Kac Interaction Clusters: A Bilinear Coagulation Equation and Phase Transition

Daniel Heydecker*, Robert I. A. Patterson†

February 21, 2019

Abstract

We consider the interaction clusters for Kac’s model of a gas with quadratic interaction rates, and show that they behave as coagulating particles with a bilinear coagulation kernel. In the large particle number limit the distribution of the interaction cluster sizes is shown to follow an equation of Smoluchowski type. Using a coupling to random graphs, we analyse the limiting equation, showing well-posedness, and a closed form for the time of the gelation phase transition \( t_g \) when a macroscopic cluster suddenly emerges. We further prove that the second moment of the cluster size distribution diverges exactly at \( t_g \). Our methods apply immediately to coagulating particle systems with other bilinear coagulation kernels.

1 Introduction and Main Results

Boltzmann [8] pictured gases as collections of billiard balls, moving in straight lines except when they “randomly” come close to one another or to the wall of a container. Under the intuition that particles should be ‘unordered at the molecular level’, he then derived the Boltzmann equation for the molecular velocity distribution; justifying this intuition, which became known as ‘molecular chaos’, has been an active area of research since. Grad helped clarify in exactly which limit a result should be possible, the Boltzmann–Grad limit, and the first result was due to Lanford [22] three quarters of a century after Boltzmann. Lanford’s result is restricted to a fraction of the mean free time, so that a positive fraction of the molecules have not collided with any other molecule; a convergence result for arbitrary finite time intervals and a wide range of initial conditions is still not available. For more detailed discussion and references to more recent results see [35].

Three quarters of a century is a long time to wait, and during this time Kac [19] introduced a simplified model for the velocities of gas molecules. In this model, position is ignored and deterministic collisions based on trajectory intersections are replaced by a stochastic collision rule, an idea also introduced by Leontovich [23]. For this simplified model, a wide body of literature [26, 15, 37, 27, 32] has shown propagation of chaos, and that the single particle marginal converges to the solution of the spatially homogeneous Boltzmann equation.

The principal challenge in analysing the full billiard model are the dependencies that arise between the molecular velocities as a result of pairs of particles either having collided or not collided. These dependencies are suppressed in the Kac model. The Lanford proof looks backward in time from a pair of molecules to the most recent collision involving either molecule, and then recursively builds a tree-like structure. For this argument, a restriction to

*University of Cambridge, dhl489@cam.ac.uk. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L016516/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis

†Robert.i.a.patterson@wias-berlin.de. This research was supported by the Deutsche Forschungsgemeinschaft (DFG) grant CRC 1114 “Scaling Cascades in Complex Systems”, Project C08.
a short, finite time is necessary, in order to guarantee that the branching process is subcritical and no infinite trees occur.

The trees appearing in the Lanford proof may be extended to a partition of the molecules into interaction clusters [33], such that any two molecules which have collided belong to the same cluster. The concept of interaction clusters was introduced by Gabrilov et al. [14], who show interesting properties of the interaction cluster size distribution by molecular dynamics simulation and in particular the formation of a giant cluster in a phase transition. We refer to this phenomenon as gelation.

The distributions of the sizes of the interaction clusters are formally derived up to the critical time in [33] in terms of the solution of the Boltzmann equation. Reducing to the case of cutoff Maxwell molecules for the spatially homogeneous Boltzmann equation, the phase transition observed in [14] can be identified precisely and the cluster size distributions observed to match those arising from the Smoluchowski coagulation equation with product kernel [25, 4, 33]. Heuristically, when a collision occurs, the corresponding clusters merge, which may be represented as a coagulation event at the level of interaction clusters.

In [34] the clusters were studied in the Kac setting and the restriction to Maxwell molecules was lifted. This allowed a general collision rate including the hard sphere case and it was formally shown in a large particle number limit that the distribution of the cluster sizes converges to a version of the Smoluchowski coagulation equation with a time-dependent product kernel.

In this work, we will consider a class of Kac processes for which the kernel is sufficiently tractable to allow a detailed analysis of the convergence and the limiting process, including the effects of a ‘gel’ formed when interaction clusters become macroscopic beyond the critical time. Our techniques rely on the properties bilinear coagulation kernels, generalising the notion of the product kernel [31], and on exploiting a random graph representation of the particle system, in the form considered by [7].

The study of gelation as the formation of a very large connected structure by joining basic building blocks goes back at least to Flory [13] whose motivation was hydrocarbon polymerisation in the manufacture of plastics. Flory understood polymerisation as the formation of a random graph, rather than in terms of coagulation, and was aware of a sharp phase transition at the emergence of a giant connected structure, which he termed ‘gel’. A rigorous proof of the random graph phase transition was provided by Erdős and Rényi [11]. The existence of a phase transition corresponding to the formation of a giant particle in a coagulation model is first discussed by Lushnikov [25], who uses this to explain the explosion of the second moment and the failure of mass conservation for the solution of the Smoluchowski equation with the product coagulation kernel. The first connection between random graph and particle approaches appears in [9], where the phase transition is proved for the particle coagulation process and an interpretation as a new proof for a phase transition in the Erdős-Rényi random graph is noted; this is also discussed in the survey article [4]. We extend this connection, and show that the bilinear form of the merger rate allows us to couple the stochastic coagulant process to inhomogenous random graphs as considered by [7].

For particles with integer masses coagulating according to a kernel bounded above by the product kernel and below by its square root Jeon [18] proved the existence of a gelation phase transition and provided an upper bound on the gelation time and presents a number of different definitions of the gelation time for coagulation–fragmentation models. The product coagulation kernel is the product of the masses of the coagulation partners, which have no properties other than mass. Norris [30, 31] replaced mass with a general function growing no more than linearly in particle mass and suitable for use in models where particles have internal structure, a step that is important for the present work. A lower bound for the gelation time was proved in [30] and an upper bound was added under appropriate assumptions in [31]; however, these bounds do not coincide in general. Normand [28] obtained explicit results concerning the blowup of a second moment for a sexed model which gives a lower bound on the gelation time, and in a later work [29] finds explicit expressions for the gelation time for a selection of models with arms. Consequently, ours is one of the first models for which the gelation time can be found exactly; moreover, several aspects of our analysis extend what was previously known about the Smoluchowski equation, using the connection to random graphs [7]. We also propose a notion of bilinear coagulation kernels, which would also be amenable to our analysis.
1.1 Definitions

1.1.1 Markov jump processes

The Kac process with parameter $N$ defines Markovian dynamics for $N$ molecules, each with mass $N^{-1}$ and a velocity $v_i(t) \in \mathbb{R}^d$, undergoing pairwise collisions that conserve momentum and energy. We write $B$ for the kernel on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ underlying the dynamics of the Kac process, and $(v_i^N(t))_{i=1}^N$ for the state of the Kac process at time $t \geq 0$. At time $t$ the instantaneous collision rate between an unordered pair of molecules $i$ and $j$ is given by $\frac{2}{N}B(v_i^N(t) - v_j^N(t), \mathbb{S}^{d-1})$, and the post-collision velocities are given by

$$
\tilde{v}_i^N(t) = \frac{v_i^N(t) + v_j^N(t)}{2} + \sigma \frac{|v_i^N(t) - v_j^N(t)|}{2},
$$
(1)

$$
\tilde{v}_j^N(t) = \frac{v_i^N(t) + v_j^N(t)}{2} - \sigma \frac{|v_i^N(t) - v_j^N(t)|}{2},
$$
(2)

with $\sigma$ distributed according to $B(v_i^N(t) - v_j^N(t), \cdot)$. For the purposes of this work we consider kernels whose total mass is of the form

$$
B(v, \mathbb{S}^{d-1}) = [\kappa + \gamma|v|^2]; \quad \kappa, \gamma \geq 0.
$$
(3)

The case $\gamma = 0$ is that of cutoff Maxwell molecules, which arises in $d$ dimensions when molecules are modelled as repelling each other with a force inversely proportional to the $(2d-1)^{th}$ power of their separation. Replacing $3$ with $B(v, \mathbb{S}^{d-1}) = |v|$ would give the hard spheres model, which can be seen as the large exponent limit of an inverse power law force model, and arises when molecules are modelled as non-overlapping ‘billiards’. The contribution from $\gamma|v|^2$ is a ‘very hard spheres’ term, which we will refer to as quadratic.

The quadratic model has appeared previously in the literature [24, 34] and is mathematically interesting, although it does not arise from a physical interaction model [38]. Since Maxwell molecules and the quadratic model (with $\kappa = 0$) are the cases $B(v, \mathbb{S}^{d-1}) \propto |v|^n$ for $n = 0, 2$ respectively it is reasonable to regard the hard spheres ($n = 1$) as an intermediate case and to suppose that much qualitative behaviour shared by the Maxwell and quadratic cases will also be seen for hard spheres. This supposition is supported by numerical observations in [34]. For our analysis, the Maxwell and quadratic cases have the crucial property that the collision rates can be expressed entirely in terms of mass, momentum and energy, which are all conserved during collisions, leading to closed form expressions for the gelation time and other quantities of interest, and we restrict ourselves to this setting. To avoid triviality, we assume throughout that $\kappa + \gamma > 0$.

Formally, we write $\sim_t$ for the equivalence relation on $\{1, 2, \ldots, N\}$ generated by the relation containing pairs $(i, j)$ such that molecules $i$ and $j$ have collided at, or before, time $t$. The (Kac-$N$) interaction clusters at time $t$ are the equivalence classes of $\sim_t$. If $I$ and $J$ are two distinct, unordered interaction clusters, the instantaneous merger rate is

$$
\frac{2}{N} \sum_{i \in I} \sum_{j \in J} (\kappa + \gamma|v_i^N(t) - v_j^N(t)|^2)
$$

$$
= \frac{2}{N} \kappa(\#I)(\#J) + \frac{2}{N} \gamma(\#I) \left( \sum_{j \in J} |v_j^N(t)|^2 \right)
$$
(4)

$$
+ \frac{2}{N} \gamma(\#J) \left( \sum_{i \in I} |v_i^N(t)|^2 \right) - \frac{4}{N} \gamma \left( \sum_{i \in I} v_i^N(t) \right) \cdot \left( \sum_{j \in J} v_j^N(t) \right).
$$

From (4) one sees that Kac interaction cluster dynamics do not require a full knowledge of the Kac process, but only the number of molecules, combined momentum and combined kinetic energy of the molecules indexed by each interaction cluster; for $t \geq 0$, let $I_j^N(t) : 1 \leq j \leq k^N(t)$ be an enumeration of the equivalence classes of
\( ~t, \) and \( x_j^N(t) \) be the associated data of the clusters
\[
x_j^N(t) = \sum_{i \in \mathcal{I}_j^N(t)} \left( 1, v_i^N(t), \frac{1}{2} |v_i^N(t)|^2 \right).
\] (5)

Thanks to the conservation properties, these quantities are added when two particles belonging to distinct clusters collide, and no change occurs when if the two particles belong to the same cluster. Therefore, the interaction cluster size dynamics can be replicated by a system of coagulating particles \((x_j^N(t) : j \leq k_N(t))\) in the state space
\[
S = \{ x = (n, p, e) \in \mathbb{N} \times \mathbb{R}^d \times [0, \infty) : |p| \leq \sqrt{2ne} \}. 
\] (6)

We study the associated empirical measure, given by
\[
\mu_t^N = \frac{1}{N} \sum_{j=1}^{k_N(t)} \delta_{x_j^N(t)}. 
\] (7)

We equip the state space \( S \) with the maps \( \pi_n, \pi_p, \pi_e \) for the projection onto the respective factors, and \( R : S \to S \) for the parity map \((n, p, e) \mapsto (n, -p, e)\). Rewriting the calculation \((4)\) in this notation, the rate at which unordered pairs of particles \( \{ x, y \} \) in \( S \) merge to form a new particle in \( A \subset S \) is \( 2K(x, y, A)/N \) with
\[
K(x, y, dz) = \overline{K}(x, y) \delta_{x+y}(dz); 
\] (8)
\[
\overline{K}(x, y) = \kappa \pi_n(x) \pi_n(y) + 2\gamma (\pi_n(x) \pi_e(y) - \pi_p(x) \pi_p(y) + \pi_e(x) \pi_n(y)). 
\] (9)

This is a Marcus–Lushnikov coagulation process \([25]\) on \( S \), which we will refer to as the \textit{stochastic coagulant}. Note that a \( 1/N \) scaling of the pair interaction rate is used, which ensures that each molecule has a total collision rate that is of order 1; this is to be expected when modelling a gas where the mean free time is a positive real number. Dividing jump rates by \( N \) is equivalent to accelerating time by the same factor and this alternative formulation means that the jump rates in the definition of the “stochastic coalescent” in \([4]\) as well as of the “stochastic \( K \)-coagulant” in \([31]\) omit the \( 1/N \) from the rates and rescale time when taking the \( N \to \infty \) limit.

Finally, we remark that the definitions \((7, 8, 9)\) make sense in the more general case when the number of particles in the underlying Kac process is replaced by a random number \( l_N \) of the same order as \( N \); in this case, the scaling factors \( N, N^{-1} \) in \((7, 8)\) remain fixed. This mild generalisation will be helpful for our arguments.

### 1.1.2 Limiting kinetic equations

We now consider various forms of the limiting Smoluchowski equation. Define a drift operator \( L \), by specifying for all bounded measurable \( f : S \to \mathbb{R} \),
\[
\langle f, L(\mu) \rangle = \frac{1}{2} \int_{S^2} \{ f(x + y) - f(x) - f(y) \} \overline{K}(x, y) \mu(dx)\mu(dy). 
\] (10)

The associated evolution equation for \((\mu_t)_{t < T}\) is that, for all \( t < T \),
\[
\mu_t = \mu_0 + \int_0^t L(\mu_s)ds. 
\] (E-G)

Following \([31]\), we say that a family \((\mu_t)_{t < T}\) of positive measures is a solution to (E-G) if the following hold:

i). For all Borel sets \( A \subset S \), the map \( t \mapsto \mu_t(A) \) is measurable;

ii). For all bounded, measurable functions \( f : S \to \mathbb{R}_+ \) of compact support, \( \langle f, \mu_0 \rangle < \infty \);
For all compact subsets $S' \subset S$ and all $t < T$,
\[
\int_0^t ds \int_{S' \times S} K(x,y)\mu_s(dx)\mu_s(dy) < \infty; \tag{11}
\]

iv). For all bounded, compactly supported functions $f : S \to \mathbb{R}$, and $t < T$,
\[
\langle f, \mu_t \rangle = \langle f, \mu_0 \rangle + \int_0^t \langle f, L_s \rangle ds \tag{12}
\]

This captures the effects of coagulations between finite clusters. However, as discussed above, we wish to include the possibility of a macroscopic component, which we term gel. To include this effect, we define a modified drift operator $L_g(\mu_t)$ by specifying, for bounded, measurable $f : S \to \mathbb{R}$,
\[
\langle f, L_g(\mu_t) \rangle = \langle f, L(\mu_t) \rangle - \int_S f(x)K(x,y)\mu_t(dx)(\mu_0 - \mu_t)(dy). \tag{13}
\]

A coagulant is then a solution $(\mu_t)_{t < T}$ to
\[
\mu_t = \mu_0 + \int_0^t L_g(\mu_s)ds. \tag{E+G}
\]

Here, the additional term comes into play only after $\mu_t$ ceases to conserve the quantities $\langle \pi_n,\mu_t \rangle$, $\langle \pi_p,\mu_t \rangle$, $\langle \pi_e,\mu_t \rangle$, and the extra term represents the interaction with the gel. This may be interpreted concretely in a similar sense to (E-G) above. This generalises the Smoluchowski coagulation equations [36] analogous to that of Flory [39]. The solution to this deterministic evolution problem is a ‘$K$-coagulant’ in the language of [31].

We write
\[
g_t = (M_t,P_t,E_t) = \langle x,\mu_0 - \mu_t \rangle = (\langle \pi_n,\mu_0 - \mu_t \rangle, \langle \pi_p,\mu_0 - \mu_t \rangle, \langle \pi_e,\mu_0 - \mu_t \rangle) \tag{14}
\]

for the mass, energy and momentum of the gel. Following remarks in [31], one may show that if $\mu_t$ is a solution to (E+G), then the maps $t \mapsto (\langle \pi_n,\mu_t \rangle, (\langle \pi_e,\mu_t \rangle$ are non-increasing, which guarantees that $M_t, E_t \geq 0$. We write $S_g$ for the continuum analogue of the state space $S$, given by
\[
S_g = [0,\infty) \times \mathbb{R}^d \times [0,\infty) \tag{15}
\]

and use the same notation $\pi_n, \pi_p, \pi_e$ for the projections onto the factors, as for $S$. When $x \in S$ and $g \in S_g$, we use $K(x,g)$ for the rate of absorption, given by (9) with the new meanings of $\pi_n(g), \pi_p(g), \pi_e(g)$. We will also write $\varphi$ for the linear combination $\varphi = \pi_n + \pi_e$, defined on both $S$ and $S_g$.

\textbf{Definition 1.1} (Conservative Solutions). As noted above, the functions $t \mapsto (\langle \pi_n,\mu_t \rangle$ and $t \mapsto (\langle \pi_e,\mu_t \rangle$ are non-increasing, whenever $(\mu_t)_{t < T}$ is a local solution to either (E-G) or (E+G). We say that a solution $(\mu_t)_{t < T}$ is conservative if both are constant on $[0,T)$, or equivalently, if $(\varphi,\mu_t)$ is constant on $[0,T)$.

Thus, any solution to (E+G) is conservative up to some time $0 \leq t_g \leq \infty$, and non-conservative thereafter.

\textbf{Definition 1.2} (Metrisation of Convergence). Let $M = M_{\leq 1}(S)$ be the space of measures on $S$ with total mass at most 1. We equip $M$ with the vague topology $\mathcal{F}(M,C_c(S))$ induced by continuous, compactly supported functions on $S$, and fix a complete metric $d_0$ compatible with this topology. Let $M^*$ be the space $M \times S_g$, and define the complete, separable metric
\[
d^*((\mu,g),(\mu',g')) := d_0(\mu,\mu') + |g-g'| \tag{16}
\]
where $|\cdot|$ is the Euclidean distance on $S_g \subset \mathbb{R}^{d+2}$. 
1.2 Statement of Results

We make the following hypotheses on the initial data $\mu_0$.

A1. The initial data $\mu_0 \in \mathcal{M}$ is equal to its pushforward under $R$, that is, $\mu_0 = \mu_0 \circ R^{-1}$.

A2. The initial data is given in terms of a sub-probability measure $m$ of particle velocities, by pushforward under the map

$$\iota : \mathbb{R}^d \to S; \quad v \mapsto (1, v, |v|^2).$$

A3. For $0 \leq k \leq 6$, we have

$$\sigma_k(m) = \langle |v|^k, m \rangle < \infty.$$ (18)

A4. The underlying measure $m$ has $m(\{0\}) = 0$.

We remark that much of our analysis can be done without the assumption (A2.), which corresponds to monodisperse initial conditions where each cluster is initially a single particle. However, this assumption is natural in the context of Kac interaction clusters, as it guarantees that the number of particles $N$ initially in the stochastic coagulant corresponds to the number of particles in the underlying Kac process.

We summarise our results on the analysis of the Smoluchowski equation (E+G) as follows.

**Theorem 1.1.** Let $\mu_0$ be an initial measure on $S$ satisfying (A1-4.), and assume that $m$ is a probability measure. Then the equation (E+G) has a unique solution $(\mu_t)_{t \geq 0}$ starting at $\mu_0$; we write $g_t = (M_t, P_t, E_t)$ for the gel data defined in (14). This solution has the following properties.

1. Phase Transition. Let $t_g$ be the first time at which the solution $\mu_t$ fails to be conservative, that is:

$$t_g := \inf \{t \geq 0 : \langle \varphi, \mu_t \rangle < \langle \varphi, \mu_0 \rangle \}.$$ (19)

Then $t_g \in (0, \infty)$, and can be given explicitly in terms of the moments of $m$ as

$$t_g = \frac{1}{\kappa + 2\gamma \sigma_2(m) + \sqrt{(\kappa + 2\gamma \sigma_2(m))^2 + 4\gamma^2(\sigma_4(m) - \sigma_2^2(m))}}.$$ (20)

2. Behaviour of the Second Moment. Consider the second moment

$$\mathcal{E}(t) = \langle \varphi^2, \mu_t \rangle.$$ (21)

Then

i). $\mathcal{E}(t)$ is finite and continuous, and so locally bounded, on $[0, \infty) \setminus \{t_g\}$.

ii). On $[0, t_g)$, $\mathcal{E}$ is monotonically increasing.

iii). At the gelation time, $\mathcal{E}(t_g) = \infty$, and $\mathcal{E}(t) \to \infty$ as $t \to t_g$.

