Gelation in cluster coagulation processes

Luisa Andreis[1], Tejas Iyer[2], Elena Magnanini [2]

submitted: August 23, 2023 (revision: March 3, 2025)

[1] Politecnico di Milano
Dipartimento di Matematica
Piazza Leonardo da Vinci 32
20133 Milano
Italy
E-Mail: luisa.andreis@polimi.it

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: tejas.iyer@wias-berlin.de
        elena.magnanini@wias-berlin.de

Gelation in cluster coagulation processes

Luisa Andreis, Tejas Iyer, Elena Magnanini

Abstract

We consider the problem of gelation in the cluster coagulation model introduced by Norris [*Comm. Math. Phys.*, 209(2):407-435 (2000)], where pairs of clusters of types (x, y) taking values in a measure space E, merge to form a new particle of type $z \in E$ according to a transition kernel $K(x, y, \mathrm{d}z)$. This model possesses enough generality to accommodate inhomogeneities in the evolution of clusters, including variations in their shape or spatial distribution. We derive general, sufficient criteria for stochastic gelation in this model. As particular cases, we extend results related to the classical Marcus–Lushnikov coagulation process, showing that reasonable 'homogenous' coagulation processes with exponent $\gamma > 1$ yield gelation; and also, coagulation processes with kernel $\bar{K}(m, n) \geq (m \wedge n) \log (m \wedge n)^{3+\epsilon}$ for $\epsilon > 0$.

1 Introduction

Models of coagulation arise widely in many scientific models, in areas ranging from physical chemistry (in the formation of polymers), to astrophysics (in modelling the formation of galaxies). A classical model for coagulation involves collections of particles (which we refer to later on, formally, as 'clusters'), each attributed with a mass. Pairs of particles coagulate to form new particles at a rate $\bar{K}(x, y)$, where x and y are *masses* of the particles. This stochastic coagulation model is known as the *Marcus–Lushnikov process* [25, 15, 24]. The limiting behaviour of the particle masses as the number of particles tends to infinity in the Marcus–Lushnikov model is generally expected to be encoded by a set of infinitely many differential equations (or measure-valued differential equations) known as the *Smoluchowski* or *Flory* equations.

Particular cases of the Marcus–Lushnikov model are closely related to other stochastic models: the case $\bar{K}(x, y) = 1$ corresponds to the Kingman's coalescent [22], the case $\bar{K}(x, y) = x + y$ has multiple interpretations, including being related to Aldous' continuum random tree [1, 4, 10], whilst the case $\bar{K}(x, y) = xy$ is closely related to the Erdős–Rényi random graph (see, for example, [20, 3, 2, 9]). We refer the reader to the review paper [5] for a more general overview, although remark that there has been a lot of progress made over the last 25 years. We also note that in this paper, we are interested in models of *pure coagulation*; whilst a lot of work in the literature also allows for the *fragmentation* of particles.

A natural question of interest related to coagulation processes is whether or not at some time $t > 0$ there is the formation of *macroscopic* or *giant* particles (*gels*), i.e. particles whose masses are on a significantly larger scale than the ones of the initial particles in the system. Motivated by the application to polymer chemistry, this is known as *gelation*. Gelation is generally defined as whether a solution of the Smoluchowski (or Flory) equation fails to 'conserve mass', which means, intuitively, that mass is lost to 'infinite-mass' particles. This is

closely linked to the appearance of 'large particles' in the Marcus–Lushnikov model as long as one knows that the trajectories of the process concentrate on solution(s) of such an equation. This fact was first used by Jeon to prove existence of gelling solutions to the Smoluchowski equation [21] whenever there exists $\alpha > 1/2$ such that $\bar{K}(x,y) \geq (xy)^\alpha$ (see also [29] for a simplified exposition). An alternative proof of gelation using analytic tools was provided by Escobedo, Mischler and Perthame [14] (see also [13]), whilst Laurencot [23] improved this to show that gelation occurs whenever $\bar{K}(x,y) \geq \sqrt{xy}(\log(x)\log(y))^{1+\epsilon}$, for $\epsilon > 0$. However, a longstanding scientific conjecture that has not previously been proven rigorously states that if $\bar{K}$ is *homogenous* with exponent $\gamma > 1$ (i.e., $\bar{K}(cx, cy) = c^\gamma \bar{K}(x, y)$ for $c > 0$) then gelation occurs (see, e.g. [5, 12, 32]).

Although an extensive literature is devoted to the classical Marcus–Lushnikov process, relatively less is known about variants of this model incorporating inhomogenieties of particles, for example, their shape, velocity, or location in space. A framework introduced by Norris in [26], called the *cluster coagulation model*, allows one to incorporate these features in a rather general way. In [26] a weak law of large numbers has also been proved, recently extended to weaker assumptions in [7]. See also [27] for a variant that incorporates diffusion of clusters. However, apart from a particular special case of the model [18], general criteria for gelation in this model that incorporate information about these inhomogeneities are lacking, despite being of interest from both the perspective of applications in physics, and mathematically.

It should be noted that whilst this model is rather general, and can incorporate spatial characteristics in the clusters, it lacks an important feature: *movement of particles in space* independent of coagulation events. There are a number of results related to other models incorporating the movement of particles as Brownian motions in space [16, 17, 34, 30], or as particles jumping across two sites [31, 33]; but we are not aware of any general results concerning criteria for gelation in these models, apart from the interesting *induced gelation* effect in a particular case of the two-site model [12]. Recently, in [8], the question of gelation in a spatial coagulation model has been approached using a different approach, with Poisson point processes and large deviations.

1.1   Overview on our contribution

In this paper we make a contribution to the aforementioned gap in the literature by providing a general sufficient criterion for gelation in the cluster coagulation model.

1 In Theorem 2.4, we provide sufficient criteria for gelation in the general setting of the cluster coagulation model. These criteria may be of interest in applied settings, where the kernels may take into account many features of a cluster, not only the mass.

2 As a consequence, in Corollary 2.8 we provide improved criteria for gelation in the classical Marcus–Lushnikov process. We show, as a particular case, that reasonable 'homogenous' coagulation processes with exponent $\gamma > 1$ yield gelation, thus, providing, as far as we now, the first rigorous proof of the aforementioned generally accepted scientific principle. We also show, that, if for $\epsilon > 0$, $\bar{K}(x, y) \geq (x \wedge y)(\log(x \wedge y))^{3+\epsilon}$, gelation occurs - a condition that depends only on the size of the minimum of merging clusters.

To prove the gelation statements, we apply techniques introduced by Jeon [21] (see also [29]), exploiting the Markovian dynamics and the generator of the process to bound the expected

value of the random time when a positive fraction of the mass of the system is made up of 'large' clusters. In the general spatial setting, we use a coarse-graining procedure, partitioning the space into regions where the interaction rates between clusters have sufficient lower bounds. The summability condition in Assumption 2.1 ensures large clusters form 'quickly' enough for gelation to arise. When the state space also incorporates, for example, the position of the clusters, this criterion guarantees gelation for kernels that are products of a function decreasing in the distance between clusters, and of a 'gelling' kernel of the masses. The improved sufficient criteria for gelation in the non-spatial Marcus-Lushnikov model, outlined in Item 2 above, are a result of finer lower bounds on the generator compared to the bounds in [21, 29].

The rest of the paper is structured as follows.

1 In Section 1.2 we introduce the cluster coagulation model and we define gelation in this context. In Section 1.3 we provide examples of natural models that fit into the framework of cluster coagulation processes.

2 In Section 2 we state our results. In particular, in Section 2.1 we state our sufficient criteria for gelation, the main theorem of this paper. In Corollary 2.8 we state how this criteria translate into the setting of classical, non-spatial, coagulation processes, improving the state of the art in this setting. In Section 2.2 we mention a connection with inhomogeneous graphs and how to derive some gelation criteria from them. Such a connection is based on a natural coupling, which we omit to prove here (we refer to an extended pre-print version of this work for it [6]).

3 Finally, Section 3 deals with the proofs of the main results.

1.2 Definition of the process and gelation

We consider the *cluster coagulation process*, introduced by Norris in [26], where *clusters* are characterised by their *type*, an element of a measurable space $(E, \mathcal{B})$. Associated with a cluster of type $x \in E$ is a *mass function* $m : E \to (0, \infty)$. Another important quantity associated with the process is a *coagulation kernel* $K : E \times E \times \mathcal{B} \to [0, \infty)$, which satisfies the following:

1 for all $A \in \mathcal{B}$ $(x, y) \mapsto K(x, y, A)$ is measurable;

2 for all $x, y \in E$ $K(x, y, \cdot)$ is a measure on E;

3 *symmetric:* for all $A \in \mathcal{B}, x, y \in E$ $K(x, y, A) = K(y, x, A)$;

4 K is *finite:* for all $x, y \in E$ $\bar{K}(x, y) := K(x, y, E) < \infty$;

5 K *preserves mass:* for all $x, y \in E$, $m(z) = m(x) + m(y)$ for $K(x, y, \cdot)$-a.a. $z \in E$.

Suppose that we begin with a configuration of clusters labelled by an index set I. Then,

■ to each labelled pair of clusters, independently, with types $x, y \in E$ and $\bar{K}(x, y) > 0$ we associate an exponential clock (exponential random variable) with parameter $\bar{K}(x, y)$;
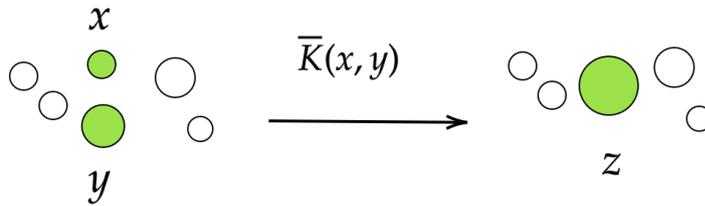
Figure 1: Two clusters of types x and y merge at rate $\bar{K}(x, y)$ into a particle of type z, sampled from the probability measure $K(x, y, \mathrm{d}z)/\bar{K}(x, y)$, with mass $m(z) = m(x) + m(y)$.

