

Universality in a class of fragmentation-coalescence processes

A. E. Kyprianou, S. W. Pagett and T. Rogers

Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK.

E-mail: a.kyprianou@bath.ac.uk; s.pagett@bath.ac.uk; t.c.rogers@bath.ac.uk

Received 23 December 2015; revised 30 March 2017; accepted 30 March 2017

Abstract. We introduce and analyse a class of fragmentation-coalescence processes defined on finite systems of particles organised into clusters. Coalescent events merge multiple clusters simultaneously to form a single larger cluster, while fragmentation breaks up a cluster into a collection of singletons. Under mild conditions on the coalescence rates, we show that the distribution of cluster sizes becomes non-random in the thermodynamic limit. Moreover, we discover that in the limit of small fragmentation rate these processes exhibit a universal cluster size distribution regardless of the details of the rates, following a power law with exponent $3/2$.

Résumé. Nous introduisons et étudions une classe de processus de fragmentation-coalescence définis sur des systèmes finis de particules organisées en amas. Lors d'un événement de coalescence, de multiples amas fusionnent simultanément en un amas plus gros, et dans le cas d'un événement de fragmentation, un amas est détruit en une collection de singletons. Sous des hypothèses faibles sur les taux de coalescence, nous montrons que la loi des tailles des amas devient déterministe dans la limite thermodynamique. De plus, nous montrons que dans la limite des petits taux de fragmentation, ces processus admettent une loi universelle pour la les tailles des amas, indépendante de la valeur précise des taux, et qui suit une loi de puissance avec un exposant $3/2$.

MSC: 60G99

Keywords: Fragmentation-coalescence; Cluster distribution; Thermodynamic limit

1. Introduction

Processes of coalescence and its reverse, fragmentation, have been widely studied in physical chemistry since the seminal work of Smoluchowski [19] a century ago (see *e.g.* [1,12,18] and references therein). Aside from chemical systems, these processes also have an important role to play in modelling genealogy [13,20] and even the dynamics of group formation [8]. The microscopic specification of a fragmentation-coalescence process is usually given in terms of a stochastic process acting on a finite number of constituent particles, introduced by Marcus [16], Lushnikov [14, 15] and Gueron [7]. However, in applied work it is common to write deterministic 'mean-field' equations that are intended to describe behaviour in thermodynamic limit of large system size. An important question arises: are these processes self-averaging so that mean-field calculations are relevant [6,11]?

Another key concept in the understanding of large-scale interacting systems is that of universality – that certain important macroscopic properties often do not depend on the detailed features of the particles and dynamics involved, but rather a much smaller set of properties determine how these processes behave in the thermodynamic limit. This is particularly true of critical phenomena, a famous example being the directed percolation phase transition [10]. In fragmentation-coalescence processes, the antagonistic nature of the driving mechanisms can give rise to self-organised criticality in certain limits, whereby the system evolves to a stationary distribution that exhibits scale-invariant behaviour normally characteristic of a phase transition [17]. It is therefore natural to ask if this behaviour is universal.

Here we answer both of the above questions in the affirmative. We examine a class of fragmentation-coalescence processes that allow simultaneous coagulation of multiple clusters with different rates, combined with independent fragmentation events that shatter clusters to singletons. The fragmentation and coalescence rates are taken to be independent of cluster size, and we make a mild technical assumption on the coalescence rates. We show that processes of this class approach a non-random limit as the system size grows, and moreover, that in the limit of small fragmentation rate we observe a universal cluster size distribution with a power-law tail of exponent $3/2$.

1.1. Finite fragmentation-coalescence processes

Consider a collection of n identical particles, grouped together into some number of clusters. We define a stochastic dynamical process as follows:

1. Every subset of k clusters coalesces at rate $\alpha(k)n^{1-k}$, independently of everything else that happens in the system. The coalescing clusters are merged to form a single cluster with size equal to the sum of the sizes of the merged clusters.
2. Clusters fragment at rate $\lambda_n > 0$, independently of everything else that happens in the system. Fragmentation of a cluster of size k results in k ‘singleton’ clusters of size one.

The standard choice of initial condition is the state with n singleton clusters and this will be the case throughout this paper. The factor of n^{1-k} appearing in the coalescence rates is included to compensate for the combinatorial explosion in the number of subsets of k clusters as n gets larger. In this way, when there are order n clusters, the global rates of fragmentation and coalescence of any number of clusters are all order n (in the case where $\lambda_n \equiv \lambda \in (0, \infty)$). Note that this choice is necessary to ensure that there is a single dominant timescale for the dynamics.

For each $n \in \mathbb{N}$, and $k \in \{1, \dots, n\}$, the state of the system is specified by the number of clusters of size k at time t . To that end we introduce the random variables

$$w_{n,k}(t) := \frac{1}{n} \#\{\text{clusters of size } k \text{ at time } t\}, \quad 1 \leq k \leq n,$$

which we take to be right continuous with left limits, and combine these into the random vector

$$\mathbf{w}_n(t) := (w_{n,1}(t), \dots, w_{n,n}(t)),$$

which takes values on the simplex

$$\mathcal{S}_n := \left\{ \mathbf{w}_n = (w_{n,1}, \dots, w_{n,n}) : nw_{n,i} \in \{1, \dots, n\} \text{ for } i \in \{1, \dots, n\}, \sum_{k=1}^n kw_{n,k} = 1 \right\}.$$

Another natural quantity is the empirical cluster size distribution, defined by

$$p_{n,k}(t) := \frac{\#\{\text{clusters of size } k \text{ at time } t\}}{\#\{\text{clusters at time } t\}}, \quad 1 \leq k \leq n. \tag{1.1}$$

From the description of the dynamics at the beginning of this subsection we can write down the infinitesimal generator of $\mathbf{w}_n(t)$ by summing over all possible events. To do this we must consider the number of possible ways of coalescing clusters of sizes l_1, \dots, l_k . If the l_i are distinct then we simply multiply together the number of clusters of each size. If some of the l_i are the same then there is a complication as a cluster cannot coalesce with itself. To each configuration l_1, \dots, l_k we can associate a partition π of $\{1, \dots, k\}$ by $l_u = l_v$ if and only if $u, v \in \pi_i$ for some i . In this way the set \mathcal{P}_k , of all partitions of $\{1, \dots, k\}$, enumerates all the ways we could have chosen k cluster sizes with replacement. Hence,

$$\begin{aligned} \mathcal{A}_n f(\mathbf{w}_n) &= \lambda_n n \sum_{k=1}^n \left(f\left(\mathbf{w}_n - \frac{1}{n} \mathbf{e}_{n,k} + \frac{k}{n} \mathbf{e}_{n,1}\right) - f(\mathbf{w}_n) \right) w_{n,k} \\ &\quad + \sum_{k=2}^n \frac{\alpha(k)}{k! n^{k-1}} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|}=1}^n \gamma_n(f, \pi, \mathbf{w}_n), \end{aligned} \tag{1.2}$$

where

$$\gamma_n(f, \pi, \mathbf{w}_n) := \left(f \left(\mathbf{w}_n - \frac{1}{n} \sum_{i=1}^{|\pi|} |\pi_i| \mathbf{e}_{n,l_i} + \frac{1}{n} \mathbf{e}_{n, \sum |\pi_i| l_i} \right) - f(\mathbf{w}_n) \right) \prod_{i=1}^{|\pi|} (n w_{n,l_i})_{|\pi_i|}, \tag{1.3}$$

where $\mathbf{e}_{n,i}$ is a vector of length n with one in position i and zero everywhere else, $|\pi|$ is the number of blocks in the partition, $|\pi_i|$ is the number of elements in the i th block of π and $(y)_z$ is the falling factorial Pochhammer symbol.

Rather than working with \mathbf{w}_n directly, considerable simplification is possible using the empirical generating functions $G_n : [0, 1] \times \mathbb{R}^+ \rightarrow \mathbb{R}$

$$G_n(x, t) = \sum_{k=1}^n x^k w_{n,k}(t), \quad n \geq 1, \tag{1.4}$$

and $g_n : [0, 1] \times \mathbb{R}^+ \rightarrow \mathbb{R}$

$$g_n(x, t) := \sum_{k=1}^n x^k p_{n,k}(t) = \frac{G_n(x, t)}{G_n(1, t)}, \quad n \geq 1. \tag{1.5}$$

1.2. Main results

Our first main result shows that processes in the class outlined in Section 1.1 are self-averaging with respect to the distribution of cluster sizes as $n \rightarrow \infty$.