3. Representation of Gel Data. Let $m$ be the underlying distribution of initial velocities, as in (A2.) For each $t \geq 0$, there exists a unique maximal pair $c_t = (a_t, b_t) \geq 0$ such that, for all $v \in \mathbb{R}^d$,

$$a_t + b_t |v|^2 = 2t \int_{\mathbb{R}^d} (1 - e^{-a_t - b_t |w|^2})(\kappa + \gamma |v - w|^2)m(dw).$$ (22)

c_t undergoes a phase transition at time $t_g$: if $t \leq t_g$, then $c_t = 0$, and if $t > t_g$ then $a_t > 0$. If, in addition, $\gamma > 0$, then if $t > t_g$ then both components $a_t, b_t > 0$. Moreover, the map $t \mapsto c_t$ is continuous.

The gel data are given in terms of $c_t$ by

$$g_t = (M_t, P_t, E_t) = \int_{\mathbb{R}^d} \left(1, 0, \frac{1}{2} |v|^2 \right) (1 - e^{-a_t - b_t |v|^2})m(dw).$$ (23)

Therefore, if $t > t_g$ then both $M_t, E_t > 0$. Moreover, the map $t \mapsto g_t$ is continuous, and $g_{t_g} = 0$.
4. Gel Dynamics. The map $t \mapsto g_t$ is differentiable on $t \in (t_\ast, \infty)$, and

$$
\frac{d}{dt} M_t = \kappa \langle \pi_n^2, \mu_t \rangle M_t + 2\gamma \langle \pi_n \pi_e, \mu_t \rangle M_t + \langle \pi_n^2, \mu_t \rangle E_t; \quad (24)
$$

$$
\frac{d}{dt} E_t = \kappa \langle \pi_n \pi_e, \mu_t \rangle M_t + 2\gamma \langle \pi_e^2, \mu_t \rangle M_t + \langle \pi_n \pi_e, \mu_t \rangle E_t. \quad (25)
$$

5. Order of the Phase Transition, and the Size-Biasing Effect. The map $t \mapsto c_t = (a_t, b_t)$ is right-differentiable at $t_\ast$, and $a'_t > 0$. The ratio of the components is

$$
\frac{b'_t}{a'_t} = \lambda = \frac{\sqrt{\kappa^2 + 4\gamma(\kappa \sigma_2(m) + \gamma \sigma_4(m))} - \kappa}{2(\kappa \sigma_2(m) + \gamma \sigma_4(m))}. \quad (26)
$$

As a consequence, the phase transition is first order; that is, the right-derivative of the gel data $M_t, E_t$ exists and is positive at $t_\ast$:

$$
M'_t > 0; \quad E'_t > 0 \quad (27)
$$

and we have the following size-biasing effect, which further quantifies the way in which gelation is driven by the fast particles:

$$
\lim_{t \uparrow t_\ast} \frac{E_t}{M_t} = \frac{1}{2} \frac{\sigma_2(m) + \lambda \sigma_4(m)}{1 + \sigma_2(m) \lambda}. \quad (28)
$$

In particular, unless $\gamma = 0$ or $|v|$ is constant $m$-almost everywhere, we have a positive bias

$$
\lim_{t \uparrow t_\ast} \frac{E_t}{M_t} > \frac{1}{2} \int_{\mathbb{R}^d} |v|^2 m(dv). \quad (29)
$$

A notable case of physical interest is when we take $m$ to be a Gaussian density $N_d(0, \sigma^2 I)$, which certainly satisfies (A1-4.). In this case, the gelation time evaluates to

$$
\tau = \frac{1}{\kappa + 2\gamma d\sigma^2 + \sqrt{2(\kappa + 2\gamma d\sigma^2)^2 + 8\gamma^2 d\sigma^2}}. \quad (30)
$$

We also prove the following convergence theorem, relating the stochastic coagulant to the solution of the limit equation. Firstly, following ideas of [31], we show that the empirical measure $\mu^N_t$ converges to the limiting solution $(\mu_t)_{t \geq 0}$ in the vague topology, uniformly in time.

One might also ask about the connection between the stochastic coagulants and the limiting gel $(g_t)_{t \geq 0}$; since each stochastic coagulant preserves $\langle \varphi, \mu^N_t \rangle$, there is no natural analogue of the representation (14). However, we will show that the limiting gel $(g_t)_{t \geq 0}$ is closely related to the data of the largest (by particle number) cluster. In particular, the convergence of this stochastic gel may be viewed as a phase transition at time $t_\ast$.

**Theorem 1.2.** Let $\mu_0$ be a probability measure on $S$ satisfying (A1-4.), for some probability measure $m$, and let $(\mu_t)_{t \geq 0}, (g_t)_{t \geq 0}$ be the associated solution to $(E+G)$ and corresponding gel. For $N \geq 1$, let $\mu^N_t$ be the stochastic coagulant, where the initial velocities of particles are sampled independently from $m$. Define

$$
g^N_t = (M^N_t, P^N_t, E^N_t) \quad (31)
$$

as the data of the largest cluster in the stochastic system, normalised by $N^{-1}$. Then we have the convergence

$$
\sup_{t \geq 0} d^* ((\mu^N_t, g^N_t), (\mu_t, g_t)) \rightarrow 0 \quad (32)
$$

in probability. In particular, we have the following phase transition:
i). If \( t \leq t_g \), then the largest cluster has size of the order \( o_p(N) \);

ii). If \( t > t_g \), the largest cluster has size of the order \( \Theta_p(N) \).

Moreover, if \( \xi_N \) is any sequence with \( \xi_N \to \infty \) and \( \frac{\xi_N}{N} \to 0 \), then we may define \( \tilde{g}_t^N = (\tilde{M}_t^N, \tilde{P}_t^N, \tilde{E}_t^N) \) by summing the data of all clusters with mass at least \( \xi_N \), and normalising by \( N \). Then the same result holds when we replace \( g_t^N \) by \( \tilde{g}_t^N \) in (32).

Here, and throughout, we use the notation \( o_p(\cdot), O_p(\cdot), \Theta_p(\cdot) \) for the probabilistic equivalents of \( o(\cdot), O(\cdot), \Theta(\cdot) \). Precise definitions can be found in [17].

1.3 Connection of results to the Literature

We briefly discuss Theorems 1.1 and 1.2 in relation to some other results in the literature.

The starting point for this project was the conjecture raised in [33] that gelation occurs before or at the mean free time \( t_{mf} \). In [33], this is supported by numerical evidence for a variety of collision kernels, including the hard spheres and quadratic cases. For the case of the quadratic kernel (3) considered above, the mean free time is

\[
t_{mf} = \frac{1}{2\kappa + 4\gamma \sigma_2(m)}.
\]

(33)

The first item Theorem 1.1 therefore gives a positive, analytical solution to this conjecture. Moreover, except in the special cases where \( \gamma = 0 \) or \( |v| \) is constant \( m \)-almost everywhere, we have the strict inequality \( t_g < t_{mf} \).

This potentially counterintuitive result may be understood as saying that large velocity tails of \( m \) reduce the gelation time more than they reduce the mean-free time; heuristically, gelation is driven by the fastest particles.

Together with Theorem 1.2, item 2 of Theorem 1.1 follows exactly the idea of Lushnikov: the formation of a giant particle at \( t_g \) corresponds exactly to blowup of the second moment \( E(t) \) and breakdown of conservation. We go further, and show that the second moment is finite after \( t_g \), since the giant particle does not correspond to anything in the limit measure. Indeed, the only time when the second moment diverges is at the critical point \( t_g \) when a giant particle is about to form; following the percolation literature, this may be thought of as an ‘incipient giant component’ [3].

Theorem 1.2, concerning the convergence of the stochastic coagulant and the giant particle, follows ideas of [31, Theorem 4.1]. However, in our case, we can use the uniqueness statement from Theorem 1.1 to conclude a local uniform convergence result in Lemma 3.1. The remainder of this theorem follows from a careful analysis of the gel through an associated random graphs process in Sections 9, 11.

Following well-known results on Erdős-Renyi graphs (see, for example, [6]), we could further ask about the size of the largest component at and below criticality, and of the small components above criticality. The results of [7] address both of these results for certain classes of graph processes, but unfortunately the results do not cover our kernel in general.

1.4 A Note on Generalisations.

As commented above, our analysis rests on the bilinear form of the total rate \( K \), which allows us to connect the Smoluchowski equation to random graphs in Sections 4, 5. This motivates the following definition.

**Definition 1.3.** For any measurable space \( S \), we say that a kernel \( K \) on \( S \times S \times S \) is bilinear if there exists a finite set of measurable mappings \( (\pi_i)_{i \leq n+m} \) and a symmetric real matrix \( (a_{ij})_{i,j \leq n+m} \) satisfying the following.

i). For all \( i \leq n+m \) and all \( x, y \in S \),

\[
\pi_i = \pi_i(x) + \pi_i(y) \quad K(x, y, \cdot) - \text{almost everywhere.}
\]

(34)
ii). For \( i \leq n \), the map \( \pi_i : S \rightarrow \mathbb{R} \) takes only nonnegative values.

iii). There exists a measurable map \( R : S \rightarrow S \) such that \( R \circ R \) is the identity on \( S \), and

\[
\pi_i \circ R = \begin{cases} 
\pi_i, & 1 \leq i \leq n; \\
-\pi_i, & n + 1 \leq i \leq n + m.
\end{cases}
\] (35)

iv). There exists a constant \( C \) such that, for all \( x \in S \),

\[
\sum_{i>n} \pi_i(x)^2 \leq C \sum_{i \leq n} \pi_i(x)^2.
\] (36)

v). For all \( i, j \leq n, a_{ij} \geq 0 \), and there is at least one pair \( i, j \leq n \) such that \( a_{ij} > 0 \).

vi). For all \( x, y \in S \), the total rate \( \overline{K}(x,y) = K(x,y,S) \) may be expressed as

\[
\overline{K}(x,y) = \sum_{i,j \leq n+m} a_{ij} \pi_i(x) \pi_j(y).
\] (37)

In this setting, we define \( \varphi := \sum_{i \leq n} \pi_i \), and seek processes of measures \( (\mu_t)_{t \geq 0} \) on \( S \) solving the equation analogous to (E+G). Similarly, one can consider a stochastic coagulant \( (\mu^0_t)_{t \geq 0} \) defined analogously to (8, 9), with \( \lfloor N\mu_0(S) \rfloor \) particles initially sampled independently from \( \mu_0(S)^{-1} \mu_0(\cdot) \). In this context, we would require the following conditions on the initial data \( \mu_0 \):

(A1'). The measure \( \mu_0 \) is invariant under the transformation \( R \) in point iii): \( \mu_0 \circ R^{-1} = \mu_0 \).

(A3'). For all \( i \leq n \), we have \( \langle \pi_i^3, \mu_0 \rangle < \infty \).

(A4'). For all \( i \leq n \),

\[
\mu_0(x \in S : \pi_i(x) = 0) = 0.
\] (38)

In this setting, no assumption analogous to (A2.) is necessary. As remarked above this assumption is included in Theorems 1.1, 1.2 in order to guarantee that the number of particles initially in the stochastic coagulant coincides with the number of particles in the underlying Kac process.

Under these hypotheses, all of the arguments used in this paper may be adapted to prove results analogous to Theorems 1.1, 1.2. In this context, a closed form expression for the gelation time \( t_g \), right-derivative \( \lim_{t \downarrow t_g} \frac{\mu_i}{t-t_g} \), or quantities analogous to the bias \( \lim_{t \downarrow t_g} \frac{\overline{K}_t}{\overline{M}_t} \) may not be available, but these will instead be characterised in terms of finite-dimensional, explicit eigenvalue problems, and can therefore be found numerically if an explicit value is desired.

1.5 Plan of the Paper.

Our programme will be as follows.

1. In Section 2, we will prove that the limiting equation (E+G) has unique, globally defined solutions, based on a truncation argument from [30, 31].

2. In Section 3, we prove an initial result, Lemma 3.1, on the convergence of the stochastic coagulant, using the ideas of [31, Theorem 4.1]. This will later be used to prove later points of Theorem 1.1 based on probabilistic arguments.

3. In Sections 4, 5, we introduce the theory of inhomogeneous random graphs set out in [7], and show how a particular example of these graphs may be coupled to the stochastic coagulant. The critical time \( t_c \) for these graphs may be found exactly, leading to the explicit expression in Theorem 1.1.
4. A weakness of the preceding sections is that, a priori, the critical time $t_c$ for the graph processes may differ from the gelation time $t_g$; in Section 6, we show that this cannot happen. This is based on a preliminary version of Theorem 1.2, which shows convergence of $(\mu^N_t, \varphi^N_t)$ at a single fixed time $t \geq 0$.

5. Section 7 is dedicated to a proof of item 2 of Theorem 1.1. The statements about the subcritical regime follow general ideas in [30, 31], while statements about the critical and supercritical cases use additional ideas from the theory of random graphs.

6. Section 8 uses the ideas of previous sections to prove items 3 and 4 of Theorem 1.1, concerning the gel data $g_t$ beyond the critical point.

7. Section 9 uses the analysis of the gel to extend Lemma 3.1 to show that convergence is uniform in time.

8. Section 10 proves item 5 of Theorem 1.1, concerning the behaviour near the critical point. This completes the proof of this theorem.

9. To finish the proof of Theorem 1.2, we revisit the ideas of Section 6 to prove convergence of the stochastic gel $g^N_t$, uniformly in time. This is the focus of Section 11, and builds further on ideas of previous sections.

2 Well-Posedness of the Limiting Equation

This chapter is dedicated to a first analysis of the Smoluchowski equations (E-G, E+G), following Norris [30, 31]. Our goal in this section is to prove the following lemma on the well-posedness of (E+G).

**Lemma 2.1.** For any measure $\mu_0 \in \mathcal{M}$ satisfying (A1.), the equation with gel (E+G) has a unique global solution $(\mu_t)_{t \geq 0}$ starting at $\mu_0$. Moreover, the momentum $P_t = 0$ for all time $t \geq 0$.

**Corollary 2.2.** Suppose $(\mu'_t)_{t < T}$ is a conservative local solution to the equation without gel, (E-G), starting at $\mu_0$. Then $\mu_t = \mu'_t$ for all $t < T$, and $T < t_g$. Hence, (E-G) has a unique maximal conservative solution, given by $(\mu_t)_{t < t_g}$.

Our proof of Lemma 2.1 is an adaptation of the arguments in [30, Section 2] and [31, Section 2] and is based on a truncation argument. Recalling that $\varphi = \pi_n + \pi_e$, we see that $\mathbf{K}(x, y) \leq \Delta \varphi(x)\varphi(y)$ for some $\Delta = \Delta(\kappa, \gamma)$. For all $\xi > 0$, we define the truncated particle space

$$S_\xi = \{x \in S : \varphi(x) \leq \xi\}. \quad (39)$$

We consider the following ‘truncation at level $\xi$’: in the empirical measure, we track only those particles inside $S_\xi$, and consider all other particles to belong to a ‘truncated gel’. Although the particles in the truncated gel affect the dynamics in $S_\xi$, these contributions depend only on the total mass, momentum and energy $g^\xi$ of the truncated gel, due to the bilinear form of the kernel. This leads to an ordinary differential equation with Lipschitz coefficients in an infinite dimensional space.

We formalise this intuition as follows. For a measure $\mu^\xi$ supported on $S_\xi$ and $g^\xi \in S_\xi$, we define a signed measure $L^\xi_g(\mu^\xi, g^\xi)$ on $S_\xi$ by specifying, for all $f \in C_c(S)$,

$$\langle f, L^\xi_g(\mu^\xi, g^\xi) \rangle = \frac{1}{2} \int_{S^\xi_\xi} [f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y)] \mathbf{K}(x, y)\mu^\xi(dx)\mu^\xi(dy) \quad (40)$$

$$- \int_{S_\xi} f(x) \mathbf{K}(x, g^\xi)\mu^\xi(dx).$$
This corresponds to the dynamics of particles inside $S_\xi$. The rate of change of the truncated gel data is given by

$$\hat{L}_g^{\xi}(\mu^{\xi}, g^{\xi}) = \frac{1}{2} \int_{S_\xi} (x + y) 1[\varphi(x + y) > \xi |K(x, y|d\mu^{\xi}(dy)
+ \int_{S_\xi} xK(x, g^{\xi})d\mu^{\xi}(dx).$$

(41)

We now seek measures $\mu_t^{\xi}$ supported on $S_\xi$ and gel data $g_t^{\xi} = (M_t^{\xi}, P_t^{\xi}, E_t^{\xi}) \in S_\xi$ such that, for all bounded measurable $f$ on $S_\xi$,

$$\langle f, \mu_t^{\xi} \rangle = \langle f, \mu_0^{\xi} \rangle + \int_0^t \langle f, L_g^{\xi}(\mu_s^{\xi}, g_s^{\xi}) \rangle ds;$$

$$g_t^{\xi} = g_0^{\xi} + \int_0^t \hat{L}_g^{\xi}(\mu_s^{\xi}, g_s^{\xi})ds.$$  

(42)

We will use the following existence and uniqueness result for the restricted dynamics ($E_t^{\xi, 1}$, $E_t^{\xi, 2}$).

**Lemma 2.3.** [Existence and Uniqueness of Restricted Dynamics] Suppose $\mu_0^{\xi}$ is a finite measure on $S_\xi$ which satisfies (A1.), and $g_0^{\xi} = (M_0^{\xi}, 0, E_0^{\xi}) \in S_\xi$. Then there exists a unique map $(\mu_t^{\xi}, g_t^{\xi})$ on $[0, \infty)$, which solves the restricted dynamics ($E_t^{\xi, 1}$, $E_t^{\xi, 2}$). Moreover, for all $t \leq T$, $\mu_t^{\xi}$ is a positive, finite measure on $S_\xi$, $P_t^{\xi} = 0$ and $g_t^{\xi} \in S_\xi$.

**Sketch Proof of Lemma 2.3.** This may be proved by a trivial modification of the arguments in [30, Proposition 2.2]. We define Picard iterates $(\mu_t^{\xi, n}, g_t^{\xi, n}) : n \geq 0, t \geq 0)$ by

$$\langle \mu_t^{\xi, 0}, g_t^{\xi, 0} \rangle = (\mu_0^{\xi}, g_0^{\xi});$$

$$\langle \mu_t^{\xi, n+1}, g_t^{\xi, n+1} \rangle = (\mu_t^{\xi, n}, g_t^{\xi, n}) + \int_0^t \langle L_g^{\xi}(\mu_s^{\xi, n}, g_s^{\xi, n}), (\mu_s^{\xi, n}, g_s^{\xi, n}) \rangle ds.$$  

(43)

One then uses bilinear continuity arguments in total variation norm $|| \cdot ||$ to show that, given a bound $\langle \varphi, \mu_0^{\xi} \rangle + M_0^{\xi} + E_0^{\xi} \leq C$, there is a positive time $T = T(\xi, C) > 0$ such that the Picard iterates $(\mu_t^{\xi, n})_{t \leq T}$ converge uniformly in total variation on $[0, T]$, and that the limit $\mu_T^{\xi}$ solves ($E_t^{\xi, 1}$, $E_t^{\xi, 2}$), possibly allowing $\mu_t^{\xi}$ to be a signed measure. This argument also implies that the solution is unique on this interval. Now, we note that the quantity $\langle \varphi, \mu_T^{\xi} \rangle + M_T^{\xi} + E_T^{\xi}$ is constant in time, and therefore this construction can be repeated on $[T, 2T]$, $[2T, 3T]$, etc., which proves global existence and uniqueness. Finally, an integrating factor is introduced to argue that $\mu_T^{\xi}$ is a positive measure. In our case, it is also straightforward to see that the gel data $M_T^{\xi}, E_T^{\xi} \geq 0$, and that $P_T^{\xi} = 0$ from the symmetry (A1.)