- upon the elapsure of the next exponential random variable in the process, corresponding to a labelled pair with types $x, y \in E$, say, the associated clusters are removed and replaced with a new labelled cluster with type $z \in E$, where z is sampled from the probability measure

$$\frac{K(x, y, \cdot)}{\bar{K}(x, y)}.$$

We refer the reader to Section 1.3 for natural examples of coagulation models that fit into this framework. Figure 1 illustrates the coagulation mechanism.

Throughout this paper we consider sequences of cluster coagulation processes, depending on a parameter $N \in \mathbb{N}$, that is for each process of the sequence we have a coagulation kernel that depends on the parameter (we denote it with K_N). One may consider this parameter N, up to random fluctuations, as the total *initial mass* of the system. We analyse the process in limiting regimes as $N \to \infty$. For each $N \in \mathbb{N}$, we consider the process as a measure valued Markov process $(\mathbf{L}_t^{(N)})_{t \geq 0}$ on the space of *finite point measures* on E, i.e., the space of positive, integer valued, finite measures on E. The point measure $\mathbf{L}_t^{(N)}$ encodes the configuration of clusters at time t, so that for $A \subseteq E$, $a \in (0, \infty)$, $\mathbf{L}_t^{(N)}(A \cap m^{-1}([a, \infty)))$ denotes the random number of clusters of mass at least a belonging to A at time t. Suppose $\mathcal{M}_+(E)$ denotes the set of finite, positive measures on E. Then, the infinitesimal generator $\mathcal{A}_N$ associated with the process is defined as follows: for any bounded measurable test function $F : \mathcal{M}_+(E) \to \mathbb{R}$, we have

$$\mathcal{A}_N F(\xi) = \frac{1}{2} \int_{E \times E \times E} \xi(\mathrm{d}x)\, (\xi - \delta_x)\, (\mathrm{d}y) K_N(x, y, \mathrm{d}z) \left(F(\xi^{(x,y) \to z}) - F(\xi) \right) \qquad \forall \xi \in \mathcal{M}_+(E),$$

(1)

where $\xi^{(x,y) \to z} := \xi + (\delta_z - \delta_x - \delta_y)$. Note that, the measure ξ in (1) is always non-negative, since, we always assume the process is initiated by a finite point measure. Thus, for any $x \in E$, by assumption $\xi(\{x\}) \in \mathbb{N}_0$ is a non-negative integer, representing the number of clusters of type x. The measure $\xi^{(x,y) \to z}$, therefore, describes the configuration of the system after a coagulation involving the two clusters $x, y \in E$ merging to form a cluster z with $m(z) = m(x) + m(y)$, for $K_N(x, y, \cdot)$-a.a. $z \in E$.

Remark 1.1. Abusing notation for brevity, if we write $\xi(\mathrm{d}x)$ for $\xi(\{x\})$, the factor $\frac{1}{2}$ in front of the generator in (1) ensures that the number of pairs that interact at rate $\bar{K}_N(x, y)$ is $\xi(\mathrm{d}x)\xi(\mathrm{d}y)$ if $x \neq y$ and $\left(\begin{smallmatrix} \xi(\mathrm{d}x)(\xi(\mathrm{d}x)-1) \\ 2 \end{smallmatrix} \right)$ if $x = y$.

Whilst we allow the kernel K_N to depend on N, we are often motivated by scenarios where $K_N \equiv K$ for some fixed kernel K. In such cases, it is well known that, in order to observe non-trivial limiting behaviour of the process (for example in proving a law of large numbers)

a time rescaling is needed. This allows one to counterbalance the increase in the number of interactions as the initial mass of clusters grows with N. This motivates the definition of the following normalised process:

$$\bar{\mathbf{L}}_t^{(N)} := \mathbf{L}_{t/N}^{(N)}/N.$$

Note that the normalised process is still a Markov process on $\mathcal{M}_+(E)$ with generator $\tilde{\mathcal{A}}_N$, defined such that for $\tilde{\xi} \in \mathcal{M}_+(E)$ and any bounded measurable test function $F : \mathcal{M}_+(E) \to \mathbb{R}$, we have

$$\tilde{\mathcal{A}}_N F(\tilde{\xi}) = \frac{N}{2} \int_{E \times E \times E} \tilde{\xi}(\mathrm{d}x) \left(\tilde{\xi} - \frac{\delta_x}{N} \right) (\mathrm{d}y) K_N(x, y, \mathrm{d}z) \left(F\left(\tilde{\xi} + \frac{(\delta_z - \delta_x - \delta_y)}{N} \right) - F(\tilde{\xi}) \right).$$

(2)

The scaling by N comes from the following. First, to get the total rate at which clusters associated with x and y interact, we need to multiply each of the terms $\tilde{\xi}(\mathrm{d}x)$ and $\left(\tilde{\xi} - \frac{\delta_x}{N} \right)(\mathrm{d}y)$ by a factor of N. Then, note that the time re-scaling $t \mapsto t/N$ corresponds to 'slowing down' each of the exponential clocks in the system by a factor of N. If X is exponentially distributed with rate K, NX is exponentially distributed with rate K/N. Therefore, we need to normalise the kernel by $K_N \mapsto K_N/N$. Combining these steps leads to the above. Given a measure $\mu \in \mathcal{M}_+(E)$ and a measurable function $f : E \to \mathbb{R}$, we denote by

$$\langle f, \mu \rangle := \int_E f(x)\mu(\mathrm{d}x).$$

At the level of the stochastic process, we denote by $\mathbb{P}_N(\cdot)$ and $\mathbb{E}_N[\cdot]$ probability distributions and expectations with regards to the trajectories of the process $(\bar{\mathbf{L}}_t^{(N)})_{t\geq0}$ with generator $\tilde{\mathcal{A}}_N$ and possibly random initial condition $\bar{\mathbf{L}}_0^{(N)}$. To ensure the process is well defined, we assume throughout that $N\bar{\mathbf{L}}_0^{(N)}$ is, almost surely, a point measure. In addition, we introduce the following notation for the regular conditional distribution and expectations when the initial condition $\bar{\mathbf{L}}_0^{(N)}$ is given by a (deterministic) measure $\pi^{(N)} \in \mathcal{M}_+(E)$

$$\mathbb{P}_{N, \pi^{(N)}}(\cdot) := \mathbb{P}_N\left(\cdot \,|\, \bar{\mathbf{L}}_0^{(N)} = \pi^{(N)} \right) \quad \text{and} \quad \mathbb{E}_{N, \pi^{(N)}}[\cdot] := \mathbb{E}_N\left[\cdot \,|\, \bar{\mathbf{L}}_0^{(N)} = \pi^{(N)} \right]. \quad (3)$$

Remark 1.2. As with $\bar{\mathbf{L}}_0^{(N)}$, we always only consider initial configurations $\pi^{(N)}$ such that $N\pi^{(N)}$ is a finite point measure.

For brevity, we generally refer to sequences of cluster coagulation processes $\left\{ (\bar{\mathbf{L}}_t^{(N)})_{t\geq0} \right\}_{N\in\mathbb{N}}$ as a cluster coagulation process.

Our main interest is the presence of *gelation* in the cluster coagulation process, indicating the emergence of 'large' clusters. The following is a minor reformulation of the definitions from Jeon [21, Definition 2]. First, it is helpful to define random times indicating the presence of these large cluster. Let $\psi : \mathbb{N} \to \mathbb{N}$ be given such that $\lim_{N\to\infty} \psi(N) = +\infty$, and let $\delta \in (0, 1)$. For each $N \in \mathbb{N}$ we define the (ψ, δ)-*gelation time* $\tau_N(\psi, \delta)$ by

$$\tau_N(\psi, \delta) := \inf \left\{ t \geq 0 : \left\langle m\mathbf{1}_{m\geq\psi(N)}, \bar{\mathbf{L}}_t^{(N)} \right\rangle \geq \delta \right\}. \quad (4)$$

Definition 1.3 (Stochastic gelation and strong gelation). Take a cluster coagulation process $\{(\bar{\mathbf{L}}_t^{(N)})_{t\geq0}\}_{N\in\mathbb{N}}$:

1 for a function $\psi : \mathbb{N} \to \mathbb{N}$ such that $\lim_{N \to \infty} \psi(N) = +\infty$ and $\delta > 0$, the (ψ, δ)-*stochastic gelation time* of the cluster coagulation process $(\bar{\mathbf{L}}^{(N)})_{N \in \mathbb{N}}$ is defined by

$$T_g^{\psi, \delta} := \inf \left\{ t \geq 0 : \limsup_{N \to \infty} \mathbb{P}_N \left(\tau_N(\psi, \delta) \leq t \right) > 0 \right\};$$

2 we define the *strong gelation time* of the cluster coagulation process by

$$\inf\{ t > 0 : \exists\, 0 < \alpha, \delta \leq 1 \text{ such that } \limsup_{N \to \infty} \mathbb{P}_N \left(\tau_N(\alpha N, \delta) \leq t \right) > 0\}.$$

If, for some $\psi : \mathbb{N} \to \mathbb{N}$ such that $\lim_{N \to \infty} \psi(N) = +\infty$ and $\delta > 0$ we have $T_g^{\psi, \delta} < \infty$ we say that *stochastic gelation occurs*. If the strong gelation time is finite, we say that *strong gelation occurs*.