Theorem 1. *Suppose that the coalescence rates $\alpha : \mathbb{N} \rightarrow \mathbb{R}^+$ satisfy*

$$\alpha(k) \leq C \exp(\gamma k \ln \ln(k)), \quad \forall k, \tag{1.6}$$

for some constants $C > 0$ and $\gamma < 1$. Define $\lambda = \lim_{n \rightarrow \infty} \lambda_n$, and let $G : [0, 1] \times \mathbb{R}^+ \rightarrow \mathbb{R}$ be the unique solution of the deterministic initial value problem

$$\begin{aligned} G(x, 0) &= x, \\ \frac{\partial G}{\partial t}(x, t) &= \lambda(x - G(x, t)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G(x, t)^k - kG(1, t)^{k-1}G(x, t)). \end{aligned} \tag{1.7}$$

If $\lambda > 0$ then in the limit $n \rightarrow \infty$ the empirical generating function $G_n(x, t)$ defined in (1.4) converges to $G(x, t)$ in L^2 , uniformly in x and t . That is,

$$\sup_{x \in [0, 1], t \in \mathbb{R}^+} \mathbb{E} \left[(G(x, t) - G_n(x, t))^2 \right] \leq \frac{H}{\sqrt{n}} \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

for some constant H . If $\lambda = 0$ then the above result holds for t less than any fixed finite time T .

Remark 1. It should be noted that condition (1.6) is a sufficient technical condition, but may not be minimal. In addition, it is not known if $n^{-1/2}$ is the optimal rate of convergence.

Unpacking the limiting generating function $G(x, t)$ allows us to write Smoluchowski-type differential equations for the (rescaled) number of clusters of a given size.

Corollary 1. Under the assumptions of Theorem 1, we have L^2 convergence $w_{n,j}(t) \rightarrow w_j(t)$, obeying

$$\begin{aligned}
 w_j(0) &= \delta_{j,1}, \\
 \frac{d}{dt} w_1(t) &= \lambda(1 - w_1(t)) - w_1(t) \sum_{k=2}^{\infty} \frac{\alpha(k)}{(k-1)!} \left(\sum_{l=1}^{\infty} w_l(t) \right)^{k-1}, \\
 \frac{d}{dt} w_j(t) &= -\lambda w_j(t) - w_j(t) \sum_{k=2}^{\infty} \frac{\alpha(k)}{(k-1)!} \left(\sum_{l=1}^{\infty} w_l(t) \right)^{k-1} + \sum_{k=2}^j \frac{\alpha(k)}{k!} \sum_{\substack{l_1, \dots, l_k \\ l_1 + \dots + l_k = j}} w_{l_1}(t) \cdots w_{l_k}(t)
 \end{aligned} \tag{1.8}$$

for $j \geq 2$. Furthermore, the limiting cluster size distribution may be obtained via $p_k(t) = w_k(t) / \sum_{l=1}^{\infty} w_l(t)$.

Our second main result deals with how the limiting process behaves in long times when fragmentation events are rare. To that end, we define the stationary cluster size distribution

$$p_k := \lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} p_{n,k}(t),$$

where $p_{n,k}(t)$ is defined in (1.1) and convergence in the large n limit should be understood in terms of Theorem 1.

Theorem 2. If α satisfies (1.6) and m is the smallest integer such that $\alpha(m) > 0$, then the stationary cluster size distribution exists and is unique. In particular, the limit of this stationary distribution as $\lambda \rightarrow 0$ exists and obeys

$$\lim_{\lambda \searrow 0} p_k = \begin{cases} \frac{1}{k} \left(\frac{m-1}{m}\right)^k \left(\frac{1}{m}\right)^{\frac{k-1}{m-1}} \binom{m \left(\frac{k-1}{m-1}\right)}{\frac{k-1}{m-1}} & \text{if } m-1 \text{ divides } k-1, \\ 0 & \text{otherwise.} \end{cases}$$

In particular $\lim_{\lambda \searrow 0} p_k \sim k^{-3/2}$, if $m-1$ divides $k-1$, where $f \sim g$ means $f(k)/g(k) \rightarrow c$ for some positive constant c , as $k \rightarrow \infty$, regardless of α .

The power-law tail of cluster size distribution is reminiscent of critical behaviour occurring at the gellation time in some pure coalescent processes [1], although here it is not a transient phenomenon but rather is approached in the long time limit. This is the essential characteristic of self-organised criticality, however, there is an unusual complication here that the result depends on the strict order of the $n \rightarrow \infty$ and $\lambda \rightarrow 0$ limits; taking $\lambda_n \rightarrow 0$ we recover a pure coalescent process that does not possess this characteristic power-law.

While it is quite typical for critical processes to exhibit universality in the scaling exponents, our result says something stronger that in fact the limiting cluster size distribution is almost completely independent of the coagulation rates, depending only on the identity of the first non-zero rate. A surprising consequence is that if, for example, the model coalesces clusters in groups of three and four (but not pairs) then in the large n and small λ limit we will see no clusters of even size whatsoever in the stationary distribution. Figure 1 shows an example of this phenomenon for the model with rates $\alpha(k) = \delta_{k,3} + 2\delta_{k,4}$. The model has the apparently paradoxical feature that clusters of even size are vanishingly rare, despite the fact that some two-thirds of clusters are singletons, and $\alpha(4) > \alpha(3)$.

We give here a brief heuristic as to why this is the case. The coalescence rates are scaled so that when the number of clusters is of order n the total coalescence rate of events involving k blocks is of the same order for all k . However, Theorem 2 states as $\lambda \rightarrow 0$ that the stationary cluster size distribution follows a power law with exponent $3/2$, which suggests that the average cluster has size of order $n^{1/2}$. Hence, the number of clusters falls below order n . This causes coalescence events of the smallest number of clusters, m , to occur much more frequently than all others and so eventually we only see clusters that could have arisen from coalescing m clusters. These are clusters of size k where $m-1$ divides $k-1$. In the above example where $m=3$, this results in the distribution only giving mass to odd-sized blocks.

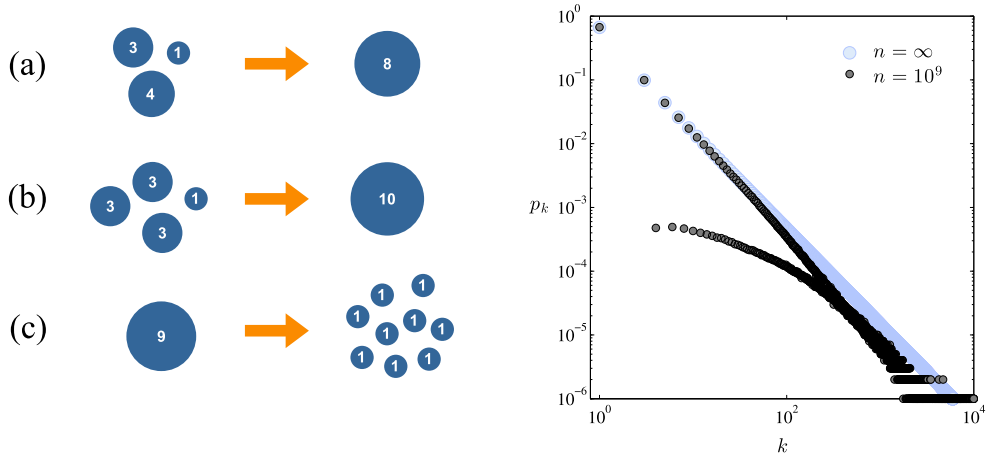


Fig. 1. Example of a fragmentation-coalescence process with fragmentation rate $\lambda = 10^{-7}$ and coalescence rates $\alpha(3) = 1$, $\alpha(4) = 2$ and $\alpha(k) = 0$ otherwise. Left: illustration of possible events (a) coalescence of three clusters, (b) coalescence of four clusters, (c) fragmentation of a cluster. Right: comparison between a single simulation run with $n = 10^9$ particles, stopped at time $T = 10^6$ (black) and the $n \rightarrow \infty$ limit given by Theorem 2 (shaded lighter). The higher points correspond to the odd sized blocks, the lower ones correspond to even sized blocks which are asymptotically vanishing in the limit $n \rightarrow \infty$, $\lambda \rightarrow 0$.