**Proof of Lemma 2.1.** We first show existence. For all $\xi < \infty$, we let $(\mu_t^{\xi}, g_t^{\xi})$ be the solution to the dynamics ($E_t^{\xi, 1}$, $E_t^{\xi, 2}$) restricted to $S_\xi$, with initial data

$$\mu_0^{\xi}(dx) = 1[x \in S_\xi] \mu_0(dx); \quad g_0^{\xi} = \int_{x \notin S_\xi} x \mu_0(dx).$$

(44)

Observe that, if $\xi < \xi'$, then $\mu_t^{\xi}, g_t^{\xi}$ given by

$$\tilde{\mu_t}^{\xi}(dx) = 1_{x \in S_\xi} \mu_t^{\xi'}(dx); \quad \tilde{g_t}^{\xi} = g_t^{\xi'} + \int_{x \in S_\xi \setminus \tilde{S}_\xi} x \mu_t^{\xi'}(dx)$$

(45)

solve the dynamics ($E_t^{\xi, 1}$, $E_t^{\xi, 2}$) with the same initial data $(\mu_0^{\xi}, g_0^{\xi})$. From uniqueness in Lemma 2.3, it follows that $\tilde{\mu_t}^{\xi} = \mu_t^{\xi}$; $\tilde{g_t}^{\xi} = g_t^{\xi}$. This shows that the measures $\mu_t^{\xi}$ are increasing in $\xi$, while the gel data $M_t^{\xi}, E_t^{\xi}$ are decreasing, and $P_t^{\xi}$ is identically 0, by symmetry (A1.). Therefore, the limits

$$\mu_t = \lim_{\xi \to \infty} \mu_t^{\xi}; \quad M_t = \lim_{\xi \to \infty} M_t^{\xi}; \quad E_t = \lim_{\xi \to \infty} E_t^{\xi}$$

(46)
exist in the sense of monotone limits; one can then check that \( \mu_t \) and \( g_t = (M_t, 0, E_t) \) satisfy the full equation \((E+G)\), with initial values \( \mu_0 \) and \( g_0 = 0 \).

To see uniqueness let \( \mu_t \) be the solution constructed above, and write \( g_t = (M_t, P_t, E_t) \) for the data of the gel. Let \( \tilde{\mu}_t \) be any solution to \((E+G)\) starting at \( \mu_0 \), and let \( \tilde{g}_t = (M_t', P_t', E_t') \) be the associated data of the gel. For all \( \xi < \infty \), it is simple to verify that

\[
\tilde{\mu}_t^\xi(dx) = 1_{x \in S_\xi} \mu_t'(dx); \quad \tilde{g}_t^\xi = \tilde{g}_t + \int_{S_\xi} x \tilde{\mu}_t(dx)
\]

is a solution to the dynamics \((E_1^\xi, E_2^\xi)\) on \( S_\xi \). By uniqueness in Lemma 2.3, it follows that \( \tilde{\mu}_t^\xi = \mu_t' \), and taking monotone limits, we see that \( \mu'_t = \lim_{\xi \to \infty} \tilde{\mu}_t^\xi = \lim_{\xi \to \infty} \mu_t^\xi = \mu_t \). The argument for \( \tilde{g} \) is identical.

3 Convergence of the Stochastic Coagulant

We now turn to a preliminary version of Theorem 1.2. In this section, we will prove local uniform convergence of the stochastic coagulant \( \mu_t^N \) to a solution \( \mu_t \) of \((E+G)\); throughout, we fix \( \mu_0 \) satisfying (A1-4.). As mentioned in the introduction, we will consider a mild generalisation of the stochastic coagulant, which will be helpful for future reference: we allow initial data of the form

\[
\mu_0^N := \frac{1}{N} \sum_{i=1}^{l_N} \delta_{(1, v_i, 1/2|v_i|^2)}
\]

where the \( (v_i)_{i=1}^{l_N} \) are the initial velocities for the underlying Kac process, \( l_N \leq N \), and we ask that the following conditions hold.

B1. As \( N \to \infty \), the initial measures \( \mu_0^N = \frac{1}{N} \sum_{i \leq N} \delta_{(1, v_i, 1/2|v_i|^2)} \) converge in probability to \( \mu_0 \) in the vague topology in probability:

\[
d_0(\mu_0^N, \mu_0) \to 0 \quad \text{in probability}.
\]

B2. We also have the convergence

\[
\langle x, \mu_0^N \rangle \to \langle x, \mu_0 \rangle \quad \text{in probability}
\]

where we recall that \( x \) is the identity function on \( S \), and

\[
\sup_{N \geq 1} \mathbb{E} \langle \varphi^2, \mu_0^N \rangle < \infty.
\]

We have the following corollary of the second point of B2. For \( \eta > 0 \), let \( X_\eta^N \) be the number of particles with kinetic energy exceeding \( \eta \):

\[
X_\eta^N = \# \left\{ i \leq l_N : \frac{1}{2} |v_i|^2 > \eta \right\}.
\]

Then, as \( \eta \to \infty \), we have the convergence

\[
\sup_{N \geq 1} \mathbb{E} \left[ \frac{1}{N} X_\eta^N \right] \to 0.
\]

For example, it is straightforward to verify that these hold in the case where \( l_N = N \), and \( v_1, ..., v_N \) are independent samples from the distribution \( m \) described in Section 1.2. However, the more general hypotheses will be useful for a ‘duality’ argument in Section 7. Our result is as follows:
Lemma 3.1. Suppose \( \mu_0 \) satisfies (A1-4.), and let \((\mu_t)_{t \geq 0}\) be the solution to (E+G) starting at \( \mu_0 \). Let \((\mu^N_t)_{t \geq 0}\) be stochastic coalescents with initial data \( \mu^N_0 \) satisfying (B1-2.). Then we have the local uniform convergence
\[
\forall t_f \geq 0 \quad \sup_{t \leq t_f} d_0(\mu^N_t, \mu_t) \to 0 \text{ in probability}
\] (54)
where recall that \( d_0 \) is a complete metric inducing the vague topology.

Remark 3.2. We will later upgrade the local uniform convergence to full uniform convergence in Lemma 9.2. We also remark that this does not immediately imply the convergence of the gel terms in Theorem 1.2, as the test functions involved are neither compactly supported nor even bounded. This will be dealt with in Sections 6, 11, where the proofs build on this result.

The proof proceeds as follows, based on the well known method of proving tightness and identifying possible limit paths. Firstly, we will argue that for any \( t_f \geq 0 \), the processes \((\mu^N_t)_{0 \leq t \leq t_f}\) are tight in the Skorohod topology of \( \mathbb{D}([0, t_f], (\mathcal{M}, d_0)) \). Then, we will argue that if \( \mu^\xi \) is any subsequential limit point, then for \( \xi \) satisfying a certain regularity condition (C1-2.), the restricted measures \( \mu^\xi_0(dx) = 1_{\xi} \mu^\xi_0 dx \) solve the restricted dynamics \((E^1_\xi, E^2_\xi)\) and that \( \mu_0^0 = \mu^\xi_0 \). By uniqueness in Lemma 2.3, it follows that \( \mu^\xi \) coincides with the solution \( \mu^\xi \) found in Lemma 2.3. For any subsequential limit \( \xi \), the required regularity condition holds for sufficiently many cutoff \( \xi \) to allow a limit \( \xi \to \infty \), to conclude that \( (\mu^\xi_t)_{0 \leq t \leq t_f} = (\mu_t)_{0 \leq t \leq t_f} \). Writing \( d_{Sk} \) for a complete Skorohod metric on \( \mathbb{D}([0, t_f], (\mathcal{M}, d_0)) \), this implies
\[
d_{Sk}((\mu^N_t)_{0 \leq t \leq t_f}, (\mu_s)_{0 \leq t \leq t_f}) \to 0 \quad \text{in probability.} 
\] (55)
Since the limit process \((\mu_t)_{0 \leq t \leq t_f}\) is continuous in the vague topology \((\mathcal{M}, d_0)\), it follows that we may upgrade to uniform convergence:
\[
\sup_{0 \leq t \leq t_f} d_0(\mu^N_t, \mu_t) \to 0 \quad \text{in probability}
\] (56)
as claimed.

We first prove some tightness results for the processes \((\mu^N_t)_{t \leq t_f}\).

Lemma 3.3. Suppose conditions (B1-3.) hold. Then

i). There process \( \mu^N \) has at most \( N - 1 \) jumps in \([0, \infty)\);

ii). the jump rates of \( \mu^N \) are bounded by
\[
\lambda^N := N \left\{ \kappa \langle \pi_n, \mu^N_0 \rangle^2 + 8\gamma \langle \pi_n, \mu^N_0 \rangle \langle \pi_e, \mu^N_0 \rangle \right\},
\] (57)

which is constant in time, and \( \lambda^N/N \in \mathcal{O}_p(1) \).

Proof. For the first claim, we note that after \( l_N - 1 \leq N - 1 \) jumps, precisely one cluster remains and so the system undergoes no further changes. The second claim is a simple calculation using the conservation laws for \( \pi_n \) and \( \pi_e \). The claim that \( \lambda^N/N \in \mathcal{O}_p(1) \) follows from (B2.).

Lemma 3.4. Suppose (B1-2.) hold, and let \( t_f > 0 \). Then the distributions of \( \mu^N \) are tight in the space of probability measures on \( \mathbb{D}([0, t_f]; (\mathcal{M}, d_0)) \).

Proof. First, we consider the real valued processes \( \langle f, \mu^N_t \rangle \) for \( f \in C_c(S) \). Compact containment is immediate, since \( \| \langle f, \mu^N_t \rangle \| \leq \| f \|_\infty \). For \( \delta > 0 \), write \( (I_i : i = 1, \ldots, \lfloor t_f/\delta \rfloor) \) for a partition of \([0, t_f]\) into intervals of length at most \( \delta \). Recalling from Lemma 3.3 that the rescaled jump rate \( \lambda^N/N \) is \( \mathcal{O}_p(1) \), it follows that, for all \( \epsilon > 0 \),
\[
\sup_{N \geq 1} \mathbb{P} \left( \mu^N \text{ makes at least } \frac{\epsilon N}{3} \text{ jumps in any } I_i : i \leq \lfloor t_f/\delta \rfloor \right) \to 0
\] (58)
as \( \delta \to 0 \). Now, since the jumps are bounded by \( \frac{3}{2} \| f \|_{\infty} / N \), for any \( \epsilon > 0 \) it follows that

\[
\sup_{N \geq 1} P (\exists s, t \leq t_1 : |s - t| < \delta, |\langle f, \mu^N_t - \mu^N_s \rangle| > \epsilon) \to 0
\]

(59)
as \( \delta \to 0 \). Therefore, for any fixed \( f \in C_c(S) \), we may apply [12, Chapter 3, Theorem 8.6c] to see that the processes \( \langle f, \mu^N_t \rangle_{0 \leq t \leq t_1} \) are tight.

We now use a general result [16, Theorem 4.6] on tightness in Skorohod spaces. Compact containment of \( \mu^N_t \) in \( (\mathcal{M}, d_0) \) is immediate, since \( \mathcal{M} \) is itself compact, and it is straightforward that the maps \( \mu \mapsto \langle f, \mu \rangle : f \in C_c(S) \) on \( (\mathcal{M}, d_0) \) satisfy the requirements of the cited theorem. Therefore, the result of the previous paragraph shows that the processes \( \mu^N \) are tight in \( D([0, t_1], (\mathcal{M}, d_0)) \), as claimed.

\[\Box\]

Having proven tightness, we now turn to the identification of possible limit paths. For \( N \in \mathbb{N} \) and \( \xi \in \mathbb{R}_+ \), we consider the processes \( \mu^N, \mu^{N, \xi} \in D([0, \infty); (\mathcal{M}, d_0)) \), where \( \mu^{N, \xi} = \mu^N 1_{S_\xi} \). We write \( g^{N, \xi} \in D([0, \infty); S_\xi) \) for the data of the gel cutoff at this level:

\[
g^{N, \xi}_t = \langle x, \mu^N_t - \mu^{N, \xi}_t \rangle
\]

(60)
where we recall that \( x \) denotes the identity function on \( S \), so that \( \langle \mu^{N, \xi}, g^{N, \xi} \rangle \in D([0, \infty); (\mathcal{M}^*, d^*) \). We first identify the martingale parts of the processes \( \langle f, \mu^{N, \xi}_t \rangle \) and \( g^{N, \xi}_t \). For \( \xi > 0 \) and \( f \in C_c(S) \), let

\[
\mathfrak{M}^{f, N, \xi}_t := \langle f, \mu^{N, \xi}_t \rangle - \langle f, \mu^N_t \rangle - \int_0^t \int_{S_\xi} f(x) \overline{K}(x, \mu^{N, \xi}_s dx) ds
\]

\[\] - \frac{1}{2} \int_0^t \int_{S_\xi} [f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y)] \overline{K}(x, y) \mu^{N, \xi}_s dx) \mu^{N, \xi}_s dy) ds
\]

(61)
and

\[
\mathfrak{M}^{g, N, \xi}_t := g^{N, \xi}_t - g^N_t - \int_0^t \int_{S_\xi} x \overline{K}(x, g^{N, \xi}_s dx) ds
\]

\[\] - \frac{1}{2} \int_0^t \int_{S_\xi} (x + y)1[\varphi(x + y) \leq \xi] \overline{K}(x, y) \mu^{N, \xi}_s dx) \mu^{N, \xi}_s dy) ds
\]

(62)
and

\[
\mathfrak{M}^{g, N, \xi}_t := g^{N, \xi}_t - g^N_t - \int_0^t \int_{S_\xi} x \overline{K}(x, g^{N, \xi}_s dx) ds
\]

\[\] + \frac{1}{2} \int_0^t \int_{S_\xi} 2x1[\varphi(2x) \leq \xi] \overline{K}(x, x) \mu^{N, \xi}_s dx) ds.

Lemma 3.5. Suppose conditions (B1-2.) hold, and write \( (\mathcal{F}^N_t)_{t \geq 0} \) for the natural filtration of \( \langle \mu^N_t \rangle_{t \geq 0} \). Then, for all \( \xi \in (0, \infty) \) and \( f \in C_c(S) \), the processes \( \mathfrak{M}^{f, N, \xi}_t, \mathfrak{M}^{N, \xi}_t \) are martingales with respect to \( (\mathcal{F}^N_t)_{t \geq 0} \). The quadratic variations are bounded by

\[
\left[ \langle f, \mu^{N, \xi}_t \rangle \right]_t = \left[ \mathfrak{M}^{f, N, \xi}_t \right]_t \leq \frac{9\| f \|_{\infty}^2}{N} \xi, \quad \left[ g^{N, \xi}_t \right]_t = \left[ \mathfrak{M}^{g, N, \xi}_t \right]_t \leq \frac{2\xi}{N} \left[ \varphi, \mu^N_0 \right]
\]

(63)
uniformly in \( t \in [0, \infty) \) and \( \xi \in [0, \infty) \).

Proof. The martingale property is a standard; see, for example, [12, Chapter 4, Proposition 1.7] or [20, Lemma 19.21]. Since \( \mathfrak{M}^{f, N, \xi} \) and \( \mathfrak{M}^{N, \xi} \) are pure jump processes, the quadratic variation is the sum of the squares of the jumps, which can be bounded using Lemma 3.3 for the first claim and by straightforward maximisation arguments for the second.

\[\Box\]

Next, we discuss a regularity condition for potential cutoff values \( \xi \). This will allow us to deduce vague convergence of the truncated measures \( \mu^{N, \xi}_t = \mu^N 1_{S_\xi} \), despite the discontinuity of the cutoff \( 1_{S_\xi} \).
Lemma 3.6 (Regularity Condition). Suppose that $\mu^N$ are stochastic coagulants satisfying (B1-2.), and that some subsequence $(\mu^N_{r})_{r \geq 1}$ of $(\mu^N)_{N \geq 1}$ converges in distribution to a random variable $\overline{\nu}$ in $\mathbb{D}([0,t],(\mathcal{M},d_0))$. Then, for almost all $\xi > 0$, the following hold:

i). Almost surely, for almost all $t \leq t_f$,
\[ \overline{\mu}_t \left( \{ x : \varphi(x) = \xi \} \right) + \overline{\mu}_t \otimes \overline{\mu}_t \left( \{ (x,y) : \varphi(x+y) = \xi \} \right) = 0; \quad (C1.) \]

ii). This also holds for $t = 0$. That is, almost surely,
\[ \overline{\mu}_0 \left( \{ x : \varphi(x) = \xi \} \right) + \overline{\mu}_0 \otimes \overline{\mu}_0 \left( \{ (x,y) : \varphi(x+y) = \xi \} \right) = 0. \quad (C2.) \]

Proof. Since $\mathbb{D}([0,t],(\mathcal{M},d_0))$ is separable, we may use the Skorohod representation theorem to realise all $\mu^N, \overline{\nu}$ on a common probability space, such that $\mu^N_r \rightarrow \overline{\nu}$ almost surely in the Skorohod topology of $\mathbb{D}([0,t],(\mathcal{M},d_0))$, and such that $\langle \varphi, \mu^N_0 \rangle \rightarrow \langle \varphi, \mu_0 \rangle$ almost surely. In particular, $\langle \varphi, \mu^N_0 \rangle$ are almost surely bounded.

We first consider the case of nonrandom $\nu^N_r \rightarrow \nu$ in $\mathbb{D}([0,t],(\mathcal{M},d_0))$, and such that, for some constant $C$,
\[ \sup_{N \geq 0} \sup_{t \leq t_f} \langle \varphi, \nu^N_t \rangle \leq C < \infty. \quad (64) \]

Let $\Psi$ be the measure on $[0,\infty)$ given by
\[ \Psi(A) = \nu_0(x : \varphi(x) \in A) + \nu_0 \otimes \nu_0((x,y) : \varphi(x+y) \in A) \]
\[ + \int_0^{t_f} (\nu_t(x : \varphi(x) \in A) + \nu_t \otimes \nu_t((x,y) : \varphi(x+y) \in A)) dt. \quad (65) \]

Since $\nu_t$ is a subprobability measure for all $t \leq t_f$, $\Psi$ has total mass at most $2(1+t_f)$, and so, for all but countably many $\xi \geq 0$, $\Psi(\{\xi\}) = 0$. For such $\xi$, we have the desired properties that
\[ \nu_0(x : \varphi(x) = \xi) + \nu_0 \otimes \nu_0((x,y) : \varphi(x+y) = \xi) = 0; \quad (66) \]
\[ \nu_t(x : \varphi(x) = \xi) + \nu_t \otimes \nu_t((x,y) : \varphi(x+y) = \xi) = 0 \quad \text{for almost all } t \leq t_f. \quad (67) \]

We now show how this may be extended to the case of random $\mu^N, \overline{\nu}$ such that $\mu^N_r \rightarrow \overline{\nu}$ almost surely in $\mathbb{D}([0,t],(\mathcal{M},d_0))$. Let $\Psi$ be the random measure in (65) corresponding to $\overline{\nu}$, and let $A(\overline{\nu})$ be the random set
\[ A(\overline{\nu}) = \{ \xi \geq 1 : \Psi(\{\xi\}) > 0 \}. \quad (68) \]

The argument above shows that, almost surely, $\int_{A(\overline{\nu})} d\xi = 0$; by Fubini, this implies that
\[ \int_1^\infty \mathbb{P}(\xi \in A(\overline{\nu})) d\xi = 0. \quad (69) \]

Therefore, for almost all $\xi \geq 1$, $\mathbb{P}(\xi \in A(\overline{\nu})) = 0$. For all such $\xi$, the following hold almost surely:
\[ \overline{\mu}_0(x : \varphi(x) = \xi) + \overline{\mu}_0 \otimes \overline{\mu}_0((x,y) : \varphi(x+y) = \xi) = 0; \quad (70) \]
\[ \overline{\mu}_t(x : \varphi(x) = \xi) + \overline{\mu}_t \otimes \overline{\mu}_t((x,y) : \varphi(x+y) = \xi) = 0 \quad \text{for almost all } t \leq t_f. \quad (71) \]

This corresponds precisely to the desired conditions (C1-2.).

This regularity condition now allows us to deduce weak convergence of the cutoff measures $\mu^N_{r,\xi}$.