Remark 1.4. Gelation is also often defined in terms of deterministic trajectories. More precisely, under suitable assumption of convergence of the sequence of kernels $\{K_N\}_N$, one expects the trajectories of $(\bar{\mathbf{L}}_t^{(N)})_{t \geq 0}$ to concentrate on solutions of the so called 'Smoluchowski' equation as $N \to \infty$, or better to its modification that includes the influence of gel, the so called 'Flory' equation, see [26, 7]. Suppose $(\bar{\mathbf{L}}_t^*)_{t \geq 0}$ is a trajectory solving this equation, with $\langle m, \bar{\mathbf{L}}_0^* \rangle < \infty$. One says $(\bar{\mathbf{L}}_t^*)_{t \geq 0}$ is a 'gelling solution' if for some $t, \varepsilon > 0$, $\langle m, \bar{\mathbf{L}}_t^* \rangle \leq \langle m, \bar{\mathbf{L}}_0^* \rangle - \varepsilon$. An equivalence principle by Jeon, which one may generalise to the cluster coagulation setting, shows that whenever $(\bar{\mathbf{L}}_t^{(N)})_{t \geq 0}$ displays this concentration, stochastic gelation is equivalent to the existence of gelling solutions to the associated equation (see [21, Theorem 5]).

1.3   Examples of cluster coagulation processes

The cluster coagulation process is general enough to encompass a large number of examples, depending on particular choices of the space E. In the following examples we fix $K_N \equiv K$ for some kernel K.

Example 1 (Classical kernel). If $E = (0, \infty)$, $K(x, y, \mathrm{d}z) = \bar{K}(x, y)\delta_{x+y}$, for a continuous symmetric function $\bar{K}(x, y)$, and the mass function $m(x) \equiv x$, the above process corresponds to the classical Marcus–Lushnikov process.

Example 2 (Historical Marcus–Lushnikov processes). One may extend E to incorporate not just the masses of clusters, but their histories. Indeed, we can take E to be a space where clusters x encode not only their mass, but the history of coagulations (a binary tree embedded in time) leading to the formation of that particle (see [19] for more details). For these processes, Jacquot in [19] proved a weak law of large numbers for the trajectories $(\bar{\mathbf{L}}_t^{(N)})_{t \in [0,T)}$ when the kernel is a function only of the associated masses and it is bounded from above by a product of sublinear functions.

Example 3 (Toy spatial coagulation models). A large number of toy models that incorporate information about the locations of clusters in 'space' fall into this framework. For example, we may take $E = \mathcal{S} \times (0, \infty)$ where $\mathcal{S} \subseteq \mathbb{R}^d$; in this case an element x of E coincides with a pair (p, n), $p \in \mathcal{S}, n \in (0, \infty)$ and we interpret p as the *location* of a cluster, and $n := m(p, n)$ as its mass. We may, then, assume that after a coagulation between clusters $x = (p, n), y = (s, o)$, the new cluster is placed at a new location, given by a measurable function of the original clusters, for example, the *centre of mass* $\frac{np+os}{n+o}$. Thus, in this case $K((p, n), (s, o), \cdot) =$

$\bar{K}((p,n),(s,o))\delta_{\frac{np+os}{n+o},n+o}$. Another alternative would be a model in which the new particle occupies one of the locations of the previous clusters with probability proportional to their mass, so that $K((p,n),(s,o),\cdot) = \bar{K}((p,n),(s,o))\left(\frac{n}{n+o}\delta_{p,n+o} + \frac{o}{n+o}\delta_{s,n+o}\right)$ (this is the way in which the 'collision operator' is defined in the model of coagulating Brownian clusters of [16]).

Example 4 (Bilinear coagulation processes). In the case that $E = [0,\infty)^d$, $A \in [0,\infty)^{d\times d}$ is a symmetric matrix with non-negative entries and $K(x,y,\mathrm{d}z) = (x^T Ay)\delta_{x+y}$, this model corresponds to the bilinear coagulation model studied in [18]. In that paper, the authors prove a weak law of large numbers for the particle system, showing that the trajectories converge to the unique solution of the Flory equation, and characterise explicitly the 'gelling time', by using comparisons between this process and *inhomogeneous random graph processes*.[1]

2   Gelation in the coagulation process

In this section, we state general sufficient conditions for stochastic gelation in the cluster coagulation model. As this model is rather general, the conditions required are more technical than conditions for the classical Marcus–Lushnikov process. The main motivation for these results is that, in applications to non-equilibrium processes inhomogenieties in the space E (corresponding to, for example, locations in space, the 'types' of cluster, or their velocities) may play a major role in whether or not gelation occurs.

2.1   Main result: sufficient criteria for stochastic gelation

In Assumption 2.1, we incorporate the inhomogeneities into the gelation criterion via a coarse-graining procedure. In other words, we assume that, for each N, we can 'partition' the space E in such a manner that we have sufficient lower bounds on the rate at which clusters belonging to a common partition interact. For example, if E is a metric space, we might partition the space into balls of a fixed radius, thus grouping together clusters that are 'close', or of a similar 'type'. Note that, in our assumptions, these partitions are allowed to grow with N, allowing for a 'finer' approximation as N is made larger.

The techniques we use extend those previously developed for the Marcus–Lushnikov process by Jeon [21] (see also Rezakhanlou [29]).

Assumption 2.1 (Assumptions on the kernel). Suppose that $\{K_N(x,y,\mathrm{d}z)\}_{N\in\mathbb{N}}$ is a sequence of kernels associated with a cluster coagulation process. For functions $\xi, \psi : \mathbb{N} \to \mathbb{N}$, with $\lim_{N\to\infty}\psi(N) = \infty$ we assume the following.

1 For each $j \leq \log_2(\psi(N))$ there exists a partition $\mathscr{P}_N^{(j)}$ of the set $m^{-1}([2^j, 2^{j+1})) \subseteq E$ such that $\left|\mathscr{P}_N^{(j)}\right| \leq \xi(N)$. Moreover, with $c_N'(P,j) := \inf\{\bar{K}_N(x,y) : x,y \in P \in \mathscr{P}_N^{(j)}\}$, we have

$$c_N'(P,j) > 0 \qquad \forall P \in \mathscr{P}_N^{(j)}. \tag{5}$$

---

[1]Actually, the model studied in [18] is slightly more general, in that clusters x belong to a metric space S, and $\bar{K}(x,y) = \pi(x)^T A\pi(y)$, where $\pi : S \to \mathbb{R}^d$ is a continuous function. Clusters x may also change values according to a kernel J on S, in such a way that $\pi(x)$ is preserved.

2 There exists a sequence $(f_j)_{j\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0}$ such that $\sum_{j=0}^{\infty} f_j < \infty$, and

$$\limsup_{N\to\infty} \sum_{j=0}^{\lceil \log_2(\psi(N)) \rceil} \frac{2^j \phi_N(j)}{f_j^2} < \infty, \tag{6}$$

where $\phi_N(j) := \sum_{P\in\mathscr{P}_N^{(j)}} \frac{1}{c'_N(P,j)}$.

3 We have $\lim_{N\to\infty} \frac{\xi(N)}{N} = 0$.

Remark 2.2. The intuition behind Assumption 2.1 can be summarised as follows.

1 Item 1 guarantees that we can partition the space E in such a way that a pair of clusters has strictly positive rate of interaction if the two clusters are close enough and have 'comparable masses'.

2 Item 2, ensures that the lower bounds in (5) are sufficiently large to render the quantity in (6) finite. This is a technical assumption that we need to guarantee that the expectation of $\tau_N(\psi,\delta)$ is bounded from above by a finite value.

3 Item 3 ensures the partitions are not too 'fine'. Under this condition, the pigeonhole principle can be used to show that there are enough clusters in each element of the partition to guarantee that the process does not get 'stuck'.

Remark 2.3. In Item 2 of Assumption 2.1, if the choices of partitions can be made in a manner independent of N, so that $\phi_N(\cdot) \equiv \phi(\cdot)$, then (6) reduces to showing that, for some sequence $(f_j)_{j\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0}$ with $\sum_{j=0}^{\infty} f_j < \infty$, we have $\sum_{j=0}^{\infty} \frac{2^j \phi(j)}{f_j^2} < \infty$.