1.3. Examples of related work

Without fragmentation one is left with a pure coalescent process [1]. The prototypical example of this class has pairs of clusters coalescing at unit rate. In this case, the $w_{n,k}(t)$ converge in probability to solutions to the Smoluchowski coagulation equations [19] with unit rate kernel:

$$\frac{d}{dt}w_k(t) = \frac{1}{2} \sum_{j=1}^{k-1} w_j(t)w_{k-j}(t) - w_k(t) \sum_{j=1}^{\infty} w_j(t), \quad k \geq 1.$$

These equations can be solved to show

$$w_k(t) = \left(1 + \frac{t}{2}\right)^{-2} \left(\frac{t}{2+t}\right)^{k-1}, \quad k \geq 1, t \geq 0,$$

and so the proportion of clusters of size k decays exponentially in k . This behaviour is typical of pure coalescence processes.

A simple modification of these dynamics is to sized-biased fragmentation-coalescence processes, a particular example of which (again with only pairs coalescing) has been studied in the context of terrorist networks [4]. A thorough analysis of this model was undertaken by Rath and Toth in [17]. For a constant fragmentation rate λ it was found that the $w_{n,k}(t)$, converge in probability as $n \rightarrow \infty$ to $w_k(t)$ solving a set of differential equations similar to those of Smoluchowski:

$$\frac{d}{dt}w_k(t) = \frac{1}{2} \sum_{j=1}^{k-1} j(k-j)w_j(t)w_{k-j}(t) - kw_k(t) - \lambda kw_k(t), \quad k \geq 2,$$

$$\frac{d}{dt}w_1(t) = -w_1(t) + \lambda \sum_{j=2}^{\infty} j^2 w_j(t).$$

And so it was shown in the long time stationary limit, as $t \rightarrow \infty$, that

$$w_k \sim \exp\left(k \log\left(1 - \frac{\lambda^2}{(1+\lambda)^2}\right)\right) k^{-5/2}.$$

As in our model a power-law tail emerges in the limit $\lambda \rightarrow 0$, this time with exponent $5/2$. Interestingly in this case the critical behaviour survives the joint limit $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$; the main effort of [17] was focused on this scaling regime.

The critical behaviour of the stationary cluster size distribution, found in Theorem 2, has also been seen in a different form of fragmentation-coalescence process studied by Bressaud and Fournier [5]. This was a mean-field approximation to a one-dimensional forest-fire model on \mathbb{Z} . Edges (each of weight one) are glued together if the site between them is occupied, which happens at rate 1, and clusters of k edges have all sites removed at rate $(k - 1)/n$. A mean-field approximation to this model, ignoring correlations, follows

$$\frac{d}{dt} w_1^{(n)}(t) = -2w_1^{(n)}(t) + \frac{1}{n} \sum_{k \geq 2} k(k - 1)w_k^{(n)}(t),$$

$$\frac{d}{dt} w_k^{(n)}(t) = -(2 + (k - 1)/n)w_k^{(n)}(t) + \left(\sum_{l \geq 1} w_l^{(n)}(t) \right)^{-1} \sum_{i=1}^{k-1} w_i^{(n)}(t)w_{k-i}^{(n)}(t), \quad k \geq 2.$$

Then, they found that the stationary cluster size distribution converges weakly to $(p_k)_{k \geq 1}$ as $n \rightarrow \infty$ ('fragmentation' rate converges to zero) where

$$p_k = \frac{2}{4^k k} \binom{2k - 2}{k - 1}$$

which matches the distribution in Theorem 2 in the case where $m = 2$.

Finally we mention an important separate class of fragmentation-coalescence processes. The models discussed so far all concern large but finite particle systems, but it is possible to define (non-size-biased) fragmentation and coalescence processes on the partitions of \mathbb{N} , as introduced in [3]. The class of processes we work with here can be seen as a natural finite counterpart to these models.

2. Thermodynamic limit

We will focus our proof on the case of fixed fragmentation rate $\lambda_n \equiv \lambda \in (0, \infty)$; the extension to the joint limit differs only in a few places. Besides standard generating function technology, our methods rely on Gronwall's inequality [22] which we reproduce below for convenience.

Lemma 1 (Gronwall's inequality). *Suppose $f : [0, \infty) \rightarrow \mathbb{R}$ is differentiable and satisfies the following differential inequality*

$$\frac{d}{dt} f(t) \leq af(t) + b, \quad t > 0,$$

where $a, b \in \mathbb{R}$. Then, we have that for all $t \geq 0$

$$f(t) \leq \begin{cases} f(0)e^{at} + \frac{b}{a}(e^{at} - 1) & \text{if } a \neq 0, \\ f(0) + bt & \text{if } a = 0. \end{cases}$$

The idea of the proof of Theorem 1 is to bound the derivative of the expected squared difference between $G(x, t)$ and $G_n(x, t)$ in such a way so as we may apply Lemma 1, where, here, a will be negative and $b = b_n$ will decay to zero as $n \rightarrow \infty$. To achieve this, we will need to prove that the derivative exists and certain error quantities (specified in the next section) converge to zero as n tends to infinity.

2.1. Mean-field calculation

We first undertake a mean-field calculation to determine a viable candidate for $\lim_{n \rightarrow \infty} G_n$ under self-averaging.

Lemma 2. For $f_0(x, \mathbf{w}_n) := \sum_{k=1}^n x^k w_{n,k}$, we have

$$\mathcal{A}_n f_0(x, \mathbf{w}_n) = \lambda(x - f_0(x, \mathbf{w}_n)) + \sum_{k=2}^n \frac{\alpha(k)}{k!} (f_0(x, \mathbf{w}_n)^k - k f_0(1, \mathbf{w}_n)^{k-1} f_0(x, \mathbf{w}_n)) + \beta_n(x, \mathbf{w}_n),$$

where

$$\sup_{\mathbf{w}_n \in \mathcal{S}_n} |\beta_n(x, \mathbf{w}_n)| \leq \frac{A}{n},$$

where A is a constant independent of n and x .

Proof. We use the infinitesimal generator from (1.2) in the special case where $f = f_0(x, \cdot)$, with x fixed. Hence

$$\begin{aligned} \mathcal{A}_n f_0(x, \mathbf{w}_n) &= -\lambda \sum_{k=1}^n x^k w_{n,k} + \lambda \sum_{k=1}^n k x w_{n,k} \\ &\quad + \sum_{k=2}^n \frac{\alpha(k)}{k! n^k} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \left(\prod_{i=1}^{|\pi|} x^{|\pi_i| l_i} - \sum_{i=1}^{|\pi|} |\pi_i| x^{l_i} \right) \prod_{i=1}^{|\pi|} (n w_{n,l_i})^{|\pi_i|}. \end{aligned}$$

As n grows large, we claim that the dominant contribution from the final product inside the third term on the right-hand side is simply $n^k w_{n,l_1} \cdots w_{n,l_k}$. If $\pi = \{1, \dots, k\}$, then this is the only term, otherwise there is a subdominant correction resulting from the fact that clusters cannot coagulate with themselves. Labelling this correction as $\beta_n(x, \mathbf{w}_n)$, which we will later bound, we obtain

$$\begin{aligned} \mathcal{A}_n f_0(x, \mathbf{w}_n) &= -\lambda \sum_{k=1}^n x^k w_{n,k} + \lambda \sum_{k=1}^n k x w_{n,k} \\ &\quad + \sum_{k=2}^n \frac{\alpha(k)}{k!} \sum_{l_1=1}^n \cdots \sum_{l_k=1}^n \left(\prod_{i=1}^k x^{l_i} - \sum_{i=1}^k x^{l_i} \right) w_{n,l_1} \cdots w_{n,l_k} \\ &\quad + \beta_n(x, \mathbf{w}_n) \\ &= \lambda(x - f_0(x, \mathbf{w}_n)) + \sum_{k=2}^n \frac{\alpha(k)}{k!} (f_0(x, \mathbf{w}_n)^k - k f_0(1, \mathbf{w}_n)^{k-1} f_0(x, \mathbf{w}_n)) + \beta_n(x, \mathbf{w}_n). \end{aligned} \tag{2.1}$$

It remains to show that $\beta_n(x, \mathbf{w}_n) \leq A/n$ for some constant $A > 0$. We know it is the remainder after subtracting the dominant term and so is

$$\beta_n(x, \mathbf{w}_n) = \sum_{k=2}^n \frac{\alpha(k)}{k!} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \left(\prod_{i=1}^{|\pi|} x^{|\pi_i| l_i} - \sum_{i=1}^{|\pi|} |\pi_i| x^{l_i} \right) \left(\prod_{i=1}^{|\pi|} \frac{(n w_{n,l_i})^{|\pi_i|}}{n^{|\pi_i|}} - \prod_{i=1}^{|\pi|} w_{n,l_i}^{|\pi_i|} \right).$$