Lemma 3.7. Suppose that $\mu^N$ are stochastic coagulants satisfying (B1-2.), and that a subsequence $\mu^{N_r}$ converges in distribution to a random variable $\overline{\nu}$ in $\mathbb{D}([0,t],(\mathcal{M},d_0))$. If $\xi$ is such that (C1-2.) hold for $\overline{\nu}$, then writing $\overline{\mu}_t^\xi$ for the cutoff measures $\overline{\mu}_t^\xi = \overline{\nu} 1_{S_t}$, we have the following convergences in distribution as $r \rightarrow \infty$. 

\[ \int_1^\infty \mathbb{P}(\xi \in A(\overline{\nu})) d\xi = 0. \quad (69) \]
i). For any $f \in C_c(S)$ and $0 \leq t \leq t_f$,

$$
\int_0^t \int_{S^2} \left[ f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y) \right] \bar{K}(x, y) \mu_s^{N_r, \xi}(dx)\mu_s^{N_r, \xi}(dy)ds
\rightarrow \int_0^t \int_{S^2} \left[ f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y) \right] \bar{K}(x, y)\bar{\mu}_s^{\xi}(dx)\bar{\mu}_s^{\xi}(dy)ds; \quad (72)
$$

and

$$
\int_0^t \int_{S^2} \left( x + y \right)1[\varphi(x + y) > \xi] \bar{K}(x, y) \mu_s^{N_r, \xi}(dx)\mu_s^{N_r, \xi}(dy)ds
\rightarrow \int_0^t \int_{S^2} \left( x + y \right)1[\varphi(x + y) > \xi] \bar{K}(x, y)\bar{\mu}_s^{\xi}(dx)\bar{\mu}_s^{\xi}(dy)ds. \quad (73)
$$

ii). We write $\bar{\mu}_s^{\xi}$ for the data of the truncated gel corresponding to $\bar{\mu}$:

$$
\bar{\mu}_s^{\xi} = \left\langle x, \mu_0 - \bar{\mu}_s^{\xi} \right\rangle.
$$

Then, for all $f \in C_c(S)$ and $t \leq t_f$,

$$
\int_0^t \int_{S^2} f(x) \bar{K}(x, g_s^{N_r, \xi})\mu_s^{N_r, \xi}(dx)ds \rightarrow \int_0^t \int_{S^2} f(x) \bar{K}(x, \bar{\mu}_s^{\xi})\bar{\mu}_s^{\xi}(dx)ds; \quad (75)
$$

$$
\int_0^t \int_{S^2} x \bar{K}(x, g_s^{N_r, \xi})\mu_s^{N_r, \xi}(dx)ds \rightarrow \int_0^t \int_{S^2} x \bar{K}(x, \bar{\mu}_s^{\xi})\bar{\mu}_s^{\xi}(dx)ds. \quad (76)
$$

iii). Let $L_s^{\xi}, \bar{L}_s^{\xi}$ be the truncated drift operators in $(40, 41)$, and write $\mu_s^{\xi}, g_s^{\xi}$ for the cutoff measure and gel corresponding to the base measure $\mu_0$. Then, for all $t \leq t_f$ and $f \in C_c(S)$,

$$
\int_0^t \left[ \left\langle f, \mu_s^{N_r, \xi} \right\rangle - \left\langle f, \mu_0^{N_r, \xi} \right\rangle - \int_0^s \left\langle f, \bar{L}_s^{\xi} \left( \mu_u^{N_r, \xi}, g_u^{N_r, \xi} \right) \right\rangle du \right] ds
\rightarrow \int_0^t \left[ \left\langle f, \bar{\mu}_s^{\xi} \right\rangle - \left\langle f, \mu_s^{\xi} \right\rangle - \int_0^s \left\langle f, L_s^{\xi} \left( \bar{\mu}_u^{\xi}, \bar{g}_u^{\xi} \right) \right\rangle du \right] ds \quad (77)
$$

and

$$
\int_0^t \left[ g_s^{N_r, \xi} - g_0^{N_r, \xi} - \int_0^s \bar{L}_s^{\xi} \left( \mu_u^{N_r, \xi}, g_u^{N_r, \xi} \right) du \right] ds
\rightarrow \int_0^t \left[ \bar{g}_s^{\xi} - g_0^{\xi} - \int_0^s \bar{L}_s^{\xi} \left( \bar{\mu}_u^{\xi}, \bar{g}_u^{\xi} \right) du \right] ds. \quad (78)
$$

Proof. As in the previous lemma, the Skorohod representation theorem shows that we may realise $\mu^N, \bar{\mu}$ on a common probability space, such that $\mu^N \rightarrow \mu$ almost surely in the Skorohod topology of $\mathbb{D}([0, t_f], (\mathcal{M}, d_0))$. Moreover, using (B1-2.) as above, $\left\langle \varphi, \mu_t^{N_r} \right\rangle$ is almost surely bounded, uniformly in $r \geq 1$ and $t \leq t_f$, and

$$
d_0(\mu_0^{N_r}, \mu_0) \rightarrow 0; \quad \left\langle x, \mu_0^{N_r} \right\rangle \rightarrow \left\langle x, \mu_0 \right\rangle \quad (79)
$$

almost surely. Since the map $(\mu_t)_{t \leq t_f} \rightarrow \mu_0$ is continuous in the Skorohod topology of $\mathbb{D}([0, t_f], (\mathcal{M}, d_0))$, the first convergence displayed above implies that $\bar{\mu}_0 = \mu_0$ almost surely. It is therefore sufficient to prove the results for the special case of nonrandom $\nu^N$, such that $\nu^N \rightarrow \nu$ almost surely, and such that the preceding results hold with $\nu^N, \nu$ in place of $\mu^N, \bar{\mu}$. We will write $\nu_t^{N_r, \xi}, g_t^{N_r, \xi}$ for the cutoff measures, and $\bar{g}_t^{\xi}, g_t^{\xi}$ for the corresponding cutoff gel, defined as above with $\nu^N, \nu$ in place of $\mu^N, \bar{\mu}$. 

16
For part i), Fix $f \in C_c(S)$ and $\xi \in [0, \infty)$. The functions

$$(x, y) \mapsto f(x)K(x, y)1[\varphi(x) \leq \xi, \varphi(y) \leq \xi]$$

and

$$(x, y) \mapsto f(x + y)K(x, y)1[\varphi(x + y) \leq \xi]$$

are compactly supported, and continuous away from the exceptional set

$$E_\xi = \{(x, y) \in S^2 : \varphi(x) = \xi \text{ or } \varphi(y) = \xi \text{ or } \varphi(x + y) = \xi\}.$$  

From condition (C1.) it follows that, for almost all $t \leq t_1$, $(\nu_t \otimes \nu_t)(E_\xi) = 0$. Since $d_0(\nu_t^N, \nu_t) \to 0$ for all but at most countably many $t \leq t_1$, it follows that

$$\int_{S^2} [f(x) + f(y)] K(x, y)\nu_t^{N_r, \xi}(dx)\nu_t^{N_r, \xi}(dy) \to 0$$

$$\int_{S^2} f(x)1[\varphi(x) \leq \xi]K(x, y)\nu_t^{N_r, \xi}(dx)\nu_t^{N_r, \xi}(dy) \to 0$$

for almost all $t \leq t_1$. Recalling that $K(x, y) \leq \Delta \varphi(x)\varphi(y)$ for some constant $\Delta = \Delta(\kappa, \gamma)$, we bound

$$\left|\int_{S^2} [f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y)] K(x, y)\nu_t^{N_r, \xi}(dx)\nu_t^{N_r, \xi}(dy)\right|$$

$$\leq 3\Delta \|f\|_\infty \langle \varphi, \nu_0^N \rangle^2$$

and so, by bounded convergence, for all $t \leq t_1$,

$$\int_0^t \int_{S^2} [f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y)] K(x, y)\nu_s^{N_r, \xi}(dx)\nu_s^{N_r, \xi}(dy)ds$$

$$\to \int_0^t \int_{S^2} [f(x + y)1[\varphi(x + y) \leq \xi] - f(x) - f(y)] K(x, y)\nu_s^{\xi}(dx)\nu_s^{\xi}(dy)ds$$

which proves the first claim. For the second claim, we note that $\nu_t^{N_r, \xi}, \nu_t^\xi$ are supported on $S_\xi$, and for all $x, y \in S_\xi$ we have the bound $|x|, |y| \leq C\xi$, for some constant $C$; the second claim now follows.

For part ii), we first claim that for $t = 0$ and almost all $t \leq t_1$, $g_t^{N_r, \xi} \to g_0^\xi$. To see this, we recall that $\langle x, \nu_0^N \rangle \to \langle x, \nu_0 \rangle = \langle x, \mu_0 \rangle$ by construction, and using conditions (C1-2.), for $t = 0$ and almost all $t \leq t_1$,

$$\langle x, \nu_t^{N_r, \xi} \rangle = \langle x1_{S_\xi}, \nu_t^{N_r} \rangle \to \langle x1_{S_\xi}, \nu_t \rangle = \langle x, \nu_t^\xi \rangle.$$  

This follows because $x1_{S_\xi}$ is compactly supported, and except for exceptional times $t$, $x1_{S_\xi}$ is continuous except on a set of $\nu_t$-measure 0. Since $\nu_0 = \mu_0$, this also implies that

$$g_0^{N_r, \xi} \to \langle x, \mu_0 - \mu_0^\xi \rangle = g_0^\xi$$

and we may therefore use $g_0^\xi$ in place of $g_0^{N_r, \xi}$.

We also observe that $\varphi(g_t^{N_r, \xi}) \leq \langle \varphi, \nu_t^N \rangle$ is bounded uniformly in $t \leq t_1$ and in $r$. The argument is now essentially identical to point i). above.
For the first claim of item iii), we first note that the the two parts above show that, for all \( t \leq t_f \),
\[
\int_0^t \langle f, L_s^\xi (\nu_s^{N_r,\xi}, g_s^{N_r,\xi}) \rangle \, ds \to \int_0^t \langle f, L_s^\xi (\tbar{\mu}_s^{\xi}, \tbar{g}_s^{\xi}) \rangle \, ds.
\] (89)
Moreover, following (85), we have the uniform bound, for all \( u \leq t_f \),
\[
\bigg| \langle f, L_u^\xi (\nu_u^{N_r,\xi}, g_u^{N_r,\xi}) \rangle \bigg| \leq 4\|f\|_\infty \langle \varphi, \nu_0^{N_r} \rangle^2.
\] (90)
It therefore follows, by bounded convergence, that for all \( t \),
\[
\int_0^t \int_0^s \langle f, L_s^\xi (\nu_s^{N_r,\xi}, g_s^{N_r,\xi}) \rangle \, duds \to \int_0^t \int_0^s \langle f, L_s^\xi (\tbar{\mu}_s^{\xi}, \tbar{g}_s^{\xi}) \rangle \, duds.
\] (91)
For the other terms, we note that \( d_0(\nu_0^{N_r}, \nu_0) \to 0 \), and that \( d_0(\nu_t^{N_r}, \nu_t) \to 0 \) for almost all \( t \leq t_f \). Therefore, by (C1-2.) and bounded convergence,
\[
\int_0^t (f, \nu_s^{N_r,\xi}) \, ds \to \int_0^t (f, \nu_0) \, ds; \quad (f, \nu_0^{N_r,\xi}) \to (f, \nu_0^\xi) = (f, \mu_0^\xi).
\] (92)
Combined with (91), these prove the first claim. An identical argument proves the second claim.

We will now show that for any subsequential limit point \( \tbar{\mu} \), the pair \((\tbar{\mu}^\xi, \tbar{\nu}^\xi)\) solves the restricted dynamics \((E|_{\xi}, E|_{\xi}^2)\) whenever \( \xi \) satisfies (C1-2.).

**Lemma 3.8.** Suppose conditions (B1-2.) hold, and that some subsequence \((\mu_r^{N_r})_{r \geq 1}\) of \((\mu_r^{N})_{N \geq 1}\) converges in distribution to a random variable \( \tbar{\mu} \) in \( D([0, t_f], (\mathcal{M}, d_0)) \). Suppose also that \( \xi \) is such that (C1-2.) hold for \( \tbar{\mu} \).
Write \( \tbar{\mu}^\xi \) for the truncated measures and \( \tbar{\nu}^\xi \) for the truncated gel, as above, and similarly \( \mu_0^\xi, g_0^\xi \). Then, almost surely, for all \( f \in C_c(S) \) and all \( t \leq t_f \),
\[
\langle f, \tbar{\mu}^\xi \rangle - \langle f, \mu_0^\xi \rangle - \int_0^t \langle f, L_s^\xi (\tbar{\mu}^{\xi}, \tbar{g}^{\xi}) \rangle \, ds = 0;
\] (93)
\[
\tbar{\mu}^\xi - \mu_0^\xi - \int_0^t L_s^\xi (\tbar{\mu}^{\xi}, \tbar{g}^{\xi}) \, ds = 0.
\] (94)

It follows that, almost surely, \((\tbar{\mu}^\xi, \tbar{\nu}^\xi)_{0 \leq t \leq t_f}\) is the unique solution \((\mu_t^\xi, g_t^\xi)_{0 \leq t \leq t_f}\) to \((E|_{\xi}, E|_{\xi}^2)\) started at \((\mu_0^\xi, g_0^\xi)\).

In particular, \((\mu_0^\xi, g_0^\xi)_{0 \leq t \leq t_f}\) is deterministic.

**Proof.** Fix \( f \in C_c(S) \). We first estimate the martingale terms \( \mathfrak{M}_t^{f,N,\xi}, \mathfrak{M}_t^{N,\xi} \). From [20, Thrm. 26.12] and Lemma 3.5, there exists a constant \( c = c(\kappa, \gamma) > 0 \) such that
\[
\mathbb{E} \left[ \sup_{t \in [0, \infty]} \left( \mathfrak{M}_t^{f,N,\xi} \right)^2 \right] \leq c \mathbb{E} \left[ \sup_{t \in [0, \infty]} \left( \mathfrak{M}_t^{f,N} \right)^2 \right] \leq \frac{9c\|f\|_\infty^2}{N}. \] (95)
It follows that, for all \( t \leq t_f \),
\[
\int_0^t M_s^{f,N,\xi} \, ds \to 0
\] (96)
in distribution. We now estimate the self-interaction term: for some other constant \( c \), we bound
\[
\left| \frac{1}{2N} \int_0^t \int_S |f(2x)1[\varphi(2x) \leq \xi] - 2f(x)| \mathcal{K}(x, x) \mu_s^{N,\xi} (dx) \, ds \right| \leq \frac{c}{N} \|f\varphi^2\|_\infty.
\] (97)
Therefore, from (61), it follows that for all \( t \leq t_f \),
\[
\int_0^t \left[ \langle f, \mu_s^{N,\xi} \rangle - \langle f, \mu_0^{N,\xi} \rangle - \int_0^s \langle f, L_s^\xi (\mu_s^{N,\xi}, g_s^{N,\xi}) \rangle \, du \right] \, ds \to 0
\] (98)
in distribution. Comparing this to item iii). of Lemma 3.7, we identify the two limits to conclude that almost surely, for all \( t \leq t_f \),
\[
\int_0^t \left[ \langle f, \mu_s^\xi \rangle - \langle f, \mu_0^\xi \rangle - \int_0^s \langle f, L_s^\xi \mu_s \rangle \right] ds = 0.
\]  
(99)

Since \( t \leq t_f \) is arbitrary and the integrand is right-continuous, this implies that, almost surely, for all \( t \leq t_f \),
\[
\langle f, \mu_t^\xi \rangle - \langle f, \mu_0^\xi \rangle - \int_0^t \langle f, L_s^\xi \mu_s \rangle = 0.
\]  
(100)

Taking an intersection over \( f \) belonging to a countable dense subset of \((C_c(S), \| \cdot \|_\infty)\) proves the first claim. The argument for the second claim is identical. It follows that \((\mu^\xi_t, \mu_t^\xi)_{0 \leq t \leq t_f}\) satisfies \((E_1^\xi, E_2^\xi)\) almost surely, which implies that \((\mu^\xi_t, \mu_t^\xi)_{0 \leq t \leq t_f}\) is characterized almost surely, by uniqueness in Lemma 2.3.

We may now prove Lemma 3.1, using the well-known combination of tightness and the identification of the limit. Tightness of the processes \( \mu^N_t \) was proven in Lemma 3.4, and so it is sufficient to characterize possible limit paths. Suppose a subsequence \((\mu^N_t)_{0 \leq t \leq t_f}\) converges in distribution to a limit \((\mu_t)_{0 \leq t \leq t_f}\). From Lemma 3.6, there is an unbounded set \( \mathcal{X} \subset [1, \infty) \) such that \((C1-2.)\) hold for each \( \xi \in \mathcal{X} \). By Lemma 3.8, it follows that the limit process \((\mu^\xi_t, \mu_t^\xi)_{0 \leq t \leq t_f}\) is, almost surely, the measure part \((\mu_t)_{0 \leq t \leq t_f}\) of the unique solution to \((E_1^\xi, E_2^\xi)\) which starts at \( \mu_01\xi_t \), for each such \( \xi \). Moreover, we recall from Lemma 2.1 that the full solution \((\mu_t)_{0 \leq t \leq t_f}\) to \((E+G)\) is characterized by
\[
\mu_t = \lim_{\xi \uparrow \infty} \mu_t^\xi
\]  
(101)
in the sense of monotone limits. Therefore, almost surely, for all \( t \leq t_f \),
\[
\mu_t = \lim_{\xi \uparrow \infty; \xi \in \mathcal{X}} \mu_t^\xi = \mu_t^\xi
\]  
(102)

Therefore, the only possible subsequential limit point is the deterministic path \((\mu_t)_{0 \leq t \leq t_f}\); together with tightness, this implies that \((\mu^N_t)_{0 \leq t \leq t_f}\rightarrow(\mu_t)_{0 \leq t \leq t_f}\) in distribution in the Skorohod topology. Since the limit path is nonrandom, it follows that, for all \( \epsilon > 0 \),
\[
\mathbb{P} \left( d_{Sk}((\mu^N_t)_{0 \leq t \leq t_f}, (\mu_t)_{0 \leq t \leq t_f}) > \epsilon \right) \rightarrow 0
\]  
(103)

where we recall that \( d_{Sk} \) is a complete metric compatible with the Skorohod topology.

Returning to the cutoff dynamics in \((E_1^\xi, E_2^\xi)\), one can see that each cutoff solution \((\mu_t^\xi)_{0 \leq t \leq t_f} = (\mu_t1\xi_t)_{0 \leq t \leq t_f}\) is continuous in total variation norm. Now, if \( f \in C_c(S) \), one may choose \( \xi \) such that \( \text{supp}(f) \subset S_\xi \), so that for all \( t \leq t_f \),
\[
\langle f, \mu_t^\xi \rangle = \langle f, \mu_t^\xi \rangle.
\]  
(104)

It follows that each process \( \langle f, \mu_t \rangle \) is continuous for \( f \in C_c(S) \), which implies that \( \mu_t \) is continuous in the metric \( d_0 \). Therefore, every uniformly open neighbourhood of \((\mu_t)_{0 \leq t \leq t_f}\) contains a Skorohod-open neighbourhood of \((\mu_t)_{0 \leq t \leq t_f}\), and together with (103), this implies that, for all \( \epsilon > 0 \),
\[
\mathbb{P} \left( \sup_{t \leq t_f} d_0(\mu_t^N, \mu_t) > \epsilon \right) \rightarrow 0
\]  
(105)
as desired.

4 Introduction to Inhomogenous Random Graphs

As discussed in the introduction, the connection between gelation and random graphs is well-understood, and the multiplicative kernel corresponds to the well-known Erdős-Rényi random graphs \([13, 11, 4]\). However, for our purposes, not all particles are equal: particles with large velocities will undergo more collisions and
exhibit quantitatively different behaviour if $\gamma > 0$, and so we will need a more sophisticated model of random graphs to accommodate this inhomogeneity. In this section, we will review the theory of inhomogenous random graphs developed in [7], which will play the same rôle for our model that the Erdős-Réyni model does for the multiplicative kernel. We will introduce the relevant definitions and state, without proof, the main results of [7] which we use in our work.

Definition 4.1. A generalised vertex space is a triple $\mathcal{V} = (\mathcal{S}, m, (v_N)_{N \geq 1})$, consisting of

- A separable metric space $\mathcal{S}$, equipped with its Borel $\sigma$-algebra;
- A measure $m$ on $\mathcal{S}$, with $m(\mathcal{S}) \in (0, \infty)$;
- A family of random variables $v_N = (v_1^{(N)}, \ldots, v_N^{(N)})$ taking values in $\mathcal{S}$, and of potentially random length $l_N$, such that the empirical measures

$$m_N = \frac{1}{N} \sum_{k=1}^{l_N} \delta_{v_k^{(N)}}$$

(106)

converge to $m$ in the weak topology $\mathcal{F}(C_b(\mathcal{S}))$, in probability.

In the special case where $m(\mathcal{S}) = 1$ and $l_N = N$, we say that $(\mathcal{S}, m, (v_N)_{N \geq 1})$ is a vertex space.

Definition 4.2. A kernel is a symmetric, measurable map $k : \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$. We say that $k$ is positive if, for all distinct $x, y \in \mathcal{S}$, $k(x, y) > 0$.