This may be formulated in a more elegant way: there exists a sequence of positive values $\{f_i\}_{i\in\mathbb{N}}$ such that $\sum_{i\in\mathbb{N}} f_i < \infty$ and $\sum_{i\in\mathbb{N}} 2^j \phi(j)/f_i^2 < \infty$, if and only if $\sum_{i\in\mathbb{N}} (2^j \phi(j))^{1/3} < \infty$. Indeed, on the one hand, if $\sum_{i\in\mathbb{N}} (2^j \phi(j))^{1/3} < \infty$, then it is enough to choose $f_j = (2^j \phi(j))^{1/3}$. For the other implication, by Hölder's inequality we see that $\sum_{j\in\mathbb{N}} (2^j \phi(j))^{1/3} \leq \left(\sum_{j\in\mathbb{N}} \frac{(2^j \phi(j))}{f_j^2} \right)^{\frac{1}{3}} \left(\sum_{i\in\mathbb{N}} f_i \right)^{\frac{2}{3}}$.

In the following theorem, recall that by Remark 1.2 we only consider initial configurations $\pi^{(N)}$ that are finite point processes. Moreover, recall Definition 1.3 and the definition of τ_N from (4).

Theorem 2.4. *Suppose that* $\left\{ (\bar{\mathbf{L}}_t^{(N)})_{t\geq 0} \right\}_{N\in\mathbb{N}}$ *is a cluster coagulation process satisfying Assumption 2.1.*

1 *Suppose that* $\{\pi^{(N)}\}_{N\in\mathbb{N}}$ *is a sequence of deterministic initial conditions such that, there exists* $\varepsilon > 0$ *and* $\rho_0 \in (0,1]$, *such that for all* $N \in \mathbb{N}$, $\left\langle m\mathbf{1}_{m\geq 1}, \pi^{(N)} \right\rangle > \varepsilon$, *and*

$$\frac{\left\langle m\mathbf{1}_{m\geq 1}, \pi^{(N)} \right\rangle}{\left\langle m\mathbf{1}_{m\geq 1}, \pi^{(N)} \right\rangle} \geq \rho_0. \tag{7}$$

Then, there exists a function ψ' *with* $\lim_{N\to\infty} \psi'(N) = \infty$, *such that, for any* $\delta \in (0, \rho_0\varepsilon)$

$$\limsup_{N\to\infty} \mathbb{E}_{N,\pi^{(N)}} \left[\tau_N(\psi'(N),\delta) \right] < C, \tag{8}$$

for a constant C, independent of $\boldsymbol{\pi}^{(N)}$ and δ. Moreover, if $\lim_{N\to\infty}\left\langle m\mathbf{1}_{m\geq 1},\boldsymbol{\pi}^{(N)}\right\rangle = \infty$, then

$$\lim_{N\to\infty}\mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[\tau_N(\psi'(N),\delta)\right] = 0.$$

2 Suppose that $\{\bar{\mathbf{L}}_0^{(N)}\}_{N\in\mathbb{N}}$ is sequence of random initial conditions such that for some $\varepsilon > 0$ and $\rho_0 \in (0,1]$

$$\limsup_{N\to\infty}\mathbb{P}_N\left(\left\langle m\mathbf{1}_{m\geq 1},\bar{\mathbf{L}}_0^{(N)}\right\rangle > \varepsilon, \frac{\left\langle m\mathbf{1}_{m\geq 1},\bar{\mathbf{L}}_0^{(N)}\right\rangle}{\left\langle m\mathbf{1}_{m\geq 1},\bar{\mathbf{L}}_0^{(N)}\right\rangle} \geq \rho_0\right) > 0. \tag{9}$$

Then stochastic gelation occurs in the process.

Remark 2.5. In Theorem 2.4, the condition on $\left\langle m\mathbf{1}_{m\geq 1},\boldsymbol{\pi}^{(N)}\right\rangle$ ensures that there is enough mass bounded from below to form a gel. We can also prove stochastic gelation when the indicators $\mathbf{1}_{m\geq 1}$ may be replaced by any $\mathbf{1}_{m\geq c}$ for any $c > 0$, as long as Assumption 2.1 is satisfied for the re-scaled function $\tilde{m} := m/c$ instead of m. Indeed, gelation with the re-scaled mass function $\tilde{m}$ also implies gelation with the original mass function m.

Remark 2.6. In Theorem 2.4, we can think of N as describing the size of the system. Item 1 is a quenched result, holding for a sequence of deterministic initial conditions $\{\boldsymbol{\pi}^{(N)}\}_{N\in\mathbb{N}}$, whilst Item 2 provides criteria for stochastic gelation given a sequence of random initial conditions.

Remark 2.7. The second part of Item 1 in Theorem 2.4 proves that when $\lim_{N\to\infty}\left\langle m\mathbf{1}_{m\geq 1},\boldsymbol{\pi}^{(N)}\right\rangle = \infty$, and Assumptions 2.1 is satisfied, we have the so called instantaneous gelation.

The following corollary applies these conditions to the classical Marcus–Lushnikov process, i.e., the setting of Example 1. In particular, it provides criteria for stochastic gelation that improves those appearing in [21, Corollary 1], [29, Theorem 1.3] and [23, Proposition 3.15]. In this context, fix a kernel $K_N \equiv K$ (removing dependence on N).

Corollary 2.8. *Let $\left\{(\bar{\mathbf{L}}_t^{(N)})_{t\geq 0}\right\}_{N\in\mathbb{N}}$ be a sequence of Marcus-Lushnikov processes with kernel K, i.e., we are in the setting of Example 1. Suppose that $\exists\,\varepsilon > 0$ and $\rho_0 \in (0,1]$ such that (9) is satisfied for the sequence of initial conditions $\{\bar{\mathbf{L}}_0^{(N)}\}_{N\in\mathbb{N}}$. Moreover, with $c'(j) := \inf_{x,y\in[2^j,2^{j+1})}\bar{K}(x,y) > 0$, for all $j \in \mathbb{N}_0$, suppose that*

$$\sum_{j=0}^{\infty}\left(\frac{2^j}{c'(j)}\right)^{\frac{1}{3}} < \infty, \tag{10}$$

Then stochastic gelation occurs. In particular, when (9) is satisfied, stochastic gelation occurs if one of the following two assumptions is satisfied by the kernel $\bar{K}$:

1 *we have $\inf_{u\in[1,2]}\bar{K}(1,u) > 0$ and for all $x,y > 0$ and all $c > 0$ sufficiently large $\bar{K}(cx,cy) = c^{\gamma}\bar{K}(x,y)$, with $\gamma > 1$;*

2 *we have $\inf_{x,y\geq 1}\bar{K}(x,y) > 0$ and there exists $\epsilon > 0$ such that, for all x,y sufficiently large $\bar{K}(x,y) \geq (x\wedge y)\log(x\wedge y)^{3+\epsilon}$.*

In the space-homogeneous setting of Corollary 2.8, we have the following remarks.

Remark 2.9. Intriguingly, Equation (10) shows that gelation may occur for kernels where $\bar{K}(x,y) = 0$ whenever $y/x > 2$ (or $y/x < 1/2$). This criterion shows that interactions between particles of 'similar size' is enough to ensure gelation, as long as it is strong enough (hence the condition on its growth).

Remark 2.10. In view of Remark 2.5, if (9) is satisfied with the indicators $\mathbf{1}_{m \geq 1}$ replaced by any $\mathbf{1}_{m \geq c}$ for some $c > 0$, and (10) is satisfied with $c'(j) := \inf_{x,y \in [c2^j, c2^{j+1})} \bar{K}(x,y)$, then we can also show that stochastic gelation occurs.

Remark 2.11. We can more precisely quantify the function ψ involved in the (ψ, δ)-stochastic gelation time (i.e., the function describing the sizes of the large particles contributing to gelation - see Definition 1.3). In particular we can show that $T_g^{\psi, \delta} < \infty$ for any $\psi(N) \leq N^b$, and any $b \in (0,1)$. See also Example 6.

2.2   Gelation criteria via coupling with inhomogeneous random graphs

Here we fix $K_N \equiv K$ for some kernel K. It is well-known that for the classical multiplicative kernel $\bar{K}(x,y) = m(x)m(y)$, if the process starts with N particles of mass 1 at time $t = 0$, then the cluster masses at time $t > 0$ are in one-to-one correspondence with the sizes of components of a continuous time analogue of the Erdős–Rényi random graph, where edges appear between pairs of vertices at rate $1/N$. Likewise, the work by Patterson and Heydecker [18] shows that for cluster coagulation models of any 'bilinear' type (see Section 1.3), the cluster masses at any time t are in one-to-one correspondence with the component sizes of an *inhomogenous random graph*. Thus, in regimes where a 'giant component' arises in the inhomogeneous random graph model, gelation occurs in the coagulation process, with an explicit description of the gelation time.

Theorem 2.12 below generalises these results. First, we define some terminology required to make a connection with inhomogeneous random graphs (using the established results from [11]). We call a coagulation kernel K *graph dominating* if

$$\text{for all } x, y, q \in E, \ K(x,y,\cdot)\text{-a.a. } z \quad \text{we have} \quad \bar{K}(z,q) \geq \bar{K}(x,q) + \bar{K}(y,q).$$

If the opposite inequality holds, we say the kernel is *graph dominated*.