Each multinomial in the second bracket, when expanded, has at most 2^k terms with varying powers of $1/n$ larger than or equal to one. Also, we can see that all the terms are divisible by $w_{n,l_1} \cdots w_{n,l_{|\pi|}}$ and have a coefficient that is less than or equal to $(k-1)^j$, where j is the power of $1/n$ in that term. Hence,

$$|\beta_n(x, \mathbf{w}_n)| \leq \sum_{k=2}^n \frac{\alpha(k)}{(k-1)!} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \frac{2^k (k-1)}{n} w_{n,l_1} \cdots w_{n,l_{|\pi|}}.$$

Summing over $l_1, \dots, l_{|\pi|}$ we have that,

$$|\beta_n(x, \mathbf{w}_n)| \leq \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)2^k}{(k-2)!} \sum_{\pi \in \mathcal{P}_k} f_0(1, \mathbf{w}_n)^{|\pi|}.$$

Then noting that $f_0(1, \mathbf{w}_n)^{|\pi|} \leq 1$ for all π , we have

$$|\beta_n(x, \mathbf{w}_n)| \leq \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)2^k}{(k-2)!} B_k,$$

where B_k denotes the k th Bell number and is the size of the set \mathcal{P}_k . We bound B_k using a recent result from [2], obtaining

$$|\beta_n(x, \mathbf{w}_n)| \leq \frac{1}{n} \left| \sum_{k=2}^n \frac{\alpha(k)}{(k-2)!} \left(\frac{1.584k}{\ln(k+1)} \right)^k \right|.$$

If we look at the asymptotic behaviour of the summands for large k , we see that

$$\frac{1}{(k-2)!} \left(\frac{1.584k}{\ln(k+1)} \right)^k \sim \exp(-k \ln \ln(k) + O(k)),$$

where \sim is defined in Theorem 2. Thus, by assumption (1.6) on α , the series converges as required. □

2.2. Self-averaging

With the mean-field behaviour determined, we proceed to the proof of L^2 convergence for sample paths. The following two lemmas will help us determine the behaviour of G_n^2 .

Lemma 3. *The function $f_0(x, \mathbf{w}_n) = \sum_{k=1}^n x^k w_{n,k}$ satisfies*

$$\mathcal{A}_n f_0^2(x, \mathbf{w}_n) = 2f_0(x, \mathbf{w}_n)\mathcal{A}_n f_0(x, \mathbf{w}_n) + \frac{\lambda x^2}{n} f_2(1, \mathbf{w}_n) + h(x, \mathbf{w}_n),$$

where

$$f_2(x, \mathbf{w}_n) := \sum_{k=1}^n k^2 x^k w_{n,k}, \quad \text{and} \quad \sup_{\mathbf{w}_n \in \mathcal{S}_n} h(x, \mathbf{w}_n) \leq \frac{E}{n},$$

such that E is a constant independent of n and x .

Proof. We plug the function $f_0^2(x, \cdot)$, with x fixed, into the infinitesimal generator (1.2)

$$\begin{aligned} \mathcal{A}_n f_0^2(x, \mathbf{w}_n) &= \lambda_n n \sum_{k=1}^n \left(\left(f_0(x, \mathbf{w}_n) + \frac{kx - x^k}{n} \right)^2 - f_0^2(x, t) \right) w_{n,k} \\ &\quad + \sum_{k=2}^n \frac{\alpha(k)}{k! n^{k-1}} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}}^n \gamma_n(f_0^2(x, \cdot), \pi, \mathbf{w}_n). \end{aligned}$$

where γ is defined in (1.3). By expanding out the brackets we see that

$$\mathcal{A}_n f_0^2(x, \mathbf{w}_n) = 2f_0(x, \mathbf{w}_n)\mathcal{A}_n f_0(x, \mathbf{w}_n) + \frac{\lambda x^2}{n} f_2(1, \mathbf{w}_n) + h(x, \mathbf{w}_n),$$

where

$$\begin{aligned}
 h(x, \mathbf{w}_n) &= \frac{\lambda}{n} \sum_{k=1}^n (x^{2k} - 2kx^{k+1}) w_{n,k} \\
 &\quad + \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)}{k!} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \left(\prod_{i=1}^{|\pi|} x^{|\pi_i| l_i} - \sum_{i=1}^{|\pi|} |\pi_i| x^{l_i} \right)^2 \prod_{i=1}^{|\pi|} \frac{(n w_{n,l_i})^{|\pi_i|}}{n^{|\pi_i|}} \\
 &\leq -\frac{2\lambda}{n} x^2 + \frac{\lambda}{n} f_0(x^2, \mathbf{w}_n) + \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)}{k!} f_0(x^2, \mathbf{w}_n)^k \\
 &\quad - \frac{2}{n} \sum_{k=2}^n \frac{\alpha(k)}{k!} k f_0(x^2, \mathbf{w}_n) f_0(x, \mathbf{w}_n)^{k-1} + \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)}{k!} k f_0(x^2, \mathbf{w}_n) f_0(1, \mathbf{w}_n)^{k-1} \\
 &\quad + \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)}{(k-2)!} f_0(x, \mathbf{w}_n)^2 f_0(1, \mathbf{w}_n)^{k-2} + \frac{1}{n^2} \sum_{k=2}^n \frac{\alpha(k)k}{(k-2)!} \left(\frac{1.584k}{\ln(k+1)} \right)^k,
 \end{aligned}$$

where the bound was attained by using Lemma 4 and the same techniques as in the proof of Lemma 2. Now, using the bound $f_0(x^j, \mathbf{w}_n)^l \leq 1$ for all $j, l \geq 1$, and (2.2), we obtain

$$h(x, \mathbf{w}_n) \leq \frac{1}{n} \left(\lambda + \sum_{k=2}^n \alpha(k) \left(\frac{1}{k!} + \frac{3}{(k-1)!} + \frac{2}{(k-2)!} \right) + \frac{B}{n} \right) \leq \frac{E}{n},$$

where E is independent of n and x , as required. □

Now we need to bound the expected value of the function $f_2(1, \mathbf{w}_n(t))$.

Lemma 4. Define the function $C_n : [0, \infty) \rightarrow \mathbb{R}$ as follows

$$C_n(t) = \mathbb{E}[f_2(1, \mathbf{w}_n(t))],$$

for $n \geq 1$ and $t \geq 0$. Then we have that

$$\sup_{n \in \mathbb{N}} C_n(t) \leq D,$$

where D is a constant independent of t .

Proof. We will use Gronwall’s inequality. We look at the derivative of C_n , using Dynkin’s formula we see that

$$\frac{d}{dt} C_n(t) = \mathbb{E}[\mathcal{A}_n f_2(1, \mathbf{w}_n(t))],$$

and hence, using the infinitesimal generator (1.2),

$$\begin{aligned}
 \frac{d}{dt} C_n(t) &= \mathbb{E} \left[-\lambda \sum_{k=1}^n (k^2 - k) w_{n,k}(t) \right. \\
 &\quad \left. + \sum_{k=2}^n \frac{\alpha(k)}{k!} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \left[\left(\sum_{i=1}^{|\pi|} |\pi_i| l_i \right)^2 - \sum_{i=1}^{|\pi|} (|\pi_i| l_i)^2 \right] \prod_{i=1}^{|\pi|} \frac{(n w_{n,l_i}(t))^{|\pi_i|}}{n^{|\pi_i|}} \right],
 \end{aligned}$$

which, using a similar argument to that used when bounding $\beta_n(x, t)$, gives that

$$\begin{aligned} \frac{d}{dt}C_n(t) &\leq \lambda - \lambda C_n(t) \\ &\quad + 2\mathbb{E}\left[\sum_{k=2}^n \frac{\alpha(k)}{k!} \sum_{l_1=1}^n \cdots \sum_{l_k=1}^n \left(\sum_{i<j} l_i l_j\right) w_{n,l_1}(t) \cdots w_{n,l_k}(t)\right] \\ &\quad + 2\mathbb{E}\left[\sum_{k=2}^n \frac{\alpha(k)}{k!} \frac{2^k(k-1)}{n} \sum_{\pi \in \mathcal{P}_k} \sum_{l_1=1}^n \cdots \sum_{l_{|\pi|=1}^n} \left(\sum_{i<j} |\pi_i| |\pi_j| l_i l_j\right) w_{n,l_1}(t) \cdots w_{n,l_{|\pi|}}(t)\right] \\ &\leq \lambda - \lambda C_n(t) + 2\mathbb{E}\left[\sum_{k=2}^n \frac{\alpha(k)}{k!} \binom{k}{2} G_n(1, t)^{k-2} + \frac{1}{n} \sum_{k=2}^n \frac{\alpha(k)k}{k!} \sum_{\pi \in \mathcal{P}_k} \binom{k}{2} G_n(1, t)^{|\pi|-2}\right] \\ &\leq \lambda - \lambda C_n(t) + \sum_{k=2}^n \frac{\alpha(k)}{(k-2)!} + \frac{B}{n}, \end{aligned}$$

where

$$B = \sum_{k=2}^{\infty} \frac{\alpha(k)k}{(k-2)!} \left(\frac{1.584k}{\ln(k+1)}\right)^k. \tag{2.2}$$

Thus,

$$\frac{d}{dt}C_n(t) \leq D_1 - \lambda C_n(t),$$

where D_1 is some constant not dependent on n or t , so applying Gronwall’s inequality gives the result as required. \square

With the help of the above three lemmas we are now ready to prove Theorem 1.