This is sufficient for our results, and is a slightly requirement stronger than the notion of irreducibility in [7]; this difference allows us to eliminate the possibility of exceptional sets of measure 0.

Definition 4.3 (Inhomogenous random graphs). Given a kernel $k$ and a generalised vertex space $\mathcal{V}$, we let $G^N$ be a random graph on $\{1, 2, \ldots, N\}$ given as follows. Conditional on the values of $v_N$, the edge $e = (ij)$ is included with probability

$$p_{ij} = 1 - \exp \left( - \frac{k(v_i^{(N)}, v_j^{(N)})}{N} \right)$$

(107)

and such that the presence of different edges is (conditionally) independent. We write $G^N \sim \mathcal{G}^V(N, k)$. We also consider the vertex data $v_N = (v_i^{(N)})_{i=1}^N$ to be part of the data of $G_i^N$, so that an equality of random graphs $G = G'$ includes the equality of the vertex data.

To treat a general class of kernels $k$, additional regularity is required, to prevent pathologies. This is the content of the following definition:

Definition 4.4 (Graphical Kernel). We say that a kernel $k$ on a vertex space $\mathcal{V} = (\mathcal{S}, m, (v_N)_{N \geq 1})$ is graphical if the following hold.

i). $k$ is almost everywhere continuous on $\mathcal{S} \times \mathcal{S}$;
ii). $k \in L^1(\mathcal{S} \times \mathcal{S}, m \times m)$;
iii). If $G^N \sim \mathcal{G}^V(N, k)$, then

$$\frac{1}{N} \mathbb{E} [e(G^N)] \rightarrow \frac{1}{2} \int_{\mathcal{S} \times \mathcal{S}} k(v, w) m(dv)m(dw)$$

(108)

where $e(\cdot)$ denotes the number of edges of the graph.

Definition 4.5. Given a graph $G$, we write $C_j(G) : j = 1, 2, \ldots$ for the connected components of $G$, in decreasing order of their sizes $\#C_j(G) = C_j(G)$. If there are fewer than $j$ connected components, then $C_j(G) = \emptyset$ and $C_j(G) = 0$. 

20
The phase transition is given in terms of the convolution operator

\[(Tf)(v) = \int_{\mathbb{R}^d} k(v, w)f(w)m(dw)\]  \hspace{1cm} (109)

for functions \(f\) such that the right-hand side is defined (i.e., finite or \(+\infty\)) for \(m\)-almost all \(v\); for instance, if \(f \geq 0\) then \(Tf\) is well-defined, possibly taking the value \(\infty\). We define

\[\|T\| = \sup\{\|Tf\|_{L^2(m)} : \|f\|_{L^2(m)} \leq 1, f \geq 0\}.\]  \hspace{1cm} (110)

If \(T\) defines a bounded linear map from \(L^2(m)\) to itself, then \(\|T\|\) is precisely its operator norm in this setting; otherwise, \(\|T\| = \infty\). It is straightforward to show that if \(k \in L^2(S \times S, m \times m)\) then \(T : L^2(m) \to L^2(m)\) is a Hilbert-Schmidt operator, and that \(\|T\|_{HS} = \|k\|_{L^2(m)} < \infty\). In this case, \(\|T\|\) is certainly finite, and is the operator norm of \(T : L^2(m) \to L^2(m)\). The example of interest to us will fall into this case.

The analysis of the random graphs uses a branching process, similar to that used in the standard analysis of Erdős-Rényi graphs. Many quantities of the graph can be expressed in terms of the ‘survival probability’ \(\rho(k, v)\) when the data \(v\) of the first vertex is given. To avoid the unnecessary complication of making this into a precise definition, we use the following characterisation, which is equivalent by [7, Theorem 6.2].

**Lemma 4.1.** Let \(k\) be a positive kernel on a generalised vertex space \(V\), such that \(k \in L^1(S \times S, m \times m)\), and such that, for all \(v\),

\[\int_S k(v, w)m(dw) < \infty.\]  \hspace{1cm} (111)

Consider the nonlinear fixed-point equation

\[\forall v \in S, \quad \rho(v) = 1 - e^{-(T\rho)(v)}\]  \hspace{1cm} (112)

where \(T\) is the convolution operator (109). Then (112) has a maximal solution \(\rho_k(v) = \rho(k; v)\); that is, for any other solution \(\tilde{\rho}\),

\[\forall v \in S, \quad \tilde{\rho}(v) \leq \rho(k, v).\]  \hspace{1cm} (113)

It therefore follows that \(0 \leq \rho_k(v) \leq 1\) for all \(v\). The maximal solution is necessarily unique, and so this uniquely defines \(\rho_k\). Moreover, we have the following dichotomy:

i). If \(\|T\| \leq 1\), then \(\rho_k(v) = 0\) for all \(v\);

ii). If \(\|T\| > 1\), then \(\rho_k(v) > 0\) for all \(v\).

This can be stated dynamically as follows. Consider the survival function ‘at time \(t\)’, given by \(\rho(tk, v)\), which we will write throughout as \(\rho_t(v)\). Then

- If \(t \leq \|T\|^{-1}\), then \(\rho_t(v) = 0\) for all \(v\);
- If \(t > \|T\|^{-1}\), then \(\rho_t(v) > 0\) for all \(v\).

We can now state the main results on the phase transition, given by [7, Theorem 3.1 and Corollary 3.2].

**Theorem 4.2 (Phase Transition).** Let \(k\) be a graphical and positive kernel for a vertex space \(V\), with \(0 < \|T\| < \infty\). Let \(G^N \sim \mathcal{G}^V(N, k)\) be random graphs on a common probability space. Then we have the convergence

\[\frac{1}{N}C_1(G^N_t) \to \int_S \rho(tk, v)m(dv) \quad \text{in probability.}\]  \hspace{1cm} (114)

Therefore, if \((G^N_t)_{t \geq 0}\) is a dynamic family of random graphs \(G^N_t \sim \mathcal{G}^V(N, tk)\), then we have the following dichotomy:
i). If \( t \leq \|T\|^{-1} \), then there is no giant component, in particular
\[
\frac{C_1(G^N_t)}{N} \to 0
\]
in probability.

ii). If \( t > \|T\|^{-1} \), then there is a giant component: there exists \( c = c(t) > 0 \) such that
\[
\mathbb{P}(C_1(G^N_t) > cN) \to 1.
\]

Remark 4.3. Following [7], based on this dichotomy, we say that

i). \( G^N \) is subcritical if \( \|T\| < 1 \);

ii). \( G^N \) is critical if \( \|T\| = 1 \);

iii). \( G^N \) is supercritical if \( \|T\| > 1 \).

We also have the following result, which implies the uniqueness of the giant component [7, Theorem 3.6]. This result considers clusters of a scale \( \xi_N \ll N \), excluding the largest; we term these mesoscopic clusters.

Theorem 4.4. Let \( G^N \sim G^V(N,k) \), for a (generalised) vertex space \( V \) and an positive graphical kernel \( k \). Let \( \xi_N \) be a sequence with
\[
\xi_N \to \infty; \quad \frac{\xi_N}{N} \to 0.
\]
Then
\[
\frac{1}{N} \sum_{j \geq 2, C_j(G^N) \geq \xi_N} C_j(G^N) \to 0
\]
in probability.

We will also make use of the following monotonicity and continuity properties, from [7, Theorem 6.4].

Theorem 4.5. Let \( k \) be a kernel on a vertex space \( V \), and let \( \rho_t(\cdot) = \rho(tk, \cdot) \) be the survival function defined above. Then the map \( t \mapsto \rho_t(\cdot) \) is monotonically increasing, in the sense that for all \( 0 \leq s \leq t \) and for all \( v \), \( \rho_s(v) \leq \rho_t(v) \). We also have the following continuity property. Let \( t_n \to t \) be a monotone sequence, either increasing or decreasing. Then
\[
\rho_{t_n}(v) \to \rho_t(v) \quad \text{for } m- \text{ almost all } v, \quad \text{and}
\int_S \rho_{t_n}(v)m(dv) \to \int_S \rho_t(v)m(dv).
\]

The final result which we will need is a ‘duality’ result, connecting the supercritical and subcritical behaviours. This is given by [7, Theorem 12.1].

Theorem 4.6. Let \( k \) be an positive graphical kernel on a generalised vertex space \( V \), such that \( \|T\| > 1 \). Let \( G^N \sim G^V(N,k) \), and form \( \hat{G}^N \) by deleting all vertexes in the largest component \( C_1(G^N) \). Then, defined on the same underlying probability space, there is a generalised vertex space \( \hat{V} = (S, \hat{m}, (\hat{w}_N)_{N \geq 1}) \) with
\[
\hat{m}(dv) = (1 - \rho(k;v))m(dv)
\]
and such that \( \hat{w}_N \) is an enumeration of those \( v_i \) not belonging to the component \( C_1(G^N) \), and a random graph \( \hat{G}^N \sim G^\hat{V}(N,k) \) such that
\[
\mathbb{P} (\hat{G}^N = \hat{G}^N) \to 1.
\]
Furthermore, if \( k \in L^2(S \times S, m \times m) \), then \( \hat{G}^N \) is subcritical.

We emphasise here that we have defined the equality \( \hat{G}^N = \hat{G}^N \) includes equality of the velocities \( v_i \) associated to each vertex; this follows from the construction in [7], since the velocities \( \hat{w}_N \) associated to \( \hat{G}^N \) are exactly those \( v_i \) not belonging to the giant component. This generalises the standard ‘duality result’ of Bollobás [5] for Erdős-Rényi graphs.
5 Coupling of the Stochastic Coagulant to Random Graphs

In this section, we will show that the stochastic coagulant defined in §1.1.1 may be coupled to a dynamic version of the random graphs $G^V(N, tk)$ discussed above. This allows us to apply the results quoted above to analyse the stochastic coagulant process and the limit equation.

**Definition 5.1.** [Dynamic Inhomogenous Random Graphs] Fix a measure $\mu_0$ satisfying (A1-4.). Let $v_N = (v_i, i = 1, 2, ..., l_N)$ be random velocities such that the empirical measures

$$\mu_0^N = \frac{1}{N} \sum_{i \leq l_N} \delta_{(1, v_i, \frac{1}{2}|v_i|^2)} (123)$$

satisfy the conditions (B1-2.), and sample $\tau_e \sim \text{Exponential}(1)$, independently of each other, and of $v_N$. We define the kernel

$$k(v, w) = 2K \left( \left\{ 1, v, \frac{1}{2}|v|^2 \right\}, \left\{ 1, w, \frac{1}{2}|w|^2 \right\} \right) (124)$$

where the right-hand side is the interaction kernel defined in (8). We form the random graphs $(G_t^N)_{t \geq 0}$ on $\{1, 2, ..., l_N\}$ by including the edge $e = (ij)$ if

$$t \geq \frac{N\tau_e}{k(v_i, v_j)} (125)$$

We emphasise that the $v_i$ do not change during the dynamics.

This has the following immediate consequences. Firstly, if we define $V = (\mathbb{R}^d, m, v_N)$, then $V$ is a vertex space in the sense of Definition 4.1 and, for all times $t$, $G_t^N$ is an instance of the inhomogenous random graph $G^V(N, tk)$ defined in Definition 4.3. Moreover, the process $(G_t^N)_{t \geq 0}$ is increasing, and is a Markov process, by the memoryless property of the exponential variables $\tau_e$. We write $T$ for the convolution operator and $\|T\|$ for the associated operator norm, as defined following (109). We write also $t_c$ for the critical time for the phase transition in Theorem 4.2, which is given by $\|T\|^{-1}$.

For a cluster $C$ of the graph $G_t^N$, we will write $M(C), P(C), E(C)$ to denote the unnormalised mass, momentum and energy

$$M(C) = \#C; \quad P(C) = \sum_{i \in C} v_i; \quad E(C) = \sum_{i \in C} \frac{1}{2}|v_i|^2. (126)$$

We write $\delta(C)$ for the point mass in $S$

$$\delta(C) = \delta(\mu(C), P(C), E(C)) (127)$$

and $\mu^N(G_t^N)$ for the normalised empirical measure

$$\mu^N(G_t^N) = \frac{1}{N} \sum_{\text{Clusters}} \delta(C) (128)$$

where the sum is over all clusters $C$ of $G_t^N$. This is connected to the stochastic coagulants as follows:

**Lemma 5.1** (Coupling of Random Graphs and Stochastic Coagulants). Let $(G_t^N)$ be the random graph process described in Definition 5.1. Then the processes

$$\mu_t^N = \mu^N(G_t^N) (129)$$

are stochastic coagulants for the kernel $K$.

We also note that, since the rates are bounded, the stochastic coagulant has uniqueness in law. As a consequence, if we wish to prove any property of a general stochastic coagulant $\mu_t^N$, we may assume that it is given by $\mu_t^N = \mu^N(G_t^N)$, and appeal to an analysis of the random graphs.
Sketch of proof of Lemma 5.1. One may verify that the two processes undergo the same transitions at the same rates; this essentially follows from the calculation (4). We will give here an alternative proof, which we feel offers more insight.

From the calculation (4), the rates of the stochastic coagulants do not depend on the collision kernel $B$ beyond the total rates $B(v, S^{d-1})$. In particular, if $B, \hat{B}$ are two kernels with the same total rates, then given random initial velocities $(v_i(0))_{i=1}^N$, one may couple the Kac processes $(v_i(t) : t \geq 0)_{i=1}^N, (\tilde{v}_i(t) : t \geq 0)_{i=1}^N$ so that the two processes have the same coagulation structure, and the initial velocities coincide. We also note that the total mass, momentum and energy for each cluster are the same in each model, thanks to the conservation properties of the kernel. Therefore, the processes of empirical measures $\mu^N_t$ coincide for the two models.

Now, we choose $\hat{B}$ to be the degenerate kernel such that the outgoing velocities are the same as the incoming velocities; in particular, each $v_i(t)$ is constant in time. For this case, we construct a graph on $\{1, ..., N\}$ by including $(i, j)$ if particles $i, j$ have collided before, or at, time $t$, according to the Kac process for $\hat{B}$. It is immediate that the resulting graph $G^N_t$ is distributed according to the distribution $\mathcal{G}^N(N, tk)$ constructed above, and by construction,

$$\mu^N_t = \mu^N(G^N_t). \quad (130)$$

Since this is true for a particular realisation of the two processes, we have equality in law, as claimed. $\square$

Combining this with the approximation result Lemma 3.1 for the stochastic coagulant, we may connect the random graph process to the limit equation as follows.

**Lemma 5.2** (Convergence of the Random Graphs). Let $(G^N_t)_{t \geq 0}$ be the random graph processes constructed above, such that the initial velocities $v_N = (v_1, ..., v_N)$ satisfy (B1-2.). Let $\mu_0$ be the corresponding initial measure on $S$ under (A2.), and assume that (A1-4.) hold. Let $(\mu_t)_{t \geq 0}$ be the solution to the Smoluchowski Equation $(E+G)$ starting at $\mu_0$; then we have the local uniform convergence

$$\sup_{s \leq t} d_0(\mu^N(G^N_t), \mu_t) \to 0 \quad (131)$$

in probability, for all $t < \infty$, where we recall that $d_0$ is a metric for the vague topology $\mathcal{F}(\mathcal{M}_{\leq 1}(S), C_c(S))$. We emphasise here that we do not require exactly $N$ particles, or that $\mu_0$ is a probability measure.

We can also compute the critical time associated to $G^N_t$ explicitly:

**Lemma 5.3** (Computation of critical time). Suppose condition (A3.) holds for $m$, and let $G^N_t$ be the random graphs defined above. Then $T$ is a bounded linear map from $L^2(m)$ to itself, and the critical time for the graph phase transition is

$$t_c = \left( \kappa \sigma_0(m) + 2\gamma \sigma_2(m) + \sqrt{(\kappa \sigma_0(m) + 2\gamma \sigma_2(m))^2 + 4\gamma^2(\sigma_0(m)\sigma_4(m) - \sigma_2^2(m))} \right)^{-1} \quad (132)$$

where we recall the notation $\sigma_k(m)$ for the $k$th moment $\langle |v|^k, m \rangle$.

**Remark 5.4.** We remark that, in the case where $m$ is a probability measure, this is the closed form claimed for $t_\kappa$ in Theorem 1.1. However, we have not yet established that $t_c = t_\kappa$; this is the content of Lemma 6.1.

**Proof of Lemma 5.3.** Firstly, by (A3.), it is easy to see that $k \in L^2(\mathbb{R}^d \times \mathbb{R}^d, m \times m)$. Therefore, as remarked in Section 4, $\|T\|$ is precisely the operator norm of $T : L^2(m) \to L^2(m)$.

Construct a basis $\{e_n\}_{n \geq 1}$ of $L^2(m)$, such that

$$e_i(v) = v_i, \quad i = 1, 2, ..., d; \quad (133)$$

$$e_{d+1}(v) = 1, \quad e_{d+2}(v) = |v|^2 \quad (134)$$
and such that, for \( n \geq d + 3 \), \( e_n \) is orthogonal to \( E = \text{Span}(e_1, \ldots, e_{d+2}) \). By expanding the quadratic term \( |v - w|^2 \), we see that, for all \( f \in L^2(m) \),

\[
(Tf)(v) = 2 \left[ \kappa(f, e_{d+1})e_{d+1}(v) + \gamma(f, e_{d+1})e_{d+2}(v) - 2\gamma \sum_{i=1}^{d} \langle f, e_i \rangle e_i(v) + \gamma(f, e_{d+2})e_{d+1}(v) \right]
\]  

(135)

where \( \langle \cdot, \cdot \rangle \) denotes the \( L^2(m) \) inner product. Therefore, \( T \) maps into the subspace \( E \), and \( T(e_n) = 0 \) if \( n \geq d+3 \). It follows from the definition (109) that \( T \) is self-adjoint, and so the operator norm \( ||T|| \) is given by the largest modulus of an eigenvalue of \( T|_E \). We write \( \sigma_{2,i}(m) \) for the second moment \( \sigma_{2,i}(m) = \int_{\mathbb{R}^d} v_i^2 m(dv) \). In this notation, and with respect to the basis \( \{ e_1, \ldots, e_d, e_{d+1}, e_{d+2} \} \), \( T|_E \) has the matrix representation

\[
[T|_E] = \begin{bmatrix}
-4\gamma\sigma_{2,1}(m) & 0 & \cdots & 0 & 0 & 0 \\
0 & -4\gamma\sigma_{2,2}(m) & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & -4\gamma\sigma_{2,d}(m) & 0 & 0 \\
0 & 0 & \cdots & 0 & 2\kappa\sigma_0(m) + 2\gamma\sigma_2(m) & 2\kappa\sigma_2(m) + 2\gamma\sigma_4(m) \\
0 & 0 & \cdots & 0 & 2\gamma\sigma_0(m) & 2\gamma\sigma_2(m)
\end{bmatrix}
\]  

(136)

Therefore, the eigenvalues are \(-4\gamma\sigma_{2,i}(m), i = 1, \ldots, d\), and the two roots \( \lambda_{\pm} \) to a quadratic equation which simplifies to

\[
\lambda^2 - 2\lambda(\kappa\sigma_0(m) + 2\gamma\sigma_2(m)) + 4\gamma^2(\sigma_2^2(m) - \sigma_0(m)\sigma_4(m)) = 0,
\]  

(137)

where \( \sigma_0(m), \sigma_2(m), \sigma_4(m) \) are as defined in Assumption (A3.). The two roots are

\[
\lambda_{\pm} = \kappa\sigma_0(m) + 2\gamma\sigma_2(m) \pm \sqrt{(\kappa\sigma_0(m) + 2\gamma\sigma_2(m))^2 + 4\gamma^2(\sigma_0(m)\sigma_4(m) - \sigma_2^2(m))}.
\]  

(138)

It is straightforward to check that the largest eigenvalue in modulus is \( \lambda_+ \), so we find

\[
||T|| = \lambda_+ = \kappa\sigma_0(m) + 2\gamma\sigma_2(m) + \sqrt{(\kappa\sigma_0(m) + 2\gamma\sigma_2(m))^2 + 4\gamma^2(\sigma_0(m)\sigma_4(m) - \sigma_2^2(m))}.
\]  

(139)

This gives the critical time as

\[
t_c = \left( \kappa\sigma_0(m) + 2\gamma\sigma_2(m) + \sqrt{(\kappa\sigma_0(m) + 2\gamma\sigma_2(m))^2 + 4\gamma^2(\sigma_0(m)\sigma_4(m) - \sigma_2^2(m))} \right)^{-1}
\]  

(140)

as claimed. \( \square \)

It is straightforward to show that (111) holds, and we may there define \( \rho_t \) as the survival function from Lemma 4.1, for kernel \( tk \). We note, for future use, the following properties where \( k \) is the kernel given above.