In Theorem 2.12 we assume the following. First, the initial conditions are *monodispersed*, i.e., $\mathrm{Supp}\left(\mathbf{L}_0^{(N)}\right) \subseteq \{x \in E : m(x) = 1\}$ for all $N \in \mathbb{N}$. Next, assume that there exists a deterministic $\bar{\mathbf{L}}_0^* \in \mathcal{M}_+(E)$ such that

$$\bar{\mathbf{L}}_0^{(N)} \to \bar{\mathbf{L}}_0^* \quad \text{weakly, in probability.} \tag{11}$$

Moreover, assume that the mass function m is continuous and E is a separable metric space, on which $\bar{\mathbf{L}}_0^*$ is a Borel probability measure. Note that when the initial conditions are monodispersed and (11) holds then $\mathrm{Supp}\left(\bar{\mathbf{L}}_0^*\right) \subseteq \{x \in E : m(x) = 1\}$. Now, define the following quantities

$$T_{\bar{K}, \bar{\mathbf{L}}_0^*} f(x) := \int_E f(y) \bar{K}(x,y) \bar{\mathbf{L}}_0^*(\mathrm{d}y), \quad \Sigma(\bar{K}, \bar{\mathbf{L}}_0^*) := \sup_{f \in L^2(\bar{\mathbf{L}}_0^*), \|f\|_{L^2(\bar{\mathbf{L}}_0^*)} = 1} \left\| T_{\bar{K}, \bar{\mathbf{L}}_0^*} f \right\|_{L^2(\bar{\mathbf{L}}_0^*)}$$

$$\text{and} \quad t^* := \inf\left\{ t > 0 : t\Sigma(K, \bar{\mathbf{L}}_0^*) > 1 \right\}.$$

Theorem 2.12. *Suppose that the above conditions are satisfied. Then the following hold.*

    1 If $(\bar{\mathbf{L}}_t^{(N)})_{t\geq 0}$ is graph dominating, then strong gelation occurs, with strong gelation time at most t^*.

    2 If $(\bar{\mathbf{L}}_t^{(N)})_{t\geq 0}$ is graph dominated then, for any $\psi : \mathbb{N} \to \mathbb{N}$ non-decreasing with $\lim_{N\to\infty} \psi(N) = \infty$, and for any $\delta > 0$ we have $T_g^{\psi,\delta} \geq t^*$.

<div align="right">■</div>

Remark 2.13. Theorem 2.12 works by coupling the coagulation process with the following inhomogeneous random graph process. The vertices of the graph are given by the initial clusters, with atoms of $\mathbf{L}_0^{(N)}$ determining *types* of the initial set of vertices (all of mass 1). Then, edges in the graph appear between vertices of type $x, y \in E$ at rate $\bar{K}(x, y)$. The condition of being graph dominating (similarly graph dominated) allows this coupling to be carried out in such a way that a coagulation between clusters precedes the appearance of an edge between associated connected components in the inhomogeneous random graph model. Thus, gelation occurs before the appearance of a giant component in the graph, well-known to emerge at time t^*. As this coupling argument is quite straightforward, we omit the details in this paper (see, however, the extended pre-print [6]). Nevertheless, we believe that examples, such as the following, mean that this monotonicity is important to be aware of in applied contexts.

Example 5. Suppose we are in the same setting as Example 3 with clusters moving to their centre of mass. Let $\rho : \mathcal{S} \to [0, \infty)$ be an even, concave function. Then the coagulation process with kernel $\bar{K} : (\mathcal{S} \times \mathbb{N}_0)^2 \to [0, \infty)$ defined by $\bar{K}((p, m), (s, n)) := mn\rho(p - s)$ is graph dominating.

3   Proofs of main results

This section is dedicated to the proofs of Theorem 2.4 and Corollary 2.8. For the proof of Item 1 of Theorem 2.4 we fix a sequence of initial conditions $\{\pi^{(N)}\}_{N\in\mathbb{N}}$ satisfying the assumptions of the theorem for some $\varepsilon, \rho_0 > 0$, and fix $\delta \in (0, \rho_0\varepsilon)$ (this is the threshold involved in definition (4)). Recall from (3) that $\mathbb{P}_{N,\pi^{(N)}}(\cdot)$ and $\mathbb{E}_{N,\pi^{(N)}}[\cdot]$ denote probabilities and expectations associated with a coagulation process that starts with such an initial condition.

Since $\delta \in (0, \varepsilon\rho_0)$ we have $\delta = \varepsilon\rho$, for some $\rho \in (0, \rho_0)$. Note that if $(f_k)_{k\in\mathbb{N}_0}$ satisfies the requirement of Item 2 of Assumption 2.1, we may re-scale $(f_k)_{k\in\mathbb{N}_0}$ so that $\sum_{i=0}^{\infty} f_i = \rho_0 - \rho > 0$. This re-scaling does not affect Assumption 2.1. Therefore, with $(f_i)_{i\in\mathbb{N}_0}$ defined in this way, define a strictly decreasing sequence $(\rho_k)_{k\in\mathbb{N}_0}$ as follows: define ρ_0 according to (7), and for $k \in \mathbb{N}$ set

$$\rho_k := \rho + \sum_{i\geq k} f_i, \tag{12}$$

so that $\rho_k \downarrow \rho > 0$.

We recall that in Theorem 2.4 we assume that there exists a function $\psi(N)$ satisfying Assumption 2.1. Ideally, with ψ' being the function appearing in Item 1 of Theorem 2.4, we would have $\psi' = \psi$. However, from a technicality in the proof of Item 1 of Theorem 2.4 (more

specifically to apply Lemma 3.1) we need to define ψ' by

$$\psi'(N) := \psi(N) \wedge \kappa(N) \quad \text{where} \quad \kappa(N) := \max \left\{ k \in \mathbb{N} : \frac{2k}{\min_{0 \leq h \leq \log_2(k)-1} f_h} \leq \frac{N \left\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{\xi(N)} \right\}. \tag{13}$$

In this way we ensure that the following inequality is satisfied:

$$\frac{2\psi'(N)}{\min_{0 \leq h \leq \log_2(\psi'(N))-1} f_h} \leq \frac{N \left\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{\xi(N)} \quad \text{for all } N \in \mathbb{N}. \tag{14}$$

The function $\psi'(N)$ satisfies the conditions of Assumption 2.1 since we have $\lim_{N \to \infty} \psi'(N) = \infty$. Indeed, by definition,

$$\liminf_{N \to \infty} \frac{\kappa(N) + 1}{\min_{0 \leq h \leq \log_2(\kappa(N)+1)-1} f_h} > \liminf_{N \to \infty} \frac{N \left\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2\xi(N)} = \infty, \tag{15}$$

where the last equality follows by Item 3 of Assumption 2.1. If $\kappa(N)$ were bounded infinitely often, the left-side of (15) would also be bounded. Hence $\lim_{N \to \infty} \kappa(N) = \infty$, and since $\lim_{N \to \infty} \psi(N) = \infty$, from (13) we also have $\lim_{N \to \infty} \psi'(N) = \infty$.

Now, in order to prove Theorem 2.4, we define the family of functions $F_k : \mathcal{M}_+(E) \to \mathbb{R}^+$ such that, for each $\mathbf{L} \in \mathcal{M}_+(E)$

$$F_k(\mathbf{L}) := \frac{\langle m \mathbf{1}_{m \geq 2^{k+1}}, \mathbf{L} \rangle}{\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \rangle}. \tag{16}$$

We also define an associated family of stopping times $(\mathcal{T}_k)_{k \in \mathbb{N}}$ such that

$$\mathcal{T}_k := \inf \left\{ t > 0 : \left\langle m \mathbf{1}_{m \geq 2^i}, \bar{\mathbf{L}}_t^{(N)} \right\rangle / \left\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle \geq \rho_i \text{ for all } i = 0, 1, \ldots, k \right\}. \tag{17}$$

In words, $\mathcal{T}_k$ represents the first time that the total mass of clusters with mass at least 2^i exceeds $\rho_i \left\langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle$ for each $i = 0, \ldots, k$. Note that the functions F_k and the times $\mathcal{T}_k$ depend on $\boldsymbol{\pi}^{(N)}$ and on N, but for brevity of notation, we will exclude this dependence in the remainder of the section.

Lemma 3.1. *Suppose that $\vec{v} = (v_1, \ldots, v_n)$, $\vec{c} = (c_1, \ldots, c_n)$ are such that $v_h \in \mathbb{N}$ and $c_h > 0$, for all $h = 1, \ldots, n$. Then, if $\sum_{h=1}^n v_h \geq \kappa_1 > n$ and $\sum_{h=1}^n \frac{1}{c_h} = \kappa_2$ then*

$$\sum_h c_k \left(v_h^2 - v_h \right) \geq \frac{(\kappa_1 - n)^2}{2\kappa_2}. \tag{18}$$

Proof. Re-writing the left side of Equation (18) we get

$$\sum_h c_h(v_h^2 - v_h) \mathbf{1}_{\{v_h > 1\}} \geq \frac{1}{2} \sum_h \frac{(v_h \mathbf{1}_{\{v_h \geq 2\}})^2}{1/c_h} \geq \frac{\left(\sum_h v_h \mathbf{1}_{\{v_h \geq 2\}} \right)^2}{2\kappa_2},$$

where the last inequality is due to the Cauchy–Schwarz inequality. Since $\sum_h v_h \mathbf{1}_{\{v_h \geq 2\}} \geq \kappa_1 - n \geq 1$, we deduce the result.

$\square$

We now prove Theorem 2.4.