Proof of Theorem 1. Recall from (1.7) that we write $G(x, t)$ for the solution of the differential equation

$$\frac{\partial G}{\partial t}(x, t) = \lambda(x - G(x, t)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G(x, t)^k - kG(1, t)^{k-1}G(x, t)),$$

$$G(x, 0) = x.$$

The existence and uniqueness of $G(x, t)$ is straightforward to establish via the Picard–Lindelöf Theorem, as the right hand side is uniformly Lipschitz in G, x and t . Specifically, starting with the case $x = 1$, we have

$$\frac{\partial G(1, t)}{\partial t} = F_1(G(1, t)), \tag{2.3}$$

where

$$F_1(g) = \lambda(1 - g) + \sum_{k \geq 2} \frac{\alpha(k)}{k!} (1 - k)g^k. \tag{2.4}$$

The derivative of F_1 is uniformly bounded on $g \in [0, 1]$ since

$$F'_1(g) = -\lambda - \sum_{k \geq 2} \frac{\alpha(k)}{(k-2)!} g^{k-1}, \tag{2.5}$$

and thus

$$\sup_{g \in [0,1]} |F'_1(g)| = \lambda + \sum_{k \geq 2} \frac{\alpha(k)}{(k-2)!} = L_1 < \infty. \tag{2.6}$$

Said another way, F_1 is uniformly Lipschitz with constant L_1 , and thus (2.3) has a unique solution. A very similar argument using furthermore the well-posedness of (2.3) applies to the case of general $x \in (0, 1]$, where we have

$$\frac{\partial G(x, t)}{\partial t} = F_x(G(x, t), t), \tag{2.7}$$

for

$$F_x(g, t) = \lambda(x - g) + \sum_{k \geq 2} \frac{\alpha(k)}{k!} (g^k - kG(1, t)^{k-1}g). \tag{2.8}$$

The derivative is again bounded:

$$F'_x(g, t) = -\lambda + \sum_{k \geq 2} \frac{\alpha(k)}{(k-1)!} (g^{k-1} - G(1, t)^{k-1}), \tag{2.9}$$

and since $G(1, t) \in [0, 1]$ we obtain

$$\sup_{g \in [0,1], t \geq 0} |F'_x(g, t)| = \lambda + 2 \sum_{k \geq 2} \frac{\alpha(k)}{(k-1)!} = L < \infty. \tag{2.10}$$

As this constant is independent of x, g and t , global existence and uniqueness follows.

Moving on now to compare to the finite system generating function $G_n(x, t)$, we start by noting that in the trivial case $x = 0$, we have

$$G_n(0, t) = G(0, t) = 0, \quad \forall t \geq 0.$$

Now, for strictly positive x we define the function $Y_{n,x}(t) : [0, \infty) \rightarrow \mathbb{R}$ as follows:

$$Y_{n,x}(t) = \mathbb{E}[(G(x, t) - G_n(x, t))^2], \quad t \geq 0.$$

We will use Gronwall's inequality, so we look at the derivative of $Y_{n,x}(t)$ with respect to t , using Dynkin's formula (bearing in mind that pre-limit in n , the system is finite) we see that

$$\begin{aligned} \frac{dY_{n,x}}{dt}(t) &= \mathbb{E}[\mathcal{A}_n(G(x, t) - G_n(x, t))^2] \\ &= 2G(x, t) \frac{\partial}{\partial t} G(x, t) + \mathbb{E}[\mathcal{A}_n G_n^2(x, t)] - 2G(x, t) \mathbb{E}[\mathcal{A}_n G_n(x, t)] \\ &\quad - 2\mathbb{E}[G_n(x, t)] \frac{\partial}{\partial t} G(x, t). \end{aligned} \tag{2.11}$$

Noting that $G_n(x, t) = f_0(x, \mathbf{w}_n(t))$, we may use Lemma 2, Lemma 3, and the definition of G in (1.7) to see that

$$\begin{aligned} \frac{d}{dt} Y_{n,x}(t) &= 2G(x, t) \left(\lambda(x - G(x, t)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G(x, t)^k - kG(1, t)^{k-1}G(x, t)) \right) \\ &\quad + 2\mathbb{E} \left[G_n(x, t) \left(\lambda(x - G_n(x, t)) + \beta_n(x, \mathbf{w}_n(t)) \right) \right] \end{aligned}$$

$$\begin{aligned}
 & + \sum_{k=2}^n \frac{\alpha(k)}{k!} (G_n(x, t)^k - kG_n(1, t)^{k-1}G_n(x, t)) \Bigg] \\
 & - 2G(x, t)\mathbb{E} \left[\lambda(x - G_n(x, t)) + \beta_n(x, \mathbf{w}_n(t)) \right. \\
 & \left. + \sum_{k=2}^n \frac{\alpha(k)}{k!} (G_n(x, t)^k - kG_n(1, t)^{k-1}G_n(x, t)) \right] \\
 & - 2\mathbb{E}[G_n(x, t)] \left(\lambda(x - G(x, t)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G(x, t)^k - kG(1, t)^{k-1}G(x, t)) \right) \\
 & + \frac{\lambda x^2}{n} C_n(t) + \mathbb{E}[h(x, \mathbf{w}_n(t))] \\
 = & -2\lambda Y_{n,x}(t) + 2\mathbb{E}[(G_n(x, t) - G(x, t))\beta_n(x, \mathbf{w}_n(t))] + \frac{\lambda x^2}{n} C_n(t) + \mathbb{E}[h(x, \mathbf{w}_n(t))] \\
 & + 2 \sum_{k=2}^n \frac{\alpha(k)}{k!} \mathbb{E}[(G(x, t) - G_n(x, t))(G(x, t)^k - G_n(x, t)^k) \\
 & - k(G(1, t)^{k-1}G(x, t) - G_n(1, t)^{k-1}G_n(x, t))] \\
 & + 2 \sum_{k=n+1}^{\infty} \frac{\alpha(k)}{k!} \mathbb{E}[(G(x, t) - G_n(x, t))(G(x, t)^k - kG(1, t)^{k-1}G(x, t))]. \tag{2.12}
 \end{aligned}$$

□

Note that the right hand side of (2.12) depends on $G_n(1, t)$. For the boundary case $x = 1$, however, we have a closed expression in $G_n(1, t)$. The plan is thus to first show the theorem holds for $x = 1$ and use this to complete the proof for general x . Define $X_n(t) := Y_{n,1}(t)$.

Lemma 5. *We have*

$$\sup_{t \geq 0} X_n(t) \leq \frac{H_n}{2\lambda}, \tag{2.13}$$

where

$$H_n := \frac{1}{n} \left(2A + D\lambda + E + \sum_{k=2}^{\infty} \frac{2\alpha(k)}{(k-2)!} \right).$$

In particular, $H_n \rightarrow 0$ as $n \rightarrow \infty$ because of assumption (1.6).