**Lemma 5.5.** The survival function \( \rho_t(v) = \rho(tk, v) \) takes the form

\[
\rho_t(v) = 1 - e^{-a_t - b_t|v|^2}
\]  

(141)

for some \( a_t, b_t \geq 0 \). Moreover, the functions \( t \mapsto a_t, t \mapsto b_t \) are continuous.

This proves the first two assertions of item 4 of Theorem 1.1.

**Proof.** Using the symmetry \( k(-v, -w) = k(v, w) \) and hypothesis (A1.), it is simple to verify that \( \tilde{\rho}(v) := \rho_t(-v) \) also satisfies the fixed point equation (112). By maximality of \( \rho_t \), we must have \( \rho_t(-v) \leq \rho_t(v) \) for all \( v \in \mathbb{R}^d \), which implies that \( \rho_t \) is an even function of \( v \in \mathbb{R}^d \).

Using the identification of the range of \( T \) as in Lemma 5.3, we see that there exist \( c_i^t : 1 \leq i \leq d + 2 \) such that

\[
(T\rho_t)(v) = \sum_{i=1}^{d} c_i^t v_i + c_{d+1}^t + c_{d+2}^t |v|^2
\]  

(142)
and therefore, from the equation (112) defining $\rho$,

$$
\rho_t(v) = 1 - \exp \left( \sum_{i=1}^{d} c_i^d v_i + c_i^{d+1} + c_i^{d+2} |v|^2 \right).
$$

(143)

Since $\rho_t$ is even, the linear term $\sum_{i \leq d} c_i^d v_i$ must identically vanish, which gives the claimed representation of $\rho_t$ by relabelling $a_t = c_t^{d+1}, b_t = c_t^{d+2}$. Since $\rho_t(v) \leq 1$ everywhere, it follows that

$$
\forall v \in \mathbb{R}^d \quad a_t + b_t |v|^2 \geq 0
$$

(144)

which is only possible if $a_t, b_t \geq 0$. The continuity follows immediately from Theorem 4.5.

\[\square\]

## 6 Equality of the Critical Times

In this section, we will prove that the critical time $t_c$ or the graph process, introduced in Section 5, coincides with the gelation time for the limiting equation, defined in Section 2 as the time at which mass and energy begin to escape to infinity.

**Lemma 6.1.** Let $\mu_0$ be a measure on $S$ satisfying (A1-4.). Let $(\mu_t)_{t \geq 0}$ be the solution to $(E+G)$ starting at $\mu_0$, with associated mass $M_t$ and energy $E_t$ of the gel; recall that $t_g$ is defined by

$$
t_g := \inf \{ t \geq 0 : \langle \varphi, \mu_t \rangle < \langle \varphi, \mu_0 \rangle \} = \inf \{ t \geq 0 : M_t + E_t > 0 \}.
$$

(145)

Let $(G_t^N)$ be the random graph processes constructed above, and suppose that (B1-2.) hold for $G_t^N, \mu_0$. Then the critical time $t_c$ for the graph transition process coincides with the gelation time $t_g$: $t_c = t_g$.

The following is a straightforward corollary.

**Corollary 6.2.** Let $\mu_0$ be a sub-probability measure on $S$ satisfying (A1-4.), for a base measure $m$. Let $(\mu_t)_{t \geq 0}$ be the solution to $(E+G)$ starting at $\mu_0$, and $t_g$ the associated gelation time. Then $t_g$ is given explicitly by the closed-form expression (132). In the case where $m$ is a probability measure, this reduces to the expression (20).

**Proof of Corollary 6.2.** Let $l_N = \lceil Nm(\mathbb{R}^d) \rceil$, and form $v_N$ by sampling $l_N$ velocities independently from $m(\cdot)/m(\mathbb{R}^d)$. It is immediate that the resulting vertex space $V = (\mathbb{R}^d, (v_N)_{N \geq 1}, m)$ satisfies (B1-2.) for the measure $\mu_0$, and the critical time $t_c$ of the associated graphs $G_t^N$ is given by the claimed expression (132). From the previous lemma, it now follows that the gelation time $t_g = t_c$, which proves the claimed result.

\[\square\]

The proof of Lemma 6.1 is based on the following weak version of the convergence of the gel in Theorem 1.2, and may be taken as preliminary reading for the more detailed arguments in Section 11.

**Lemma 6.3.** Let $(\mu_t)_{t \geq 0}, M_t, E_t$ and $G_t^N$ be as above. Fix $t > 0$, and write $M_t^N, E_t^N$ for the scaled mass and energy of the largest component of $G_t^N$, as in Theorem 1.2:

$$
M_t^N := \frac{1}{N} C_1(G_t^N); \quad E_t^N = \frac{1}{N} E(C_1(G_t^N)).
$$

(146)

Then $M_t^N \to M_t$ and $E_t^N \to E_t$ in probability.

We first show that Lemma 6.3 implies Lemma 6.1; the remainder of this section is dedicated to the proof of Lemma 6.3.
Proof of Lemma 6.1. Throughout, let \((v_i)_{i=1}^N\) be the velocities associated to the nodes of the random graph process, which we recall are independent of time.

Firstly, suppose for a contradiction that \(t_g < t_c\). Then \(M_{t_c} + E_{t_c} > 0\), but we bound

\[
M_{t_c}^N + E_{t_c}^N \leq \left( \frac{1}{N} C_1(G_{t_c}^N) \right)^{\frac{1}{2}} \left( \frac{1}{N} \sum_{i=1}^N (1 + \frac{1}{2} |v_i|^2)^2 \right)^{\frac{1}{2}}.
\]  

(147)

The first term converges to 0 in probability, by definition of the phase transition, and the second term is bounded in \(L^2\) by hypothesis (B2.). This implies that \(M_{t_c}^N + E_{t_c}^N \to 0\) in probability, which contradicts Lemma 6.3; we must therefore have that \(t_g \geq t_c\).

Conversely, if \(t < t_g\), then \(M_t = 0\) by definition. Now, the convergence

\[
\frac{1}{N} C_1(G_t^N) \leq M_t^N \to 0
\]

(148)

in probability implies that the largest cluster is of the order \(o_p(N)\), which is only possible if \(t \leq t_c\). Since \(t < t_g\) was arbitrary, we must have \(t_g \leq t_c\), and together with the previous argument, we have shown that \(t_g = t_c\) as claimed.

6.1 Preparatory Lemmas

The proof of Lemma 6.3 is based on the following argument. We know, from Theorem 4.4, that any ‘mesoscopic’ clusters contain negligible mass; thanks to the integrability assumption (A3.), the same is true for the energy. Therefore, almost all mass and energy either belongs to the ‘microscopic’ scale, whose convergence is quantified by Lemma 3.1, or the giant component, whose convergence is the subject of interest here. Therefore, with a suitable approximation argument, the claimed convergence will follow from the quoted results.

We begin with some preparatory lemmas; throughout, we will assume the notation of Lemma 6.3. Firstly, we bound the size of the largest cluster below, even in the cases where there is no giant component.

Lemma 6.4. Fix \(t > 0\) and \(r \in \mathbb{N}\), and let \(G_t^N\) be the random graph process described above. Then

\[
\mathbb{P}(C_1(G_t^N) \leq r) \to 0.
\]

(149)

Proof. We remark that, due to (A1, A4.), \(m\) is not a point mass. Therefore, we can find open sets \(V, W\) of positive \(m\)-measure such that \(\inf_{v \in V, w \in W} |v - w| > 0\); define

\[
I_1 = \{i \in \{1, 2, ..., N\} : v_i \in V\}; \quad N_1 = \#I_1
\]

(150)

and similarly, \(I_2, N_2\), with \(V\) replaced by \(W\). By weak convergence (B1.) and openness of \(V\) and \(W\), \(N_1 \geq cN\) and \(N_2 \geq cN\) with high probability, for some constant \(c > 0\). Now, for all \(i \in I_1, j \in I_2\) the edge \(e = (ij)\) is present in \(G_t^N\) with probability

\[
\mathbb{P}(e \text{ present in } G_t^N) \geq \left(1 - \exp \left( - \frac{\epsilon}{N} \right) \right) \geq \frac{\delta}{N}
\]

(151)

for all \(N\) large enough, for some positive \(\epsilon, \delta > 0\), possibly depending on \(t, \kappa, \gamma, V, W\). Therefore, we can construct a random bipartite graph \(H_t^N\), with vertex sets size \(cN\) and independent edges occurring with probability \(\frac{\delta}{N}\), such that \(H_t^N \subset G_t^N\) with high probability. It is straightforward to see that the maximum degree \(\Delta(H_t^N) \to \infty\) in probability, which implies the claim. \(\square\)

For the proof of of Lemma 6.1, and later Theorem 1.2, we will wish to study the convergence of integrals \(\langle \varphi f, \mu_t^N \rangle\), for bounded continuous functions \(f\) with non-compact support. However, the convergence result Lemma 3.1 only gives us information when the support of \(f\) is compact. Our second preparatory lemma allows us to approximate the integrals \(\langle \varphi f, \mu_t^N \rangle\) for functions whose support is bounded in the \(\pi_n\)-direction.
Lemma 6.5 (A step towards uniform integrability). Suppose that \( \mu^N \) are stochastic coagulants satisfying (B1-2.). Then, for every positive \( r \in \mathbb{N} \),
\[
\beta(r, \eta) := \sup_{N \geq 1} \mathbb{E} \left[ \sup_{t \geq 0} \left\langle \varphi 1[\pi_e(x) > \eta, \pi_n(x) \leq r], \mu_t^N \right\rangle \right] \to 0 \quad \text{as } \eta \to \infty. \tag{152}
\]

Proof. Let \( \mu_t^N \) be a stochastic coagulant coupled to a random graphs process \( G_t^N \). Using Cauchy-Schwarz we note firstly that
\[
\sup_{t \geq 0} \left\langle \varphi 1[\pi_e(x) > \eta, \pi_n(x) \leq r], \mu_t^N \right\rangle
\leq \left( \frac{1}{N} \sum_{i=1}^{N} \right. \sum_{\xi \in \mathcal{C}(G_t^N)} \frac{1}{4} |v_i|^4 \right)^{\frac{1}{2}} \left( \sup_{t \geq 0} \frac{1}{N} \sum_{i=1}^{N} \left[ E(\mathcal{C}_j(G_t^N)) > \eta, \mathcal{C}_j(G_t^N) \leq r \right] \right)^{\frac{1}{2}} \tag{153}
\]
\[
= \left( \frac{1}{N} \sum_{i=1}^{N} \frac{1}{4} |v_i|^4 \right)^{\frac{1}{2}} \left( \sup_{t \geq 0} \left\langle \pi_n 1[\pi_e(x) > \eta, \pi_n(x) \leq r], \mu_t^N \right\rangle \right)^{\frac{1}{2}}. 
\]
As remarked in Definition 5.1, the velocities \( v_i \) associated with the graph nodes are constant in time, so the first factor is independent of \( t \geq 0 \), and is bounded in \( L^2 \) by the second assertion of (B2.). Therefore, it is sufficient to prove the claim with \( \varphi \) replaced by \( \pi_n \).

Recall from the discussion below (B2.) that \( X_{\eta/r}^N \) is given by
\[
X_{\eta/r}^N = \# \left\{ i \leq N : \frac{1}{2} |v_i|^2 > \frac{\eta}{r} \right\}. \tag{154}
\]
Fix \( t \geq 0 \). Now, if a cluster at time \( t \) contains at most \( r \) particles, but has total kinetic energy greater than \( \eta \), then it must contain a particle of velocity \( v \) with \( \frac{1}{2} |v|^2 > \frac{\eta}{r} \). Since each such cluster contains at most \( r \) particles, the contribution of these clusters is at most
\[
\left\langle \pi_n 1[\pi_e(x) > \eta, \pi_n(x) \leq r], \mu_t^N \right\rangle \leq \frac{r}{N} X_{\eta/r}^N. \tag{155}
\]
Now, the right-hand side is independent of \( t \geq 0 \), and so is an upper bound when we maximise over \( t \). Taking expectations, we obtain
\[
\mathbb{E} \left[ \sup_{t \geq 0} \left\langle \pi_n 1[\pi_e(x) > \eta, \pi_n(x) \leq r], \mu_t^N \right\rangle \right] \leq \frac{r}{N} \mathbb{E} \left[ X_{\eta/r}^N \right]. \tag{156}
\]
From (53), this vanishes as \( \eta \to \infty \), uniformly in \( N \).

\[\square\]

6.2 Proof of Lemma 6.3

Using the two preparatory lemmas developed above, we now prove Lemma 6.3.

Proof of Lemma 6.3. Throughout, we let \((\mu_t^N)_{t \geq 0}\) be a stochastic coagulant coupled to a random graph process \((G_t^N)_{t \geq 0}\), as described in Section 5. We write \((v_i)_{i=1}^N\) for the velocities associated to the graph vertices. The case \( t = 0 \) is trivial, and can be omitted. We deal first with the mass term \( M_t^N \) and show later how this may be modified for the energy \( E_t^N \).

Fix \( t > 0 \), and let \( \xi_N \) be a sequence, to be constructed later, such that
\[
\xi_N \to \infty; \quad \frac{\xi_N}{N} \to 0; \quad \mathbb{P}(C_1(G_t^N) \geq \xi_N) \to 1. \tag{157}
\]
We now construct ‘bump functions’ as follows. Let $\eta_r \to \infty$ be a sequence growing sufficiently fast that, in the notation of Lemma 6.5, $\beta(r, \eta_r) \to 0$, and let

$$S_{(r)} := \{ x \in S : \pi_n(x) < r, |\pi_p(x)| \leq \sqrt{2r} \eta_r, \pi_e(x) \leq \eta_r \}. \quad (158)$$

Let $\tilde{g}_r$ be the indicator $\tilde{g}_r = 1[\pi_n(x) < r]$, and construct a continuous, compactly supported function $\tilde{f}_r$ such that

$$0 \leq \tilde{f}_r \leq 1; \quad \tilde{f}_r = 1 \text{ on } S_{(r)}; \quad \tilde{f}_r(x) = 0 \text{ if } \pi_n(x) \geq r. \quad (159)$$

The final condition is compatible with continuity because $\pi_n : S \to \mathbb{N}$ is continuous and integer valued. We define $f_N = \tilde{f}_{\xi_N}$ and $g_N = \tilde{g}_{\xi_N}$. We now decompose the difference $M_t^N - M_t$:

$$M_t^N - M_t = \left( (\langle \pi_n, \mu_t \rangle - \langle \pi_n f_N, \mu_t \rangle) + \langle \pi_n f_N, \mu_t - \mu_t^N \rangle \right) := T_1^4$$

$$+ \langle \pi_n (f_N - g_N), \mu_t \rangle + \langle \pi_n g_N, \mu_t^N \rangle - \langle \pi_n, \mu_0^N - M_t \rangle := T_2^4$$

$$+ \langle \pi_n, \mu_0^N - \mu_0 \rangle := T_3^4$$

where we recall that $M_t = \langle \pi_n, \mu_0 - \mu_t \rangle$. We now estimate the errors $T_i^4, i = 1, 3, 4, 5$; the remaining term $T_2^4$ will be dealt with separately, and requires careful construction of the sequence $\xi_N$.

1. **Estimate on $T_1^4$.** Let $h_N = 1_{S(\xi_N)}$, so that $h_N \leq f_N \leq 1$. As $N \to \infty$, $\pi_n h_N \uparrow \pi_n$, and so by monotone convergence, $\langle \pi_n h_N, \mu_t \rangle \uparrow \langle \pi_n, \mu_t \rangle$ This implies immediately that the (nonrandom) error $T_1^4 \to 0$.

2. **Estimate on $T_3^4$.** From the definitions of $f_N, g_N$, we observe that

$$|T_3^4(t)| = \langle \pi_n (g_N - f_N), \mu_t^N \rangle \leq \langle \pi_n 1[\pi_n(x) < \xi_N, \pi_e(x) > \eta_N], \mu_t^N \rangle. \quad (161)$$

Therefore, in the notation of Lemma 6.5,

$$\mathbb{E} \left[ |T_3^4(t)| \right] \leq \beta(\xi_N, \eta_N). \quad (162)$$

By construction of $\eta_r$, and since $\xi_N \to \infty$, it follows that $\mathbb{E}[|T_3^4(t)|] \to 0$, which implies convergence to 0 in probability.

3. **Estimate on $T_2^4$.** By the choice (157) of $\xi_N$, we have that $C_1(G_t^N) \geq \xi_N$ with high probability. On this event, we have the equality

$$\langle \pi_n g_N, \mu_t^N \rangle = \langle \pi_n, \mu_t^N \rangle - \langle \pi_n 1_{\pi_n \geq \xi_N}, \mu_t^N \rangle$$

$$= \langle \pi_n, \mu_0^N \rangle - \frac{1}{N} \sum_{j \geq 1 : C_j(G_t^N) \geq \xi_N} \sum_{i \in C_j(G_t^N)} 1 \quad (163)$$

$$= \langle \pi_n, \mu_0^N \rangle - M_t^N - \frac{1}{N} \sum_{j \geq 2 : C_j(G_t^N) \geq \xi_N} \sum_{i \in C_j(G_t^N)} 1. $$

Therefore, with high probability,

$$T_2^4(t) \leq \frac{1}{N} \sum_{j \geq 2 : C_j(G_t^N) \geq \xi_N} \sum_{i \in C_j(G_t^N)} (1 + \frac{1}{2} |v_i|^2) \quad (164)$$
and we bound, on this event,

\[ T_N^4(t) \leq \left( \frac{1}{N} \sum_{j \geq 2; C_j(G_i^N) \geq \xi_N} C_j(G_i^N) \right)^{\frac{1}{2}} \left( \frac{1}{N} \sum_{i=1}^N (1 + \frac{1}{2} |v_i|^2) \right)^{\frac{1}{2}}. \]  

(165)

The first term is the mass of the mesoscopic clusters, which converges to 0 in probability, by Theorem 4.4, and the second term is bounded in \( L^2 \) by the second part of (B2.). Together, these imply that \( T_N^4(t) \to 0 \) in probability.

4. **Estimate on \( T_N^5 \).** Using the first part of (B2.), we have the convergence in distribution

\[ \langle \pi_n, \mu_0^N \rangle \to \langle \pi_n, \mu_0 \rangle \]  

(166)

which implies that \( T_N^5 \to 0 \) in probability as desired.

5. **Construction of \( \xi_N \), and convergence of \( T_N^2 \).** It remains to show how a sequence \( \xi_N \) can be constructed such that \( T_N^2 \to 0 \) in probability. Let \( A_{r,N}^1, A_{r,N}^2 \) be the events

\[ A_{r,N}^1 = \left\{ |\langle \bar{\varphi}_f, \mu_i^N - \mu_t \rangle| < \frac{1}{r} \right\}; \quad A_{r,N}^2 = \{ C_1(G_i^N) \geq r \}. \]  

(167)

Then, as \( N \to \infty \) with \( r \) fixed, both \( \mathbb{P}(A_{r,N}^1), \mathbb{P}(A_{r,N}^2) \to 1 \), by Lemma 3.1 and Lemma 6.4. We now define \( N_r \) inductively for \( r \geq 1 \) by setting \( N_1 = 1 \), and letting \( N_{r+1} \) be the minimal \( N > N_r \) such that, for all \( N' \geq N \),

\[ N \geq (r + 1)^2; \quad \mathbb{P}(A_{r+1,N'}^1) > \frac{r}{r + 1}; \quad \mathbb{P}(A_{r+1,N'}^2) > \frac{r}{r + 1}. \]  

(168)

Now, we set \( \xi_N = r \) for \( N \in [N_r, N_{r+1}) \cap \mathbb{N} \). It follows that \( \xi_N \to \infty \) and \( \xi_N \leq \sqrt{N} \ll N \), and

\[ \mathbb{P} \left( C_1(G_i^N) \geq \xi_N \right) \geq 1 - \frac{1}{\xi_N} \to 1. \]  

(169)

Therefore, \( \xi_N \) satisfies the requirements (157) above. Moreover,

\[ \mathbb{P} \left( |T_N^2| < \frac{1}{\xi_N} \right) \geq \mathbb{P} \left( A_{1, N}^1 \right) > 1 - \frac{1}{\xi_N} \to 1 \]  

(170)

and so, with this choice of \( \xi_N \), \( T_N^2 \to 0 \) in probability. Since we have now dealt with every term appearing in the decomposition (160), it follows that \( M_i^N \to M_i \) in probability, as claimed.