3.1 Proof of Item 1 of Theorem 2.4

Proof. We first fix a positive integer $k \leq \lceil \log_2(\psi'(N)) \rceil - 1$, where we recall that the function ψ' comes from (13) (the reason for this choice of k is to be able to apply Item 1 of Assumption 2.1 later on). Note that the function F_k defined in (16) is bounded above by 1 almost surely on the path of the process, and is clearly measurable. Now, recalling the generator in (2), as F_k only considers the masses of clusters larger than 2^{k+1}, ignoring other features, and the mass of the coagulated cluster z is $m(z) = m(x) + m(y)$, we deduce that

$$\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t^{(N)}) = \frac{1}{2 \langle m \mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \rangle} \int_{E \times E} \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x) \left(\bar{\mathbf{L}}_t^{(N)} - \frac{1}{N} \delta_x \right) (\mathrm{d}y) \bar{K}_N(x, y) \tag{19}$$
$$\times \left[(m(x) + m(y)) \mathbf{1}(m(x) + m(y) \geq 2^{k+1}) - m(x) \mathbf{1}(m(x) \geq 2^{k+1}) - m(y) \mathbf{1}(m(y) \geq 2^{k+1}) \right].$$

Now,

$$M_{F_k}^{(N)}(t) := F_k(\bar{\mathbf{L}}_t^{(N)}) - F_k(\boldsymbol{\pi}^{(N)}) - \int_0^t \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_s^{(N)}) \mathrm{d}s$$

is a martingale with respect to the natural filtration of the process[2] initiated by the measure $\boldsymbol{\pi}^{(N)}$. Let $I, J > 0$ be given. By Doob's optional sampling theorem (see, for example, [28, Corollary 2.3.6 and Theorem 2.3.2]) applied to the bounded stopping times $I \wedge \mathcal{T}_k \leq I \wedge \mathcal{T}_{k+1}$ we have

$$\mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[F_k(\bar{\mathbf{L}}_{I \wedge \mathcal{T}_{k+1}}) \right] = \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[F_k(\bar{\mathbf{L}}_{I \wedge \mathcal{T}_k}) \right] + \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\int_{I \wedge \mathcal{T}_k}^{I \wedge \mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right]$$
$$\geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\int_{I \wedge \mathcal{T}_k}^{I \wedge \mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right]$$
$$\geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k \leq J} \int_{I \wedge \mathcal{T}_k}^{I \wedge \mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right].$$

Noting that $\tilde{\mathcal{A}}_N F_k$ is non-negative and the integrand on the left is monotone in I, by monotone convergence in I, and Fatou's lemma we have

$$\mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[F_k(\bar{\mathbf{L}}_{\mathcal{T}_{k+1}}) \right] \geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k \leq J} \liminf_{I \to \infty} \int_{I \wedge \mathcal{T}_k}^{I \wedge \mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right]$$
$$= \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k \leq J} \int_{\mathcal{T}_k}^{\mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right].$$

Now, by applying monotone convergence in J, we deduce that

$$\mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[F_k(\bar{\mathbf{L}}_{\mathcal{T}_{k+1}}) \right] \geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k < \infty} \int_{\mathcal{T}_k}^{\mathcal{T}_{k+1}} \tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t) \mathrm{d}t \right]. \tag{20}$$

We now seek lower bounds for the quantity $\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t)$ appearing in (20), when $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$ (notice that $\mathcal{T}_{k+1}$ might even be equal to $+\infty$). First, note that

$$m(x)\mathbf{1}(m(x) \geq 2^{k+1}) + m(y)\mathbf{1}(m(y) \geq 2^{k+1}) \leq (m(x) + m(y))\mathbf{1}(m(x) \text{ or } m(y) \geq 2^{k+1}). \tag{21}$$

---

[2]The fact that $M_{F_k}^{(N)}(t)$ is a martingale follows from the definition of the infinitesimal generator of the process, see, for example, [28, Proposition 7.1.6]

Using (21), we may bound (19) from below so that

$$
\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t^{(N)}) \geq \frac{1}{2\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \int_{E\times E} \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\left(\bar{\mathbf{L}}_t^{(N)} - \frac{1}{N}\delta_x\right)(\mathrm{d}y)\bar{K}_N(x,y)
$$
$$
\times (m(x)+m(y))\mathbf{1}(m(x)+m(y) \geq 2^{k+1} > m(x), m(y))
$$
$$
\geq \frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \int_{E\times E} \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\left(\bar{\mathbf{L}}_t^{(N)} - \frac{1}{N}\delta_x\right)(\mathrm{d}y)\bar{K}_N(x,y)
$$
$$
\times \mathbf{1}(m(x)+m(y) \geq 2^{k+1} > m(x), m(y))
$$
$$
\geq \frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \sum_{P\in\mathscr{P}_N^{(k)}} \int_{P\times P} \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\left(\bar{\mathbf{L}}_t^{(N)} - \frac{1}{N}\delta_x\right)(\mathrm{d}y)\bar{K}_N(x,y)
$$
$$
\times \mathbf{1}(m(x)+m(y) \geq 2^{k+1} > m(x), m(y)),
$$

where the last inequality follows from restricting the integral from the space $E\times E$ to the space $\bigcup_{P\in\mathscr{P}_N^{(k)}} P\times P$, with the partition $\mathscr{P}_N^{(k)}$ as defined in Item 1 of Assumption 2.1. Now, using the definition of $c_N'(P,k)$ in Item 1 of Assumption 2.1 we bound this further by

$$
\frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k) \int_{P\times P} \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\left(\bar{\mathbf{L}}_t^{(N)} - \frac{1}{N}\delta_x\right)(\mathrm{d}y) \quad (22)
$$
$$
\times \mathbf{1}(m(x)+m(y) \geq 2^{k+1} > m(x), m(y)).
$$

Moreover, since $P\in\mathscr{P}_N^{(k)}$, by Item 1 of Assumption 2.1 for $(x,y)\in P\times P$ we have $2^k \leq m(x), m(y) < 2^{k+1}$. This implies that the indicator $\mathbf{1}(m(x)+m(y) \geq 2^{k+1} > m(x), m(y))$ in (22) is one. Thus, we can re-write (22) as

$$
\frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k) \int_P \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\left(\left(\int_P \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)\right) - \frac{1}{N}\right).
$$

For brevity of notation, we now write $\langle f, \bar{\mathbf{L}}_t^{(N)}\rangle$ for terms of the form $\int_E \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}x)f(x)$ or $\int_E \bar{\mathbf{L}}_t^{(N)}(\mathrm{d}y)f(y)$. Then, summing over the possible values $m(x), m(y) \in [2^k, 2^{k+1})$, we deduce the lower bound

$$
\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t^{(N)}) \geq \frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \sum_{P\in\mathscr{P}_N^{(k)}} \sum_{n_1,n_2=2^k}^{2^{k+1}-1} c_N'(P,k) \left\langle \mathbf{1}_{n_1\leq m(x)<n_1+1, x\in P}, \bar{\mathbf{L}}_t^{(N)}\right\rangle
$$
$$
\times \left(\left\langle \mathbf{1}_{n_2\leq m(x)<n_2+1, x\in P}, \bar{\mathbf{L}}_t^{(N)}\right\rangle - \frac{1}{N}\mathbf{1}_{\{n_1=n_2\}}\right)
$$
$$
= \frac{2^k}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle} \quad (23)
$$
$$
\times \frac{1}{N^2} \sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)\left(\left(N\sum_{n=2^k}^{2^{k+1}-1}\left\langle \mathbf{1}_{n\leq m(x)<n+1, x\in P}, \bar{\mathbf{L}}_t^{(N)}\right\rangle\right)^2 - N\sum_{n=2^k}^{2^{k+1}-1}\left\langle \mathbf{1}_{n\leq m(x)<n+1, x\in P}, \bar{\mathbf{L}}_t^{(N)}\right\rangle\right).
$$

We now wish to apply Lemma 3.1 to the sum indexed by the elements P of $\mathscr{P}_N^{(k)}$ in the above display, where the integer valued random variables $\sum_{n=2^k}^{2^{k+1}-1} N\left\langle \mathbf{1}_{n\leq m(x)<n+1, x\in P}, \bar{\mathbf{L}}_t^{(N)}\right\rangle$ play the role of v_h in the lemma. These integers count the number of clusters at time t in each

element P of the partition $\mathscr{P}_N^{(k)}$. Since $|\mathscr{P}_N^{(k)}| \leq \xi(N)$, $\xi(N)$ plays the role of n appearing in Lemma 3.1.