Proof. Substituting $x = 1$ into (2.12) we see that

$$\begin{aligned}
 \frac{dX_n}{dt}(t) = & -2\lambda X_n(t) + \frac{\lambda}{n} C_n(t) + \mathbb{E}[h(1, \mathbf{w}_n(t))] + 2\mathbb{E}[(G_n(1, t) - G(1, t))\beta_n(1, \mathbf{w}_n(t))] \\
 & + 2 \sum_{k=2}^n \frac{\alpha(k)}{k!} \mathbb{E}[(G(1, t) - G_n(1, t))(G(1, t)^k - G_n(1, t)^k)](1 - k) \\
 & + 2 \sum_{k=n+1}^{\infty} \frac{\alpha(k)}{k!} \mathbb{E}[(G(1, t) - G_n(1, t))G(1, t)^k](1 - k).
 \end{aligned}$$

Now, $(G(1, t) - G_n(1, t))(G(1, t)^k - G_n(1, t)^k) \geq 0$, which means the second line above is negative. Hence, together with Lemma 3 and Lemma 4, we have

$$\frac{dX_n}{dt}(t) \leq -2\lambda X_n(t) + \frac{1}{n}(D\lambda + E + 2A) + \sum_{k=n+1}^{\infty} \frac{2\alpha(k)}{(k-1)!} \leq -2\lambda X_n(t) + H_n, \quad (2.14)$$

so that applying Gronwall's inequality gives the result as required. \square

We can now use this bound to complete the proof of Theorem 1.

Proof of Theorem 1 (cont.). Continuing from (2.12), we have that

$$\frac{d}{dt}Y_{n,x}(t) \leq -2\lambda Y_{n,x}(t) + 2 \sum_{k=2}^{\infty} S_k + H_n, \quad (2.15)$$

where

$$S_k = \frac{\alpha(k)}{k!} \mathbb{E} \left[G(x, t)^{k-1} (G(x, t)^2 - G(x, t)G_n(x, t)) + G_n(x, t)^{k-1} (G_n(x, t)^2 - G(x, t)G_n(x, t)) \right. \\ \left. - kG(1, t)^{k-1} (G(x, t)^2 - G(x, t)G_n(x, t)) - kG_n(1, t)^{k-1} (G_n(x, t)^2 - G(x, t)G_n(x, t)) \right].$$

We will apply Lemma 5 to bound the sum of the S_k . First, it is necessary to bound the sum in terms of $\mathbb{E}[|G(1, t) - G_n(1, t)|]$, and remove any terms involving just $G(x, t)$ and $G_n(x, t)$. To do this we create terms that contain the positive term $(G(x, t) - G_n(x, t))^2$, so that if we pre-multiply them by something negative we can discard it for an upper bound. For example, looking at the first term above we can do the following by adding a zero

$$G(x, t)^{k-1} \mathbb{E} [G(x, t)^2 - G(x, t)G_n(x, t)] = G(x, t)^{k-1} \mathbb{E} [G(x, t)^2 - G(x, t)G_n(x, t)] \\ - G(x, t)^{k-1} \mathbb{E} [G(x, t)G_n(x, t) - G_n(x, t)^2] \\ + G(x, t)^{k-1} \mathbb{E} [G(x, t)G_n(x, t) - G_n(x, t)^2] \\ = G(x, t)^{k-1} \mathbb{E} [(G(x, t) - G_n(x, t))^2] \\ + G(x, t)^{k-1} \mathbb{E} [G(x, t)G_n(x, t) - G_n(x, t)^2] \\ = G(x, t)^{k-1} \mathbb{E} [(G(x, t) - G_n(x, t))^2] \\ + \mathbb{E} [G(x, t)^{k-2} G_n(x, t) (G(x, t)^2 - G(x, t)G_n(x, t))].$$

In creating the square term above, we get a similar term to what we started with but with the exponent of $G(x, t)$ decreased by one and the exponent of $G_n(x, t)$ increased by one. We repeat this process until the exponent of $G_n(x, t)$ is $k-1$. We then subtract and add back in the same terms but with x replaced by 1. Specifically,

$$S_k = \frac{\alpha(k)}{k!} \mathbb{E} [G(x, t)^{k-1} (G(x, t) - G_n(x, t))^2 + G(x, t)^{k-2} G_n(x, t) (G(x, t) - G_n(x, t))^2] \\ + \cdots + G(x, t)G_n(x, t)^{k-2} (G(x, t) - G_n(x, t))^2 + G_n(x, t)^{k-1} (G(x, t) - G_n(x, t))^2 \\ - \frac{\alpha(k)}{k!} \mathbb{E} [G(1, t)^{k-1} (G(x, t) - G_n(x, t))^2 + G(1, t)^{k-2} G_n(1, t) (G(x, t) - G_n(x, t))^2] \\ + \cdots + G(1, t)G_n(1, t)^{k-2} (G(x, t) - G_n(x, t))^2 + G_n(1, t)^{k-1} (G(x, t) - G_n(x, t))^2 \\ + \frac{\alpha(k)}{k!} \mathbb{E} [G(1, t)^{k-1} (G(x, t) - G_n(x, t))^2 + G(1, t)^{k-2} G_n(1, t) (G(x, t) - G_n(x, t))^2]$$

$$\begin{aligned}
& + \cdots + G(1, t)G_n(1, t)^{k-2}(G(x, t) - G_n(x, t))^2 + G_n(1, t)^{k-1}(G(x, t) - G_n(x, t))^2] \\
& - \frac{\alpha(k)}{k!}(kG(1, t)^{k-1}G(x, t)^2 - kG(1, t)^{k-1}G(x, t)\mathbb{E}[G_n(x, t)] \\
& + k\mathbb{E}[G_n(1, t)^{k-1}G_n(x, t)^2] - kG(x, t)\mathbb{E}[G_n(1, t)^{k-1}G_n(x, t)]),
\end{aligned}$$

noting that the first four lines combined give something negative. As $G(x, t) \leq G(1, t)$ and $G_n(x, t) \leq G_n(1, t)$, we have

$$\begin{aligned}
S_k & \leq \frac{\alpha(k)}{k!}\mathbb{E}[G(1, t)^{k-1}(G(x, t)^2 - 2G(x, t)G_n(x, t) + G_n(x, t)^2)] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[G(1, t)^{k-2}G_n(1, t)(G(x, t)^2 - 2G(x, t)G_n(x, t) + G_n(x, t)^2)] \\
& + \cdots + \frac{\alpha(k)}{k!}\mathbb{E}[G(1, t)G_n(1, t)^{k-2}(G(x, t)^2 - 2G(x, t)G_n(x, t) + G_n(x, t)^2)] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[G_n(1, t)^{k-1}(G(x, t)^2 - 2G(x, t)G_n(x, t) + G_n(x, t)^2)] \\
& - \frac{\alpha(k)}{k!}(kG(1, t)^{k-1}G(x, t)^2 - kG(1, t)^{k-1}G(x, t)\mathbb{E}[G_n(x, t)] \\
& + k\mathbb{E}[G_n(1, t)^{k-1}G_n(x, t)^2] - kG(x, t)\mathbb{E}[G_n(1, t)^{k-1}G_n(x, t)]).
\end{aligned}$$

We gather terms in $G(x, t)^2$, $G_n(x, t)^2$ and $G(x, t)G_n(x, t)$, each multiplied by $|G(1, t)^j - G_n(1, t)^j|$ for some j . Here we have taken absolute values in order to not worry about signs. We do this by matching each of the terms from the first four lines of the above with the corresponding term from the last two lines. Doing this we see that

$$\begin{aligned}
S_k & \leq \frac{\alpha(k)}{k!}G(x, t)^2\mathbb{E}[G(1, t)^{k-2}|G(1, t) - G_n(1, t)| + \cdots + |G(1, t)^{k-1} - G_n(1, t)^{k-1}|] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[G_n(x, t)^2(|G(1, t)^{k-1} - G_n(1, t)^{k-1}| + \cdots + G_n(1, t)^{k-2}|G(1, t) - G_n(1, t)|)] \\
& + \frac{\alpha(k)}{k!}G(x, t)\mathbb{E}[G_n(x, t)(G(1, t)^{k-2}|G(1, t) - G_n(1, t)| + \cdots \\
& + G(1, t)|G(1, t)^{k-2} - G_n(1, t)^{k-2}|)] \\
& + \frac{\alpha(k)}{k!}G(x, t)\mathbb{E}[G_n(x, t)(G_n(1, t)|G(1, t)^{k-2} - G_n(1, t)^{k-2}| + \cdots \\
& + G_n(1, t)^{k-2}|G(1, t) - G_n(1, t)|)].
\end{aligned}$$