The arguments for the energy \( E_t^N \) are identical to those above, using the same bound (164) on \( T_N^4 \).

We also note, for future use, an important corollary of this argument.

**Corollary 6.6.** At the instant of gelation, the gel is negligible: \( M_{t_g} = E_{t_g} = 0 \).

**Proof.** This follows from the critical case of Theorem 4.2 exactly as in (147).

\[ \square \]

7 **Behaviour of the Second Moment**

In this section, we consider part 2 of Theorem 1.1, concerning the behaviour of the second moment \( \mathcal{E}(t) = \langle \varphi^2, \mu_t \rangle \). Following [31], one might expect that the gelation time \( t_g \) corresponds to a divergence of \( \mathcal{E}(t) \) as \( t \uparrow t_g \); by an approximation argument, we will show that this is indeed the case. We also introduce a duality argument, corresponding to Theorem 4.6, which allows us to prove that \( \mathcal{E} \) is finite on \( (t_g, \infty) \). The final assertion follows from the fact that \( M_{t_g} = E_{t_g} = 0 \), which is the content of Corollary 6.6.

30
7.1 Subcritical Regime

We first deal with the subcritical regime \([0, t_g]\), to show that the second moment \(E(t)\) is finite and increasing on this interval, and that \(t_g\) is exactly the first time at which \(E\) diverges.

**Lemma 7.1.** Suppose \(\mu_0\) satisfies (A1-4.), and let \((\mu_t)_{t \geq 0}\) be the corresponding solution to \((E+G)\). The second moment \(E(t) = \langle \varphi^2, \mu_t \rangle\) is finite and increasing on \([0, t_g]\), and increases to infinity as \(t \uparrow t_g\), where \(t_g\) is the associated gelation time.

The ideas of this argument follow [31], where there is a similar result for approximately multiplicative kernels, for which the total rate \(\overline{T}(x,y)\) is bounded above and below by nonzero multiples of \(\tilde{\varphi}(x)\tilde{\varphi}(y)\), where \(\tilde{\varphi}\) is a mass function playing the same rôle as our \(\varphi\). Unfortunately, this cannot be applied directly, for two reasons.

i). Firstly, the total rate in (9) contains the term \(\pi_p(x) \cdot \pi_p(y)\) of indefinite sign.

ii). Secondly, the remaining combination of \(\pi_n, \pi_e\) is not in an approximately multiplicative form: particles of either very low or very high energy prevent the desired lower bound from holding for any positive constant.

We deal with these as follows. To deal with item i), we introduce a symmetrised kernel \(K^m\), and use the symmetry (A1.) to argue that the solutions coincide exactly with solutions to the original equation (E-G). To deal with the degeneracy in point ii), we consider a truncated state space \(S^e\), which excludes particles of either very high or very low energy. In this context, the kernel \(K^m\) is approximately multiplicative, so the results of [31] apply; we then carefully justify taking the limit \(\epsilon \downarrow 0\).

We first consider the symmetrised equation. Let \(K^m\) be the kernel on \(S \times S \times S\) given by

\[
K^m(x,y,dz) = \frac{1}{4}K(Rx,y,dz) + \frac{1}{2}K(x,y,dz) + \frac{1}{4}K(x,Ry,dz). \tag{171}
\]

Let \(L^m\) be the drift operator for the modified kernel \(K^m\), and consider the modified equation

\[
\mu_t = \mu_0 + \int_0^t L^m(\mu_s)ds. \tag{mE-G}
\]

The total rate of the modified kernel is

\[
\overline{K^m}(x,y) = \kappa \pi_n(x)\pi_n(y) + 2\gamma (\pi_n(x)\pi_e(y) + \pi_e(x)\pi_n(y)). \tag{172}
\]

Consider a modified state space, which truncates the velocity distribution by excluding clusters with extreme kinetic energies: for \(\epsilon > 0\), let

\[
S^\epsilon = \{ x \in S : \epsilon \pi_n(x) \leq \pi_e(x) \leq \epsilon^{-1} \pi_n(x) \}. \tag{173}
\]

Note that this state space is preserved under both kernels \(K, K^m\). Moreover, on the reduced state space \(S^e\), the modified kernel \(K^m\) is approximately multiplicative [31] in the sense that, for some \(\delta_\epsilon > 0\) and \(\Delta_\epsilon < \infty\), we have

\[
\delta_\epsilon \varphi(x)\varphi(y) \leq \overline{K^m}(x,y) \leq \Delta_\epsilon \varphi(x)\varphi(y) \tag{174}
\]

for all \(x, y \in S^\epsilon\). For any \(\epsilon\), let \(\mu_0^\epsilon\) denote the restriction \(\mu_0^\epsilon(dx) = 1_{x \in S^\epsilon} \mu_0(dx)\). We now appeal to [31, Theorem 2.2], on existence and uniqueness for approximately multiplicative kernels, to obtain the following, which provides the connection between gelation and explosion of a second moment.

**Lemma 7.2.** Suppose (A1-4.) hold, and let \(\mu_0^\epsilon\) be as above. For all \(\epsilon > 0\), there is a unique maximal conservative solution \((\nu_t^\epsilon)_{0 \leq t \leq t_c^\epsilon}\) in \(S^\epsilon\) to the modified equation \((mE-G)\), starting from \(\mu_0^\epsilon\). Moreover, the map \(t \mapsto \langle \varphi^2, \nu_t^\epsilon \rangle\) is finite and increasing on \([0, t_c^\epsilon]\), and increases to \(\infty\) as \(t \uparrow t_c^\epsilon\).
Similarly, we can also apply Corollary 2.2 to see that there exist maximal conservative solutions \((\mu^\epsilon_t)_{t < t_g^\epsilon}\) to (E-G) starting at \(\mu_0^\epsilon\), which are given by initial segments of a global solution \((\mu^\epsilon_t)_{t \geq 0}\) to (E+G). Repeatedly exploiting uniqueness, we show that these coincide with the solution to (mE-G):

**Lemma 7.3 (Relationship of equations).** Let \(\mu_0^\epsilon\) be as above, for initial data \(\mu_0\) satisfying (A1-4.). Then the maximal conservative solutions \((\mu^\epsilon_t)_{t < t_g^\epsilon}\) and \((\nu^\epsilon_t)_{t < t_g^\epsilon}\), to (E-G) and (mE-G) respectively, coincide. In particular, \(t_e = t_g^\epsilon\), and the map

\[
t \mapsto \langle \varphi^2, \mu^\epsilon_t \rangle
\]

is finite and increasing on \([0, t_e]\), and increases to \(\infty\) as \(t \uparrow t_g^\epsilon\).

**Proof.** Firstly, we note that \((\mu^\epsilon_t \circ R^{-1})_{t < t_g^\epsilon}\) also solves (E-G), starts at \(\mu_0^\epsilon\) by (A1.), and is conservative. Therefore, by uniqueness in Corollary 2.2, we must have \(\mu^\epsilon_t \circ R^{-1} = \mu^\epsilon_t\) for all \(t < t_g^\epsilon\). Therefore, for any bounded, measurable function \(f\), and \(t < t_g^\epsilon\), it follows from elementary manipulations that

\[
\int_{S^3} (f(z) - f(x) - f(y))K(x, y, dz)\mu^\epsilon_t(dx)\mu^\epsilon_t(dy) = \int_{S^3} (f(z) - f(x) - f(y))K(Rx, y, dz)\mu^\epsilon_t(dx)\mu^\epsilon_t(dy)
\]

Combining these, we see that \((\mu^\epsilon_t)_{t < t_g^\epsilon}\) solves the modified equation (mE-G), and so, by uniqueness of the maximal conservative solution \((\nu^\epsilon_t)_{t < t_g^\epsilon}\) in Lemma 7.2, we have

\[
t_g^\epsilon \leq t_e; \quad \mu^\epsilon_t = \nu^\epsilon_t \quad \forall t < t_g^\epsilon. \tag{177}
\]

The other implication is identical, using the uniqueness of the maximal conservative solution \((\nu^\epsilon_t)_{t < t_g^\epsilon}\) in Lemma 7.2 to deduce that \(\nu^\epsilon_t = \nu^\epsilon_t \circ R^{-1}\) for \(t < t_g^\epsilon\). Hence, the equations (176) hold with \(\nu^\epsilon_t\) in place of \(\mu^\epsilon_t\), for any bounded, measurable \(f\) and \(t < t_g^\epsilon\). Therefore, \((\nu^\epsilon_t)_{t < t_g^\epsilon}\) is a conservative solution to the unmodified equation (E-G), and so by Corollary 2.2,

\[
t_e \leq t_g^\epsilon; \quad \nu^\epsilon_t = \mu^\epsilon_t \quad \forall t < t_g^\epsilon. \tag{178}
\]

Using previous arguments, we make the following remark on the gelation times \(t_g^\epsilon\).

**Lemma 7.4.** Suppose \(\mu_0\) satisfies (A1-4.). Let \(\mu_0^\epsilon\) be as above, and let \(t_g^\epsilon\) be the corresponding gelation time. Then \(t_g^\epsilon \to t_g\) as \(\epsilon \downarrow 0\).

**Proof.** This follows from the calculation in Lemma 5.3. Let \(m^\epsilon\) be the cutoff measure

\[
m^\epsilon (dv) = 1\{\sqrt{2\epsilon} \leq |v| \leq \sqrt{2\epsilon^{-1}}\} \, m(dv) \tag{179}
\]

so that \(m^\epsilon\) corresponds to the initial particle velocities for \(\mu_0^\epsilon\). By Corollary 6.2, the gelation times \(t_g^\epsilon, t_g\) are given by a continuous function of the moments \(\sigma_0, \sigma_2, \sigma_4\) of \(m^\epsilon, m\). Using dominated convergence and hypotheses (A3-4.), we have

\[
\sigma_k(m^\epsilon) \to \sigma_k(m); \quad 0 \leq k \leq 4 \tag{180}
\]

which implies the claimed convergence \(t_g^\epsilon \to t_g\).

We now turn to the proof of Lemma 7.1. We say that a local solution \((\nu_t)_{t < T}\) to (E-G) is **strong** if, for all times \(t < T\),

\[
\int_0^t \langle \varphi^2, \nu_s \rangle \, ds < \infty. \tag{181}
\]

We use the following result from [31] on the existence and uniqueness of strong solutions.
Lemma 7.5. Any strong solution to \((E-G)\) is conservative. For any finite measure \(\mu_0\) with \(\langle \varphi^2, \mu_0 \rangle < \infty\), there is a unique maximal strong solution \((\mu'_t)_{t \leq t_e(\mu_0)}\) to \((E-G)\), starting at \(\mu_0\), and with \(t_e(\mu_0) > 0\), such that \(\langle \varphi^2, \mu'_t \rangle\) is increasing on \([0, t_e(\mu_0))\). If \(t_e(\mu_0) < \infty\), then \(\langle \varphi^2, \mu'_t \rangle\) increases to \(\infty\) as \(t \uparrow t_e(\mu_0)\).

Here, the subscript ‘e’ denotes explosion: \(t_e(\mu_0)\) is exactly the blow-up time of the second moment. When the measure \(\mu_0\) is clear, we will omit the argument of \(t_e\).

Proof. This is almost a special case of [31, Theorem 2.1]. From the cited result, for any finite measure \(\mu_0\) with \(\langle \varphi^2, \mu_0 \rangle < \infty\), there exists a maximal strong solution \((\mu'_t)_{t \leq t_e(\mu_0)}\). Moreover, there exists a constant \(C = C(\kappa, \gamma) > 0\) such that, for all such \(\mu_0\), \(t_e(\mu_0) \geq C\langle \varphi^2, \mu_0 \rangle^{-1}\). By applying this bound to \(\mu'_t\), if \(t_e(\mu_0) < \infty\), then \(\langle \varphi^2, \mu'_t \rangle \geq (C(t_e(\mu_0) - t))^{-1}\) which implies the claimed divergence.

By Corollary 2.2, since \((\mu'_t)_{t \leq t_e}\) is conservative, it follows that \(t_e \leq t_g\), and \(\mu'_t = \mu_t\) for all \(t < t_e\). It remains to show that \(t_e \geq t_g\).

Following the ideas of [31, Proposition 2.7], we obtain the integral relations, for all \(t < t_e\),

\[
\langle \pi^2_n, \mu_t \rangle = \langle \pi^2_n, \mu_0 \rangle + \int_0^t \left[ \kappa \langle \pi^2_n, \mu_s \rangle^2 + 4\gamma \langle \pi_n \pi_e, \mu_t \rangle \langle \pi^2_n, \mu_s \rangle \right] ds; \tag{182}
\]

\[
\langle \pi_n \pi_e, \mu_t \rangle = \langle \pi_n \pi_e, \mu_0 \rangle + \int_0^t \left[ \kappa \langle \pi^2_n, \mu_s \rangle \langle \pi_n \pi_e, \mu_s \rangle + 2\gamma \langle \pi_n \pi_e, \mu_t \rangle^2 + 2\gamma \langle \pi^2_n, \mu_s \rangle \langle \pi^2_e, \mu_s \rangle \right] ds; \tag{183}
\]

\[
\langle \pi^2_e, \mu_t \rangle = \langle \pi^2_e, \mu_0 \rangle + \int_0^t \left[ \kappa \langle \pi_n \pi_e, \mu_s \rangle^2 + 4\gamma \langle \pi_n \pi_e, \mu_t \rangle \langle \pi^2_e, \mu_s \rangle \right] ds. \tag{184}
\]

These immediately imply that \(\mathcal{E}(t)\) is bounded on compact subsets of \([0, t_e)\), and in particular cannot diverge before \(t_e\). Combining this with Lemma 7.5, the maximal time \(t_e\) of existence of a strong solution is precisely the first time at which the second moment \(\mathcal{E}(t)\) diverges, or \(\infty\) if there is no divergence.

We also remark on the relationship of this result to the solutions \((\mu'_t)_{t \leq t'_e} = (\nu'_t)_{t < t'_e}\) discussed in Lemmas 7.2, 7.3. It is clear, from Lemma 7.2, that \((\nu'_t)_{t < t'_e}\) is a strong solution. Moreover, in view of the comments above, since \(\langle \varphi^2, \nu'_t \rangle \uparrow \infty\) as \(t \uparrow t'_e\), it follows that that \((\nu'_t)_{t \leq t'_e}\) is the maximal strong solution with initial data \(\mu'_0\). This justifies the use of the notation \(t'_e\) in Lemma 7.2.

As discussed at the beginning of this section, our argument is now as follows. From general considerations in [31], we argued above that the gelation time and explosion time coincide for the restricted dynamics: \(t'_e = t'_g\), and we now wish to justify the limit \(\epsilon \downarrow 0\). The limiting behaviour of \(t'_g\) is understood from Lemma 7.4, and so we wish to understand the behaviour of \(t'_e\).

Using standard regularity arguments, we may view (182 - 184) as a differential equation for the three moments \(q_t = ((\langle \pi^2_n, \mu_t \rangle, \langle \pi_n \pi_e, \mu_t \rangle, \langle \pi^2_e, \mu_t \rangle))\) and, from the discussion above, the blow-up time to the ODE system is exactly \(t_e\). An identical argument holds for \(\mu'_t\), which blows up at \(t'_e\). By analysing this system of ODEs, we will show that the explosion time is continuous in the initial data, which implies that \(t'_e \to t_e\).

Lemma 7.6. Consider the ordinary differential equation \(\dot{q}_t = b(q_t)\) in \(\mathbb{R}^3\), where \(b\) is the locally Lipschitz field given by

\[
b(q_1, q_2, q_3) = \begin{pmatrix}
    \kappa q_1^2 + 4\gamma q_1 q_2 \\
    \kappa q_1 q_2 + 2\gamma q_2 q_2 + 2\gamma q_1 q_3 \\
    \kappa q_2^2 + 4\gamma q_2 q_3
\end{pmatrix}. \tag{185}
\]

Then, for all \(q_0 \in \mathbb{R}^3\), there exists a unique maximal solution \(\psi(q_0, t)\) starting at \(q_0\), defined until time \(\zeta(q_0) \in (0, \infty]\). Consider the sets

\[
E = (0, \infty)^3; \quad E_\delta = [\delta, \infty)^3. \tag{186}
\]

Then, if \(q_0 \in E_\delta\) for some \(\delta > 0\), then the solution \((\psi(q_0, t))_{t \leq \zeta(q_0)} \subset E_\delta\). We have the following properties:
i). Let $J_\epsilon$ be the set
\[
J_\epsilon = \{ q \in E : \zeta(q) \geq \epsilon \}.
\]
If $\gamma > 0$, then for all $\epsilon, \delta > 0$, the set $E_\delta \cap J_\epsilon$ is bounded. Moreover, $\zeta < \infty$ everywhere.

ii). Suppose $q^n_0 \in E$ and $q^n_0 \to q_0 \in E$. Then $\zeta(q^n_0) \to \zeta(q_0)$.

iii). Suppose $I \subset \mathbb{R}_+$ is an open interval, and the map $q_0 : I \to E$ is continuous, and such that $t < \zeta(q_0(t))$ for all $t \geq 0$. Then the map $I \to E, t \mapsto \psi(q_0(t), t)$ is continuous.

Proof. For all three items, the case where $\gamma = 0, \kappa > 0$ may be checked by an elementary explicit calculation. For the remainder of the proof, we exclude this case, and consider only the case $\gamma > 0$.

i). Let $\zeta_0$ denote the blowup time for the dynamics (185) with $\kappa = 0$. It is straightforward to see that $\zeta(q) \leq \zeta_0(q)$ for all $q \in E$, and so it is sufficient to show that $E_\delta \cap \{ q : \zeta_0(q) \geq \epsilon \}$ is bounded. We argue using the following explicit computation.

Let $q(0) = (q_1(0), q_2(0), q_3(0)) \in E$, and let $q(t) = (q_1(t), q_2(t), q_3(t))$ be the solution to (185) with $\kappa = 0$, starting at $q(0)$. It is then straightforward to see that
\[
\frac{1}{q_1(t)} \frac{d}{dt} q_1(t) = \frac{1}{q_3(t)} \frac{d}{dt} q_3(t)
\] (188)

which implies that $q_3(t) = q_1(t)q_3(0)/q_1(0)$ for all $t \geq 0$. Now, the linear combination $\bar{q}(t)$ given by
\[
\bar{q}(t) = q_1(t) + \sqrt{\frac{q_1(0)}{q_3(0)}} q_2(t)
\] (189)

has the same blowup time as $q(t)$, and satisfies the ordinary differential equation
\[
\frac{d}{dt} \bar{q}(t) = 4\gamma \sqrt{\frac{q_3(0)}{q_1(0)}} \bar{q}(t)^2.
\] (190)

This has the unique solution
\[
\bar{q}(t) = \left( \frac{1}{\bar{q}(0)} - 4\gamma \sqrt{\frac{q_3(0)}{q_1(0)}} t \right)^{-1}, \quad t < \frac{1}{4\gamma \bar{q}(0)} \sqrt{\frac{q_1(0)}{q_3(0)}}.
\] (191)

In terms of the initial data $q(0)$, this gives the blowup time as
\[
\zeta(q(0)) = \frac{1}{4\gamma} \left( \sqrt{q_1(0)q_3(0)} + q_2(0) \right)^{-1}
\] (192)

which converges to 0 as $q(0) \to \infty$ in $E_\delta$. This shows that $E_\delta \cap \{ q : \zeta_0(q) \geq \epsilon \}$ is bounded, as claimed. The same computation also shows that $\zeta(q) < \infty$ for all $q \in E$.

ii). The lower semicontinuity of explosion times is standard, and follows from the continuous dependence on the initial data. Therefore, it is sufficient to prove that $\lim \sup_{n \to \infty} \zeta(q^n) \leq \zeta(q)$.

Suppose, for a contradiction, that for some $\epsilon > 0$, we have $\lim \sup_{n \to \infty} \zeta(q^n) > \zeta(q) + \epsilon$; write $\tau = \zeta(q)$. By passing to a subsequence, we may assume that $\zeta(q^n) > \tau + \epsilon$ for all $n$, and some fixed $\epsilon > 0$. Moreover, since $q^n \to q \in E$, we may assume that $q^n, q \in E_\delta$ for all $n$, for some $\delta > 0$, which implies that $\psi(q^n, t) \in E_\delta$ for all $t < \zeta(q^n)$ and all $n \in \mathbb{N}$.