Note, that, if $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$, by the definition in (17), we have

$$\left\langle m\mathbf{1}_{m \geq 2^k}, \bar{\mathbf{L}}_t^{(N)} \right\rangle \geq \rho_k \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle \quad \text{and} \quad \left\langle m\mathbf{1}_{m \geq 2^{k+1}}, \bar{\mathbf{L}}_t^{(N)} \right\rangle < \rho_{k+1} \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle,$$

so that

$$\left\langle m\mathbf{1}_{2^k \leq m < 2^{k+1}}, \bar{\mathbf{L}}_t^{(N)} \right\rangle \overset{(12)}{\geq} f_k \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle. \tag{24}$$

Therefore, we choose κ_1 in Lemma 3.1 such that

$$\sum_{h=1}^{n} v_h = \sum_{P \in \mathscr{P}_N^{(k)}} \sum_{n=2^k}^{2^{k+1}-1} N \left\langle \mathbf{1}_{n \leq m(x) < n+1, x \in P}, \bar{\mathbf{L}}_t^{(N)} \right\rangle = N \left\langle \mathbf{1}_{2^k \leq m < 2^{k+1}}, \bar{\mathbf{L}}_t^{(N)} \right\rangle$$

$$\overset{(24)}{\geq} N \frac{f_k \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2^{k+1}} =: \kappa_1$$

For the assumptions of Lemma 3.1 to hold, we further require

$$\kappa_1 = N \frac{f_k \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2^{k+1}} > \xi(N).$$

This follows from the fact that

$$\kappa_1 = N \frac{f_k \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2^{k+1}} \geq \frac{(\min_{0 \leq \ell \leq \lceil \log_2(\psi'(N)) \rceil - 1} f_\ell) N \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{\psi'(N)} \overset{(14)}{\geq} 2\xi(N) > \xi(N) \tag{25}$$

where the first inequality in (25) is obtained by inserting the value $k = \log_2(\psi'(N)) - 1$ in the denominator. Therefore, applying Lemma 3.1 to the sum in (23), with $\kappa_2 := \sum_{P \in \mathscr{P}_N^{(k)}} c_N'(P, k)^{-1}$ we obtain the lower bound:

$$\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t^{(N)}) \geq \frac{2^k}{\left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle N^2} \times \frac{(\kappa_1 - n)^2}{2\kappa_2} = \frac{2^{k-1} \left(\left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle \frac{f_k}{2^{k+1}} - \frac{\xi(N)}{N} \right)^2}{\left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle \sum_{P \in \mathscr{P}_N^{(k)}} c_N'(P, k)^{-1}},$$

$\forall t \in [\mathcal{T}_k, \mathcal{T}_{k+1})$. Since (25) implies that $\frac{\xi(N)}{N} \leq \frac{1}{2} \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle \frac{f_k}{2^{k+1}}$, we finally deduce the bound

$$\tilde{\mathcal{A}}_N F_k(\bar{\mathbf{L}}_t^{(N)}) \geq \frac{f_k^2 \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2^{k+5} (\sum_{P \in \mathscr{P}_N^{(k)}} c_N'(P, k)^{-1})}, \quad \text{for all } k \leq \log_2(\psi'(N)) - 1 \text{ and } t \in [\mathcal{T}_k, \mathcal{T}_{k+1}).$$

Now, recalling that $0 \leq F_k(\bar{\mathbf{L}}_t^{(N)}) \leq 1$, by applying (20) we have

$$1 \geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[F_k(\bar{\mathbf{L}}_{\mathcal{T}_{k+1}}^{(N)}) \right]$$

$$\geq \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k < \infty} \int_{\mathcal{T}_k}^{\mathcal{T}_{k+1}} \tilde{\mathcal{A}} F_k(\bar{\mathbf{L}}_t^{(N)}) \mathrm{d}t \right]$$

$$\geq \frac{f_k^2 \left\langle m\mathbf{1}_{m \geq 1}, \boldsymbol{\pi}^{(N)} \right\rangle}{2^{k+5} (\sum_{P \in \mathscr{P}_N^{(k)}} c_N'(P, k)^{-1})} \mathbb{E}_{N, \boldsymbol{\pi}^{(N)}} \left[\mathbf{1}_{\mathcal{T}_k < \infty} (\mathcal{T}_{k+1} - \mathcal{T}_k) \right]$$

which in turn yields

$$\mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[\mathbf{1}_{\mathcal{T}_k<\infty}(\mathcal{T}_{k+1}-\mathcal{T}_k)\right] \leq \frac{2^{k+5}(\sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)^{-1})}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle f_k^2}. \tag{26}$$

Now, by assumption in Equation (7), we have $\mathcal{T}_0 = 0 < \infty$ almost surely. Thus, by induction and the bound in (26), for all $k \leq \lceil\log_2(\psi'(N))\rceil - 1$, we have $\mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}[\mathcal{T}_k] < \infty$ and hence $\mathcal{T}_k < \infty$ almost surely. Therefore, for all $k \leq \lceil\log_2(\psi'(N))\rceil - 1$ we may re-write (26) as

$$\mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[\mathcal{T}_{k+1}-\mathcal{T}_k\right] \leq \frac{2^{k+5}(\sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)^{-1})}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle f_k^2}. \tag{27}$$

Then, by summing (27), we deduce that

$$\mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[\mathcal{T}_{\lceil\log_2(\psi'(N))\rceil}\right] = \sum_{k=0}^{\lceil\log_2(\psi'(N))\rceil-1} \mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[(\mathcal{T}_{k+1}-\mathcal{T}_k)\right]$$

$$\leq \frac{1}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle}\sum_{k=0}^{\lceil\log_2(\psi'(N))\rceil-1} \frac{2^{k+5}(\sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)^{-1})}{f_k^2}. \tag{28}$$

Now, recall that according to the definition in (12) and just above it, $\delta = \rho\varepsilon$ and $\rho_k > \rho$ for all k. Therefore we have

$$\mathcal{T}_{\lceil\log_2(\psi'(N))\rceil} = \inf\left\{t > 0 : \left\langle m\mathbf{1}_{m\geq 2^i}, \bar{\mathbf{L}}_t^{(N)}\right\rangle / \left\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\right\rangle \geq \rho_i, \forall i = 0, 1, \ldots, \lceil\log_2(\psi'(N))\rceil\right\}$$

$$\geq \inf\left\{t > 0 : \left\langle m\mathbf{1}_{m\geq\psi'(N)}, \bar{\mathbf{L}}_t^{(N)}\right\rangle / \left\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\right\rangle \geq \rho\right\}$$

$$\geq \inf\left\{t > 0 : \left\langle m\mathbf{1}_{m\geq\psi'(N)}, \bar{\mathbf{L}}_t^{(N)}\right\rangle \geq \rho\varepsilon\right\} \overset{(4)}{=} \tau_N(\psi'(N), \delta), \tag{30}$$

almost surely.

Combining (30) with (28), we deduce that

$$\mathbb{E}_N\left[\tau_N(\psi'(N), \delta)\right] \leq \mathbb{E}_{N,\boldsymbol{\pi}^{(N)}}\left[\mathcal{T}_{\lceil\log_2(\psi'(N))\rceil}\right]$$

$$\leq \frac{1}{\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\rangle}\sum_{k=0}^{\lceil\log_2(\psi'(N))\rceil-1} \frac{2^{k+5}(\sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)^{-1})}{f_k^2}$$

$$\leq \frac{1}{\varepsilon}\sum_{k=0}^{\lceil\log_2(\psi'(N))\rceil-1} \frac{2^{k+5}(\sum_{P\in\mathscr{P}_N^{(k)}} c_N'(P,k)^{-1})}{f_k^2}. \tag{31}$$

By Equation (6), the right-side of (31) is a constant $C = C(\varepsilon)$, independent of $\boldsymbol{\pi}^{(N)}$ and δ, from which we deduce (8). In addition, the right-side just above (31) tends to zero if $\left\langle m\mathbf{1}_{m\geq 1}, \boldsymbol{\pi}^{(N)}\right\rangle \to \infty$, proving the last statement in Item 1 of Theorem 2.4. $\qquad\square$

3.2   Proof of Item 2 of Theorem 2.4

Proof of Item 2 of Theorem 2.4. To prove Item 2 of Theorem 2.4 we apply Item 1 of Theorem 2.4 and anneal over the possible values of $\bar{\mathbf{L}}_0^{(N)}$. In this regard, define the event

$$\mathcal{E}^{(N)}(\varepsilon, \rho_0) := \left\{\left\langle m\mathbf{1}_{m\geq 1}, \bar{\mathbf{L}}_0^{(N)}\right\rangle > \varepsilon, \frac{\left\langle m\mathbf{1}_{m\geq 1}, \bar{\mathbf{L}}_0^{(N)}\right\rangle}{\left\langle m\mathbf{1}_{m\geq 1}, \bar{\mathbf{L}}_0^{(N)}\right\rangle} \geq \rho_0\right\}.$$

Since the bound in (8) is independent of $\pi^{(N)}$ and N, see (31), we have

$$\mathbb{E}_N\left[\tau_N(\psi'(N),\delta)\,\Big|\,\mathcal{E}^{(N)}(\varepsilon,\rho_0)\right] \leq C.$$

Using the assumption from (9), set $p_0 := \limsup \mathbb{P}_N\left(\mathcal{E}^{(N)}(\varepsilon,\rho_0)\right) > 0$. Then, there exists a subsequence $(N_j)_{j\in\mathbb{N}}$ such that $\lim_{j\to\infty}\mathbb{P}_{N_j}\left(\mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right) = p_0$. In particular, for N_j sufficiently large,

$$\mathbb{P}_{N_j}\left(\mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right) > 0.$$

By Markov's inequality,

$$\mathbb{P}_{N_j}\left(\tau_{N_j}(\psi'(N_j),\delta) \leq 2C\,\Big|\,\mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right) \geq \frac{1}{2}.$$

We thus get

$$\begin{aligned}
\limsup_{N\to\infty} \mathbb{P}_N\left(\tau_N(\psi'(N),\delta) \leq 2C\right) &\geq \lim_{j\to\infty}\mathbb{P}_{N_j}\left(\tau_{N_j}(\psi'(N_j),\delta) \leq 2C\right)\\
&\geq \lim_{j\to\infty}\mathbb{P}_{N_j}\left(\left\{\tau_{N_j}(\psi'(N_j),\delta) \leq 2C\right\} \cap \mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right)\\
&= \lim_{j\to\infty}\mathbb{P}_{N_j}\left(\tau_{N_j}(\psi'(N_j),\delta) \leq 2C\,\Big|\,\mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right)\mathbb{P}_{N_j}\left(\mathcal{E}^{(N_j)}(\varepsilon,\rho_0)\right) \geq \frac{p_0}{2}.
\end{aligned}$$

We deduce that stochastic gelation occurs according to Definition 1.3 (with $T_g^{\psi',\delta} \leq 2C$). $\quad\square$

3.3   Proof of Corollary 2.8

We finish this section with the proof of Corollary 2.8.