Hence, using the fact that $G(x, t), G(1, t), G_n(x, t), G_n(1, t) \leq 1$, we see that

$$\begin{aligned}
S_k & \leq \frac{\alpha(k)}{k!}\mathbb{E}[|G(1, t)^{k-1} - G_n(1, t)^{k-1}| + \cdots + |G(1, t) - G_n(1, t)|] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[(|G(1, t)^{k-1} - G_n(1, t)^{k-1}| + \cdots + |G(1, t) - G_n(1, t)|)] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[|G(1, t)^{k-2} - G_n(1, t)^{k-2}| + \cdots + |G(1, t) - G_n(1, t)|] \\
& + \frac{\alpha(k)}{k!}\mathbb{E}[|G(1, t)^{k-2} - G_n(1, t)^{k-2}| + \cdots + |G(1, t) - G_n(1, t)|] \\
& = 2\frac{\alpha(k)}{k!}\sum_{j=1}^{k-1}\mathbb{E}[|G(1, t)^j - G_n(1, t)^j|] + 2\frac{\alpha(k)}{k!}\sum_{j=1}^{k-2}\mathbb{E}[|G(1, t)^j - G_n(1, t)^j|],
\end{aligned}$$

$$S_k \leq 2 \frac{\alpha(k)}{k!} \sum_{j=1}^{k-1} j \mathbb{E}[|G(1, t) - G_n(1, t)|] + 2 \frac{\alpha(k)}{k!} \sum_{j=1}^{k-2} j \mathbb{E}[|G(1, t) - G_n(1, t)|],$$

where the above inequality uses $|y^j - z^j| \leq j|y - z|$ for $y, z \leq 1$. Thus

$$S_k \leq \frac{2\alpha(k)(k-1)}{k(k-2)!} \mathbb{E}[|G(1, t) - G_n(1, t)|].$$

We can plug this result back into (2.15) and see that

$$\frac{d}{dt} Y_{n,x}(t) \leq -2\lambda Y_{n,x}(t) + 4\mathbb{E}[|G(1, t) - G_n(1, t)|] \sum_{k=2}^n \frac{\alpha(k)(k-1)}{k(k-2)!} + H_n.$$

Now using Lemma 5, we have that

$$\frac{d}{dt} Y_{n,x}(t) \leq -2\lambda Y_{n,x}(t) + 4 \frac{\sqrt{H_n}}{\sqrt{\lambda}} \sum_{k=2}^n \frac{\alpha(k)(k-1)}{k(k-2)!} + H_n. \tag{2.16}$$

Then, we can use Gronwall’s inequality to see that

$$\begin{aligned} Y_{n,x}(t) &\leq \left(2 \frac{\sqrt{H_n}}{\lambda\sqrt{\lambda}} \sum_{k=2}^n \frac{\alpha(k)(k-1)}{k(k-2)!} + H_n \right) (1 - e^{-2\lambda t}) \\ &\leq 2 \frac{\sqrt{H_n}}{\lambda\sqrt{\lambda}} \sum_{k=2}^n \frac{\alpha(k)(k-1)}{k(k-2)!} + H_n \\ &\rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$, thanks to (1.6). This convergence is uniform in both x and t , as required. □

From Theorem 1 the Smoluchowski-type equations stated in Corollary 1 follow immediately:

Proof of Corollary 1. We begin by substituting $G(x, t) = \sum_j x^j w_j(t)$ into (1.7);

$$\begin{aligned} \frac{\partial G}{\partial t}(x, t) &= \lambda(x - G(x, t)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G(x, t)^k - kG(x, t)G(1, t)^{k-1}) \\ &= \lambda \left(x - \sum_{j=1}^{\infty} x^j w_j(t) \right) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} \left(\left(\sum_{j=1}^{\infty} x^j w_j(t) \right)^k - k \left(\sum_{j=1}^{\infty} x^j w_j(t) \right) \left(\sum_{j=1}^{\infty} w_j(t) \right)^{k-1} \right). \end{aligned}$$

Rearranging the right hand side to collect powers of x we obtain

$$\begin{aligned} \frac{\partial G}{\partial t}(x, t) &= x \left(\lambda(1 - w_1(t)) - w_1(t) \sum_{k=1}^{\infty} \frac{\alpha(k)}{(k-1)!} \left(\sum_{j=1}^{\infty} w_j(t) \right)^{k-1} \right) \\ &\quad + \sum_{j=2}^{\infty} x^j \left(-\lambda w_j(t) - w_j(t) \sum_{k=1}^{\infty} \frac{\alpha(k)}{(k-1)!} \left(\sum_{l=1}^{\infty} w_l(t) \right)^{k-1} \right) \\ &\quad + \sum_{k=2}^j \frac{\alpha(k)}{k!} \sum_{l_1+\dots+l_k=j} w_{l_1}(t) \cdots w_{l_k}(t), \end{aligned}$$

as required, where the final term on the right-hand side is obtained using the product formula for power series. □

3. Universal stationary distribution

Now that we know how the empirical generating function $G_n(x, t)$ converges in the large n limit, we can extract a prediction for the stationary distribution of cluster sizes. Recalling from (1.5) that the generating function of the cluster size distribution is obtained via

$$g_n(x, t) = \frac{G_n(x, t)}{G_n(1, t)},$$

we reintroduce (and later prove the existence of) the quantities

$$g_\lambda(x, t) = \frac{G(x, t)}{G(1, t)}, \quad G_\lambda(x) = \lim_{t \rightarrow \infty} G(x, t),$$

$$g_\lambda(x) = \lim_{t \rightarrow \infty} g_\lambda(x, t) = \frac{G_\lambda(x)}{G_\lambda(1)}, \quad p_k(\lambda) = \lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} p_{n,k}(t).$$

Note: as this section focuses on the limit as $\lambda \rightarrow 0$ of the above quantities, these quantities have been labelled with λ to make it easier to follow. Under self-averaging, the stationary cluster size distribution is found simply by determining the fixed point $G_\lambda(x)$ of the differential equation (1.7) and developing $g_\lambda(x) = G_\lambda(x)/G_\lambda(1)$ as a power series. In practice, this procedure is only tractable in the limit $\lambda \searrow 0$, where we find a particular limiting distribution.

Proof of the convergence of G as $t \rightarrow \infty$ is an application of the Poincaré–Bendixson Theorem (see [21], p. 223), which states that attractors of plane autonomous systems are either fixed points, collections of fixed points (with their associated linking orbits), or limit cycles. To compute the asymptotic form of the solution in small λ we employ the classical technique of series inversion that dates back to Newton, Lagrange, Bürmann and Puiseux. In particular, the following theorem is a sub-case of results of those authors (see [9], p. 183):

Theorem 3 (Series inversion). *Let $f : [0, \infty) \rightarrow [0, \infty)$ be defined by the power series $f(x) = \sum_{k \geq m} a_k x^k$ for integer $m \geq 1$ and non-negative real constants $\{a_k\}$. The inverse is expressed via the Puiseux series*

$$f^{-1}(y) = \sum_{k=1}^{\infty} b_k y^{k/m},$$

where

$$b_k = \frac{1}{k!} \lim_{z \rightarrow 0} \left[\frac{d^{k-1}}{dz^{k-1}} \left(\frac{z}{f(z)^{1/m}} \right)^k \right].$$

In particular, note that $b_1 = 1/a_1$.

Proof of Theorem 2. Recall that the limit generating function $G(x, t)$ obeys (1.7). Equivalently we may write $\partial G(x, t)/\partial t = F_x(G(x, t), t)$, where F_x is the function defined in (2.8). Since $G(1, t) \geq G(x, t)$ for all x , it follows that F_x is monotonically decreasing in its first argument, moreover, $F_x(1, t) < 0 < F_x(0, t)$. Hence, in the particular case $x = 1$, $G(1, t)$ has a unique fixed point, denoted $G_\lambda(1)$. As $G(1, t) \in [0, 1]$ for all t convergence is assured.

