **Ben Schweizer**
Homogenization of the Prager model in one-dimensional plasticity

2008-05 **Benjamin Fine, Alexei Myasnikov, Gerhard Rosenberger**
Generic Subgroups of Group Amalgams

2008-06 **Flavius Guiaş**
Generalized Becker-Döring Equations Modeling the Time Evolution of a Process of Preferential Attachment with Fitness