Proof of Corollary 2.8. To prove the first statement of the corollary we apply Item 2 of Theorem 2.4. In order to do so, we need to show that Assumption 2.1 is satisfied. Recall that in this setting $E = (0,\infty)$ and $m(x) = x$ for all $x \in (0,\infty)$. Therefore, in Assumption 2.1, for every j we can take trivial single-set partition of $[2^j, 2^{j+1})$, so that $\xi(N) \equiv 1$. Since each of the partitions $\mathscr{P}_N^{(j)}$ consist of the single set $[2^j, 2^{j+1})$, and we assumed $c'(j) = \inf_{x,y\in[2^j,2^{j+1})}\bar{K}(x,y) > 0$, we also have, for all $N \in \mathbb{N}$

$$c'_N(P,j) \equiv c'_N([2^j,2^{j+1}),j) \equiv c'(j) > 0, \qquad \forall j \in \mathbb{N}_0,$$

where $c'_N(P,j)$ are as defined in Item 1 of Assumption 2.1. Now, for Item 2, Equation (6) reduces to showing that for some sequence $(f_j)_{j\in\mathbb{N}_0} \in (0,\infty)^{\mathbb{N}_0}$ with $\sum_{j=0}^{\infty} f_j < \infty$, we have

$$\sum_{j=0}^{\infty} \frac{2^j}{f_j^2 c'(j)} < \infty,$$

where we note that the sum may be taken to infinity because in this case the term ϕ_N is independent of N. However, by Remark 2.3, this is equivalent to (10). Finally, for Item 3 of Assumption 2.1, since $\xi(N) \equiv 1$, we clearly have $\lim_{N\to\infty}\frac{\xi(N)}{N} = 0$. This concludes the proof of the first statement of the corollary.

For the second assertion, note that in either of the two cases, we have $c'(j) = \inf_{x,y\in[2^j,2^{j+1})}\bar{K}(x,y) > 0$ for all $j \in \mathbb{N}_0$. Hence, we only need to show that under each assumption, (10) is satisfied. Indeed,

1 Under Item 1, set $\kappa^* := \inf_{i \in [1,2]} \bar{K}(1,i) > 0$. By the homogeneity assumption there exists $j_0 \in \mathbb{N}$ such that, for all $j \geq j_0$ we have (assuming without loss of generality $x < y$), $\bar{K}(x,y) = x^\gamma \bar{K}\left(1, \frac{y}{x}\right) \geq \kappa^* 2^{\gamma j} =: c'(j)$ whenever $2^j \leq x, y < 2^{j+1}$. This implies that

$$\sum_{j=0}^\infty \left(\frac{2^j}{c'(j)}\right)^{\frac{1}{3}} = \frac{1}{(\kappa^*)^{1/3}} \sum_{j=0}^\infty 2^{(1-\gamma)/3} < \infty,$$

whenever $\gamma > 1$.

2 Under Item 2, by assumption there exists a $j_0 \in \mathbb{N}$ such that, for all $j \geq j_0$ we have $\bar{K}(x,y) \geq 2^j (j \log 2)^{3+\epsilon} =: c'(j)$ for $2^j \leq x, y < 2^{j+1}$. This implies that

$$\sum_{j=0}^\infty \left(\frac{2^j}{c'(j)}\right)^{\frac{1}{3}} = \frac{1}{(\log 2)^{1+\epsilon/3}} \sum_{j=0}^\infty \frac{1}{j^{1+\epsilon/3}} < \infty.$$

$\square$

Example 6. As mentioned in Remark 2.11, in the setting of Corollary 2.8, we can quantify the sizes of the large clusters contributing to gelation. Indeed, the proof of Item 1 of Theorem 2.4 can be retraced with the choice of $f_k := \frac{1}{k^{1+\epsilon}}$ (recall the assumptions introduced around Equation (12)). In particular condition (14) reduces to check whether

$$\frac{N}{4} \geq \psi'(N)(\lceil \log_2(\psi'(N)) \rceil)^{1+\epsilon} \qquad \forall N \in \mathbb{N}$$

(recall that $\xi(N) \equiv 1$). This can be achieved by taking for instance any $\psi'(N) \leq N^b$ for any $b \in (0,1)$, and ϵ sufficiently small.

References

[1] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[2] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

[3] D. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3:no. 3, 59, 1998.

[4] D. Aldous and J. Pitman. Tree-valued Markov chains derived from Galton-Watson processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(5):637–686, 1998.

[5] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[6] L. Andreis, T. Iyer, and E. Magnanini. Gelation in cluster coagulation processes. *arXiv preprint arXiv:2308.10232v2*, 2023.

[7] L. Andreis, T. Iyer, and E. Magnanini. Hydrodynamic limits in cluster coagulation processes. arXiv preprint, 2024.

[8] L. Andreis, W. König, H. Langhammer, and R. Patterson. Spatial particle processes with coagulation: Gibbs-measure approach, gelation and smoluchowski equation. arXiv preprint arXiv:2401.06668, 2024.

[9] L. Andreis, W. König, and R. Patterson. A large-deviations principle for all the cluster sizes of a sparse Erdős-Rényi graph. *Random Structures & Algorithms*, 59(4):522–553, 2021.

[10] J. Bertoin. Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.*, 29(1):344–360, 2001.

[11] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31(1):3–122, 2007.

[12] A. Eibeck and W. Wagner. Stochastic particle approximations for Smoluchoski's coagulation equation. *Ann. Appl. Probab.*, 11(4):1137–1165, 2001.

[13] M. Escobedo, P. Laurençot, S. Mischler, and B. Perthame. Gelation and mass conservation in coagulation-fragmentation models. *J. Differential Equations*, 195(1):143–174, 2003.

[14] M. Escobedo, S. Mischler, and B. Perthame. Gelation in coagulation and fragmentation models. *Comm. Math. Phys.*, 231(1):157–188, 2002.

[15] D. T. Gillespie. The stochastic coalescence model for cloud droplet growth. *J. Atmos. Sci.*, 29:1496–1510, 1972.

[16] A. Hammond and F. Rezakhanlou. The kinetic limit of a system of coagulating Brownian particles. *Arch. Ration. Mech. Anal.*, 185(1):1–67, 2007.

[17] A. Hammond and F. Rezakhanlou. Moment bounds for the Smoluchowski equation and their consequences. *Comm. Math. Phys.*, 276(3):645–670, 2007.

[18] D. Heydecker and R. I. A. Patterson. Bilinear coagulation equations. arXiv preprint arXiv:1902.07686, 2019.

[19] S. Jacquot. A historical law of large numbers for the Marcus-Lushnikov process. *Electron. J. Probab.*, 15:no. 19, 605–635, 2010.

[20] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures Algorithms*, 4(3):231–358, 1993. With an introduction by the editors.

[21] I. Jeon. Existence of gelling solutions for coagulation-fragmentation equations. *Comm. Math. Phys.*, 194(3):541–567, 1998.

[22] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl.*, 13(3):235–248, 1982.

[23] P. Laurençot. Weak compactness techniques and coagulation equations. In *Evolutionary equations with applications in natural sciences*, pages 199–253. Springer, 2014.

[24] A. Lushnikov. Some new aspects of coagulation theory. *Izv. Acad. Sci. USSR, Ser. Phys. Atmos. Oceans*, 14, 01 1978.

[25] A. H. Marcus. Stochastic coalescence. *Technometrics*, 10:133–143, 1968.

[26] J. R. Norris. Cluster coagulation. *Comm. Math. Phys.*, 209(2):407–435, 2000.

[27] J. R. Norris. Measure solutions for the Smoluchowski coagulation-diffusion equation. arXiv preprint arXiv:1408.5228, 2014.

[28] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[29] F. Rezakhanlou. Gelation for Marcus-Lushnikov process. *Ann. Probab.*, 41(3B):1806–1830, 2013.

[30] F. Rezakhanlou. Pointwise bounds for the solutions of the Smoluchowski equation with diffusion. *Arch. Ration. Mech. Anal.*, 212(3):1011–1035, 2014.

[31] R. Siegmund-Schultze and W. Wagner. Induced gelation in a two-site spatial coagulation model. *Ann. Appl. Probab.*, 16(1):370–402, 2006.

[32] W. Wagner. Explosion phenomena in stochastic coagulation-fragmentation models. *Ann. Appl. Probab.*, 15(3):2081–2112, 2005.

[33] W. Wagner. Post-gelation behavior of a spatial coagulation model. *Electron. J. Probab.*, 11:no. 35, 893–933, 2006.

[34] M. R. Yaghouti, F. Rezakhanlou, and A. Hammond. Coagulation, diffusion and the continuous Smoluchowski equation. *Stochastic Process. Appl.*, 119(9):3042–3080, 2009.