For general x , note that $(G(1, t), G(x, t))$ is a plane autonomous system and, by the same argument as above, has a unique fixed point. From the Poincaré–Bendixson theorem, we conclude that we need only to rule out the case of periodic orbits to confirm convergence to the fixed point. But we have already shown that $G(1, t)$ converges to its unique fixed point $G_\lambda(1)$, so the ω -limit set of the plane system must be contained in the line $(G_\lambda(1), \cdot)$. Thus, by the Jordan Curve theorem (see [21], p. 220), the ω -limit set cannot be a regular periodic orbit. Therefore, we may conclude that the solution of (1.7) is required to converge to a stationary value $G(x, t) \rightarrow G_\lambda(x)$, where

$$0 = \lambda(x - G_\lambda(x)) + \sum_{k=2}^{\infty} \frac{\alpha(k)}{k!} (G_\lambda(x)^k - kG_\lambda(1)^{k-1}G_\lambda(x)). \tag{3.1}$$

At the point $x = 1$ we rearrange to find λ as a power series in $G_\lambda(1)$:

$$\lambda = \frac{1}{1 - G_\lambda(1)} \sum_{k \geq m} \frac{\alpha(k)}{k!} (k - 1) G_\lambda(1)^k = \frac{\alpha(m)}{m!} (m - 1) G_\lambda(1)^m + \sum_{k > m} c_k G_\lambda(1)^k, \tag{3.2}$$

where m is the least integer such that $\alpha(m) > 0$, and c_k are the appropriate constants. From Theorem 3 we thus determine that

$$G_\lambda(1) = \lambda^{1/m} \left(\frac{m!}{(m - 1)\alpha(m)} \right)^{1/m} + o(\lambda^{1/m}).$$

For general x , equation (3.1) implies the same leading order behaviour of $G_\lambda(1)$ in λ . Recalling the definition $g_\lambda(x) = G_\lambda(x)/G_\lambda(1)$ at the start of this section, we see that

$$\begin{aligned} 0 &= \lambda(x - G_\lambda(1)g_\lambda(x)) + \sum_{k=m}^{\infty} \frac{\alpha(k)}{k!} G_\lambda(1)^k (g_\lambda(x)^k - kg_\lambda(x)) \\ &= \lambda x - \left(\frac{m!}{(m - 1)\alpha(m)} \right)^{1/m} \lambda^{1+1/m} g_\lambda(x) \\ &\quad + \sum_{k=m}^{\infty} \frac{\alpha(k)}{k!} \left(\frac{m!}{(m - 1)\alpha(m)} \right)^{k/m} \lambda^{k/m} (g_\lambda(x)^k - kg_\lambda(x)) + o(\lambda) \\ &= \lambda x + \frac{\lambda}{m - 1} (g_\lambda(x)^m - mg_\lambda(x)) + o(\lambda), \quad x \in [0, 1]. \end{aligned}$$

In other words, when m is the smallest integer such that α is non-zero, we find that in the limit as $\lambda \searrow 0$,

$$mg_\lambda(x) - g_\lambda(x)^m - x(m - 1) = 0.$$

Note that this limit exists as g_λ is a Puiseux series in λ . Applying the inversion theorem once more, we obtain the series

$$g_\lambda(x) = \sum_{k=1}^{\infty} \frac{x^k}{k!} \lim_{z \rightarrow 0} \left[\frac{d^{k-1}}{dz^{k-1}} \left(\frac{z}{f(z)} \right)^k \right], \tag{3.3}$$

where $f(z) = m/(m - 1)z - 1/(m - 1)z^m$. We compute

$$\begin{aligned} \left(\frac{z}{f(z)} \right)^k &= \left(\frac{m - 1}{m} \right)^k \left(\frac{1}{1 - z^{m-1}/m} \right)^k \\ &= \left(\frac{m - 1}{m} \right)^k \left(\sum_{n=0}^{\infty} \left(\frac{1}{m} \right)^n z^{n(m-1)} \right)^k \\ &= \left(\frac{m - 1}{m} \right)^k \sum_{n=0}^{\infty} \left(\frac{1}{m} \right)^n \binom{k - 1 + n}{n} z^{n(m-1)}, \end{aligned}$$

using the product formula for series. To get the summands for (3.3) we need to look at the $(k - 1)$ th derivative of the above series, evaluated at $z = 0$. Hence we have that

$$\frac{d^{k-1}}{dz^{k-1}} \left(\frac{z}{f(z)} \right)^k \Big|_{z=0} = \begin{cases} \left(\frac{m-1}{m} \right)^k \left(\frac{1}{m} \right)^{\frac{k-1}{m-1}} \binom{m(\frac{k-1}{m-1})}{\frac{k-1}{m-1}} (k - 1)! & \text{if } m - 1 \text{ divides } k - 1, \\ 0 & \text{otherwise.} \end{cases}$$

Plugging this into (3.3) and comparing coefficients of x^k , we have

$$\lim_{\lambda \searrow 0} p_k(\lambda) = \begin{cases} \frac{1}{k} \left(\frac{m-1}{m}\right)^k \left(\frac{1}{m}\right)^{\frac{k-1}{m-1}} \binom{m \left(\frac{k-1}{m-1}\right)}{\frac{k-1}{m-1}} & \text{if } m-1 \text{ divides } k-1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, we can conclude that, for large values of k , under these conditions,

$$\lim_{\lambda \searrow 0} p_k(\lambda) \sim k^{-3/2},$$

regardless of the values of $\alpha(k)$, $k \geq 2$. □

Acknowledgements

We would like to thank Jason Schweinsberg for providing an intuitive explanation for the absence of even-sized clusters in the example of Figure 1. TR gratefully acknowledges support from the Royal Society.

References

- [1] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** (1) (1999) 3–48. [MR1673235](#)
- [2] D. Berend and T. Tassa. Improved bounds on Bell numbers and on moments of sums of random variables. *Probab. Math. Statist.* **30** (2) (2010) 185–205. [MR2792580](#)
- [3] J. Berestycki. Exchangeable fragmentation-coalescence processes and their equilibrium measures. *Electron. J. Probab.* **9** (25) (2004) 770–824. [MR2110018](#)
- [4] J. C. Bohorquez, S. Gourley, A. R. Dixon, M. Spagat and N. F. Johnson. Common ecology quantifies human insurgency. *Nature* **462** (7275) (2009) 911–914.
- [5] X. Bressaud and N. Fournier. A mean-field forest-fire model. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** (2) (2014) 589–614. [MR3296482](#)
- [6] M. Eibl, E. Hendriks, J. Spouge and M. Schreckenberg. Exact solutions for random coagulation processes. *Physik Z.* **58** (3) (1985) 219–227.
- [7] S. Gueron. The steady-state distributions of coagulation-fragmentation processes. *J. Math. Biol.* **37** (1) (1998) 1–27. [MR1636636](#)
- [8] S. Gueron and S. A. Levin. The dynamics of group formation. *Math. Biosci.* **128** (1) (1995) 243–264.
- [9] M. Hazewinkel. *Encyclopaedia of Mathematics, Vol. 5*. Springer, Netherlands, 1990.
- [10] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Adv. Phys.* **49** (7) (2000) 815–958. [MR2246565](#)
- [11] I. Jeon. Existence of gelling solutions for coagulation-fragmentation equations. *Comm. Math. Phys.* **194** (3) (1998) 541–567. [MR1631473](#)
- [12] K. Koutzenogii, A. Levykin and K. Sabelfeld. Kinetics of aerosol formation in the free molecule regime in presence of condensable vapor. *J. Aerosol Sci.* **27** (5) (1996) 665–679.
- [13] A. Lambert. The branching process with logistic growth. *Ann. Appl. Probab.* **15** (2) (2005) 1506–1535. [MR2134113](#)
- [14] A. Lushnikov. Certain new aspects of the coagulation theory. *Izv. An. Fiz. Atmos. Phys.* **14** (10) (1978) 738–743.
- [15] A. Lushnikov. Coagulation in finite systems. *J. Colloid. Interf. Sci.* **65** (2) (1978) 276–285.
- [16] A. H. Marcus. Stochastic coalescence. *Technometrics* **10** (1968) 133–143. [MR0223151](#)
- [17] B. Ráth and B. Tóth. Erdős–Rényi random graphs + forest fires = self-organized criticality. *Electron. J. Probab.* **14** (45) (2009) 1290–1327. [MR2511285](#)
- [18] J. Seinfeld. *Atmospheric Chemistry and Physics of Air Pollution*. Wiley, New York, 1986.
- [19] M. Smoluchowski. Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen. *Physik Z.* **17** (1916) 557–585.
- [20] S. Tavaré. Line-of-descent and genealogical processes, and their applications in population genetics models. *Theor. Popul. Biol.* **26** (2) (1984) 119–164. [MR0770050](#)
- [21] G. Teschl. *Ordinary Differential Equations and Dynamical Systems. Graduate Studies in Mathematics* **140**. American Mathematical Society, Providence, RI, 2012. [MR2961944](#)
- [22] F. Verhulst. *Nonlinear Differential Equations and Dynamical Systems. Universitext*. Springer, Berlin, 1990. Translated from the Dutch. [MR1036522](#)