Now, if $t \leq \tau$, we have $\zeta(\psi(t, q^n)) = \zeta(q^n) - t \geq \epsilon$, which implies the containment
\[
\{ \psi(t, q^n) : t \leq \tau, n \geq 1 \} \subset E_\delta \cap J_\epsilon
\] (193)
which we know, from item i), to be bounded: for some $C < \infty$,

$$\{\psi(t, q^n) : t \leq \tau, n \geq 1\} \subset [0, C]^3. \tag{194}$$

By the lemma of leaving compact sets, there exists $s < \tau$ such that, for all $t \in (s, \tau)$, $\psi_t(q) \notin [0, C]^3$. However, if we pick $t \in (s, \tau)$, we have $\psi_t(q^n) \to \psi_t(q)$, by the continuity of the dependence in the initial conditions, which is a contradiction. Therefore, $\limsup_{n \to \infty} \zeta(q^n) \leq \zeta(q)$, which proves the claimed convergence.

iii). Firstly, we note that by ii), the map $t \mapsto \zeta(q_0(t))$ is continuous on $I$. Therefore, fixing $t \in I$, we may choose choose $\epsilon, \delta > 0$ such that, if $|t - s| \leq \delta$, then $s \in I$ and $s < \min(\zeta(q_0(s)), \zeta(q_0(t))) - \epsilon$. Now, we observe that, for $s \in [t - \delta, t + \delta],$

$$|\psi(t, q_0(t)) - \psi(s, q_0(s))| \leq |\psi(t, q_0(t)) - \psi(t, q_0(s))| + |\psi(t, q_0(s)) - \psi(s, q_0(s))|. \tag{195}$$

As $s \to t$, the first term converges to 0 by continuity of the solution $s \mapsto \psi(x_0(t), s)$; it is therefore sufficient to control the second term. We observe that, for all $s \in [t - \delta, t + \delta]$, we have $\zeta(\psi(s, q_0(s))) = \zeta(q_0(s)) - s > \epsilon$. Moreover, by compactness, there exists some $\eta > 0$ such that $q_0(s) \in E_\eta$ for all $s \in [t - \delta, t + \delta]$, and so $\psi(q_0(s), u) \in E_\eta$ for all $0 \leq u \leq \zeta(q_0(s))$. However, we showed in point i). above that the region $E_\eta \cap I = \{q \in E_\eta : \zeta(q) \geq \epsilon\}$ is compact and so there exists a constant $M = M(\epsilon)$: for all $s \in [t - \delta, t + \delta]$, and for all $u \leq t + \delta,$

$$u < \zeta(q_0(s)); \quad |b(\psi(u, q_0(s)))| \leq M. \tag{196}$$

This implies the bound, for all $s \in [t - \delta, t + \delta],$

$$|\psi(t, q_0(s)) - \psi(s, q_0(s))| \leq M|t - s| \tag{197}$$

which implies the claimed continuity.

\[ \square \]

We can now use this to prove our main result Lemma 7.1 on the second moment $\mathcal{E}(t)$ in the subcritical phase.

**Proof of Lemma 7.1.** Let $\mu_0$ be any measure on $S$ satisfying (A1-4.), and let $(\mu_t)_{t \geq 0}$ be the associated solution to $(E+G)$. From the discussion following Lemma 7.5, it is sufficient to show that $t_g = t_e < \infty$. We also recall that $t_e$ is characterised as the explosion time $\zeta$ of the ODE system (182-184).

Since the base measure $m$ of $\mu_0$ is not a multiple of the point mass $\delta$, all the quadratic moments $q$ of $\mu_0$ are strictly positive, and so $q \in E$. Therefore, by the second point of Lemma 7.6, the explosion time $t_e = \zeta(q) < \infty$. As $\epsilon \downarrow 0$, the quadratic moments $q^\epsilon$ of $\mu_0^\epsilon$ converge to the quadratic moments $q \in E$ of $\mu_0$ by dominated convergence, and using (A4.). Therefore, by Lemma 7.6, $t_e^\epsilon = \zeta(q^\epsilon) \to \zeta(q) = t_e$. By Lemma 7.3, we know that $t_e^\epsilon = t_g^\epsilon$, and by Lemma 7.4, $t_g^\epsilon \to t_g$. Together, these imply that $t_e = t_g$, as claimed.

\[ \square \]

### 7.2 The Critical Point

Using the concepts introduced above, we next consider the behaviour at and near the critical time $t_g$.

**Lemma 7.7.** Assume that $\mu_0$ is a probability measure. In the notation of Lemma 7.1, we have

$$\mathcal{E}(t_g) = \infty = \lim_{t \to t_g} \mathcal{E}(t). \tag{198}$$
Proof. We first show that $\mathcal{E}(t_g) = \infty$. Suppose, for a contradiction, that $\mathcal{E}(t_g) < \infty$. Then, applying [31, Proposition 2.7] as in Lemma 7.5, we see that, for some positive $\delta > 0$, there exists a strong solution $(\nu_t)_{t<\delta}$ to (E-G), starting at $\mu_{t_g}$. This solution is conservative, so is an initial segment of the solution $(\nu_t)_{t\geq 0}$ to (E+G) starting at $\mu_{t_g}$. By uniqueness in Lemma 2.1,

$$\nu_t = \mu_{t_g+t} \quad \text{for all } t \geq 0.$$  \hspace{1cm} (199)

By Corollary 6.6, $\langle \varphi, \mu_{t_g} \rangle = \langle \varphi, \mu_0 \rangle$, and by definition of $t_g$,

$$\langle \varphi, \mu_{t_g+t} \rangle < \langle \varphi, \mu_0 \rangle = \langle \varphi, \mu_{t_g} \rangle \quad \text{for all } t > 0.$$  \hspace{1cm} (200)

This contradicts the fact that $(\nu_t)_{t<\delta}$ is strong, which therefore shows that $\mathcal{E}(t) = \infty$.

The second point follows, because $t \mapsto \mu_t$ is continuous, and $\mu \mapsto \langle \varphi^2, \mu \rangle$ is lower semicontinuous, when $\mathcal{M}$ is equipped with the vague topology. □

7.3 The Supercritical Regime

We finally turn to the supercritical case; our result is as follows.

Lemma 7.8. In the notation of Lemma 7.1, the map $t \mapsto \mathcal{E}(t)$ is finite and continuous, and therefore locally bounded, on $(t_g, \infty)$.

The proof is based on a duality argument following Theorem 4.6, which connects the measures in the supercritical regime to an auxiliary process in the subcritical case. Let $(G^N_t)_{t\geq 0}$ be the random graph processes described in Section 5 with $N$ particles sampled independently from $m$, and fix $t > t_g$. Let $\tilde{G}^N_t$ be the graph $G^N_t$ with the giant component deleted.

Let $\rho_t(v) = \rho(t, v)$ be the survival function defined in Lemma 4.1, and let $\tilde{m}^t(dv) = (1 - \rho_t(v))m(dv)$ and $\tilde{\rho}_0^t$ the corresponding measure on $S$ under pushforward by $\iota : v \mapsto (1, v, \frac{1}{2}|v|^2)$. By Lemma 2.1, there exists a unique solution $(\tilde{\mu}_s^t)_{s \geq 0}$ to the equation (E+G) starting at $\tilde{\rho}_0^t$; write $\tilde{t}_g(t)$ for its gelation time. By Theorem 4.6, we can construct a generalised vertex space $\tilde{V} = (\mathbb{R}^d, \tilde{m}, (w_N)_{N \geq 1})$ and a graph $\tilde{G}_t^N \sim G^\tilde{V}(N, tK)$ such that $\mathbb{P}(\tilde{G}_t^N = \tilde{G}_t^N) \rightarrow 1$, and where $w_N$ is an enumeration of the vertexes $v_i$ not belonging to the giant component.

In order to appeal to Lemmas 5.2, 6.1, we will now verify that the desired regularity conditions (A1-4, B1-2.) hold for the vertex space $\tilde{V}$.

Lemma 7.9. Fix $t > 0$, and let $\mu_0, \tilde{m}, \tilde{\rho}_0^t$ and $\tilde{V}$ be as described above. Then the regularity conditions (B1-2.) hold for $w_N$ and $\tilde{\rho}_0^t$.

Proof. To ease notation, we write $\tilde{\mu}_0, \tilde{m}$ for $\tilde{\rho}_0^t, \tilde{m}^t$, $\mu_0^N$ for the initial empirical measure of the unmodified process corresponding to $v_N$, and $\tilde{\rho}_0^N$ for the reduced empirical measure corresponding to $w_N$:

$$\tilde{\rho}_0^N = \frac{1}{N} \sum_{i=1}^{i_N} \delta_{(1, w_i, \frac{1}{2}|w_i|^2)}.$$  \hspace{1cm} (201)

It is straightforward to see that $\tilde{\rho}_0^N$ inherits the properties (A1-4.) from $\mu_0$, and so it is sufficient to establish (B1-2.). We will also appeal to the fact that the unmodified empirical measures $\mu_0^N$ satisfy (B1-2.), which is straightforward to verify.

For (B1.), we note that part of the content of Theorem 4.6 is that $\tilde{V}$ is a generalised vertex space, as defined in Definition 4.1, which includes the weak convergence

$$\tilde{m}_N = \frac{1}{N} \sum_{i \leq l_N} \delta_{w_i} \rightarrow \tilde{m} \quad \text{weakly, in probability.}$$  \hspace{1cm} (202)
Since the map \( t : \mathbb{R}^d \to S \) is continuous and the vague topology is weaker than the weak topology, (B1.) follows.

We will now show that (B2.) follows from the previous point, together with the moment estimates for the original initial measure \( \mu_N^0 \).

Fix \( R < \infty \), and let \( \chi_R \in C_c(S) \) be such that \( 1_{S_R} \leq \chi \leq 1_{S_{R+1}} \). We observe that

\[
\| x, \hat{\mu}_0^N \| - \| x, \hat{\mu}_0^0 \| \leq \| x\chi_R, \hat{\mu}_0^N - \hat{\mu}_0^0 \| + \| x, 1_{S_R^c}, \hat{\mu}_0^N + \| x, 1_{S_{R+1}^c}, \hat{\mu}_0^0 \|
\]

\[
\leq \| x\chi_R, \hat{\mu}_0^N - \hat{\mu}_0^0 \| + \frac{\sqrt{2}}{R} \| \varphi^2, \mu_0^N \| + \frac{\sqrt{2}}{R} \| \varphi^2, \mu_0^0 \|.
\]

(203)

We now fix \( \epsilon, \delta > 0 \). Since \( \| \varphi^2, \mu_0^N \| \) converges almost surely by (A3.), and in particular is \( O_p(1) \), we may choose \( R < \infty \) such that the second and third terms are at most \( \epsilon/3 \) with probability exceeding \( 1 - \delta/2 \), for all \( N \). For this choice of \( R \), the first term vanishes as \( N \to \infty \) by vague convergence in probability, and so is at most \( \epsilon/3 \) with probability exceeding \( 1 - \delta/2 \) for all \( N \) large enough. Therefore, for all such \( N \), we have

\[
P\left( \| x, \hat{\mu}_0^N - \hat{\mu}_0^0 \| > \epsilon \right) \leq \delta
\]

(204)

which proves the desired convergence in probability.

For the second assertion of (B2.), we note that \( \| \varphi^2, \mu_0^N \| \leq \| \varphi^2, \mu_0^0 \| \) by the construction of \( w_N \), and \( \| \varphi^2, \mu_0^N \| \) is bounded in \( L^1 \) by (A3.), recalling that the velocities in \( v_N \) are sampled independently from \( m \). \( \square \)

We now use this preparatory result to prove Lemma 7.8.

**Proof of Lemma 7.8.** Let \( G_t^N, \tilde{G}_t^N, \tilde{G}_t^N \) be as above. Recalling that we consider equality of graphs to include equality of the vertex data, it follows from Theorem 4.6 that

\[
P(\mu^N(\tilde{G}_t^N) = \mu^N(\tilde{G}_t^N)) \to 1.
\]

(205)

From Lemmas 5.2, 7.9, we obtain the following convergences in probability:

\[
\mu^N(G_t^N) \to \mu_t; \quad \mu^N(\tilde{G}_t^N) \to \hat{\mu}_t
\]

(206)

in the vague topology, in probability. Moreover, the difference

\[
\mu^N(G_t^N) - \mu^N(\tilde{G}_t^N) = \frac{1}{N} \delta(C_1(G_t^N))
\]

(207)

converges to 0 in the vague topology in probability, since the support is eventually disjoint from any compact set, with high probability. It follows that

\[
\mu^N(\tilde{G}_t^N) \to \mu_t
\]

(208)

in the vague topology, in probability, and by uniqueness of limits, we have \( \hat{\mu}_t^t = \mu_t \). Using assumption (A3.), we can see that \( tk \in L^2(\mathbb{R}^d \times \mathbb{R}^d, dx \times dx) \), and so it follows from Theorem 4.6 that the graphs \( \tilde{G}_t^N \) are subcritical. By Lemma 6.1, it follows that that \( t < \hat{t}_e(t) \), and so by Lemma 7.1, we have

\[
\langle \varphi^2, \mu_t \rangle = \langle \varphi^2, \hat{\mu}_t^t \rangle < \infty.
\]

(209)

Using Theorem 4.5 and dominated convergence, the map

\[
t \mapsto q_t^0 = \left( \langle \rho_n^2, \hat{\mu}_0^0 \rangle, \langle \pi_n \pi_e, \hat{\mu}_0^0 \rangle, \langle \pi_e^2, \hat{\mu}_0^0 \rangle \right)
\]

\[
= \left( (1 - \rho_t), \frac{1}{2} |v|^2 (1 - \rho_t), m \right)
\]

(210)

is continuous, and it is clear that it takes values in \( E = (0, \infty)^3 \). Therefore, by the general ODE considerations in Lemma 7.6 point iii), it follows that the map

\[
t \mapsto q_t^t = \left( \langle \rho_n^2, \hat{\mu}_t^t \rangle, \langle \pi_n \pi_e, \hat{\mu}_t^t \rangle, \langle \pi_e^2, \hat{\mu}_t^t \rangle \right)
\]

(211)

is finite and continuous on \( (t_g, \infty) \). Since \( \hat{\mu}_t^t = \mu_t \), this implies that \( t \mapsto E(t) \) is finite and continuous on \( (t_g, \infty) \), which implies that it is bounded on compact subsets. \( \square \)

**Remark 7.10.** The same argument also shows that \( t \mapsto \hat{t}_e(t) \) is continuous. This fact will be used later in the proof of Lemma 11.2.
8 Representation and Dynamics of the Gel

8.1 Representation Formula

The duality construction used in the proof of Lemma 7.8 gives us a natural way to identify \( M_t, E_t \) in terms of the survival function \( \rho_t \) from Sections 4, 5. This is the content of the following lemma.

**Lemma 8.1.** Let \( \mu_0 \) be an initial data satisfying (A1-4.) for some measure \( m \), and let \( M_t, E_t \) be the gel data for the corresponding solution to \((E+G)\). Let \( \rho_t(\cdot) \) be the corresponding survival function defined in Sections 4, 5. Then we have the equalities

\[
M_t = \int_{\mathbb{R}^d} \rho_t(v) m(dv); \quad E_t = \int_{\mathbb{R}^d} \frac{1}{2} |v|^2 \rho_t(v) m(dv).
\]

In particular, both \( M_t \) and \( E_t \) are continuous, and if \( t > t_g \) then \( M_t > 0 \) and \( E_t > 0 \).

It is immediate from the symmetry (A1.), or from Lemma 2.3, that \( P_t = 0 \). Together with the identification of \( \rho_t \) in Lemma 5.5, this proves part 3 of Theorem 1.1.

**Proof.** We deal with the supercritical and subcritical/critical cases, \( t > t_g, t \leq t_g \) separately.

1. **Supercritical Case** \( t > t_g \). Let \( (\hat{\mu}_s^t)_{s \geq 0} \) and \( \hat{t}_g(t) \) be as in the proof of Theorem 7.8. Then, since \( (\hat{\mu}_s^t)_{s \geq 0} \) is conservative on \( [0, \hat{t}_g) \), and \( t < \hat{t}_g(t) \), we have

\[
\langle \pi_n, \hat{\mu}_t^t \rangle = \langle \pi_n, \hat{\mu}_0^t \rangle = \int_{\mathbb{R}^d} (1 - \rho(t, v)) m(dv).
\]

But since \( \mu_t = \hat{\mu}_t^t \), we have

\[
M_t := \langle \pi_n, \mu_0 \rangle - \langle \pi_n, \mu_t \rangle = \langle \pi_n, \mu_0 \rangle - \int_{\mathbb{R}^d} (1 - \rho(t, v)) m(dv)
\]

which implies the result for \( M_t \). The argument for \( E_t \) is identical.

2. **Subcritical and Critical Cases** \( t \leq t_g \). For \( t < t_g \), the result is immediate: we have \( M_t = E_t = 0 \) by definition of \( t_g \), and \( \rho_t = 0 \) by Theorem 4.1. The critical case is identical, recalling from Corollary 6.6 that \( M_{t_g} = E_{t_g} = 0 \).

Continuity follows from Theorem 4.5 by using dominated convergence. For the final claim, if \( t > t_g \) then at least one of \( M_t, E_t \) is strictly positive. If \( M_t > 0 \), then by (212), \( \rho_t \) is not 0 \( m \)-almost everywhere, and by (A4.),

\[
m \left( v : \frac{1}{2} |v|^2 \rho_t(v) > 0 \right) > 0
\]

and it follows that \( E_t > 0 \). The case where \( E_t > 0 \) is identical.

\[ \square \]

8.2 Gel Dynamics Beyond the Critical Time

We now obtain point 4 of Theorem 1.1 as a consequence of the previous results. We have already proven the continuity of \( M_t, E_t \) on the whole time interval \([0, \infty)\) and the finiteness of the second moments \( q_t = (\langle \pi_n^2, \mu_t \rangle, \langle \pi_n \pi_e, \mu_t \rangle, \langle \pi_e^2, \mu_t \rangle) \) in the supercritical regime. Therefore, it is sufficient to prove the following result.
Lemma 8.2. In the notation of Lemma 7.8, let $(M_t, E_t)$ be the mass and energy of the gel associated to $(\mu_t)_{t \geq 0}$. Then, for $t \geq t_\xi$, we have

\[
M_t = \int_{t_\xi}^{t} \left( \kappa \langle \pi_n^2, \mu_s \rangle M_s + 2\gamma \left[ \langle \pi_n \pi_e, \mu_s \rangle M_s + \langle \pi_n^2, \mu_s \rangle E_s \right] \right) ds; \tag{216}
\]

\[
E_t = \int_{t_\xi}^{t} \left( \kappa \langle \pi_n \pi_e, \mu_s \rangle M_s + 2\gamma \left[ \langle \pi_e^2, \mu_s \rangle M_s + \langle \pi_n \pi_e, \mu_s \rangle E_s \right] \right) ds. \tag{217}
\]

Remark 8.3. We interpret Lemma 8.2 as as saying that the growth of the gel away from $t_\xi$ is due entirely to the absorption of finite clusters, rather than an additional blow-up due to coagulation of small particles. This may be expected following the relationship between gelation and blowup of the second moment $\mathcal{E}(t)$ in Lemma 7.1, and the finiteness of $\mathcal{E}$ in the supercritical regime.

Proof. We return to the truncated dynamics $(E_t^\xi, E_t^{\xi, \mu})$ used in the proof of Lemma 2.1. We recall that, starting at

\[
\mu_0^\xi = 1_{S_\xi} \mu_0; \quad g_0^\xi = \int_{x \notin S_\xi} x \mu_0(dx) \tag{218}
\]

the solution $(\mu_t^\xi, g_t^\xi)$ to $(E_t^\xi, E_t^{\xi, \mu})$ exists and is unique, and as $\xi \uparrow \infty$, we have

\[
\mu_t^\xi \uparrow \mu_t; \quad (M_t^\xi, E_t^\xi) \downarrow (M_t, E_t) \tag{219}
\]

where $(\mu_t)_{t \geq 0}$ is the solution to $(E+G)$ starting at $\mu_0$, and $(M_t, E_t)$ are the associated gel data.

Fix $s, t$ such that $t_\xi < s < t$. It is immediate from $(E_t^\xi)$ that

\[
M_t^\xi - M_s^\xi = \int_s^t \left( \kappa \langle \pi_n^2, \mu_u \rangle M_u^\xi + 2\gamma \left[ \langle \pi_n \pi_e, \mu_u \rangle M_u^\xi + \langle \pi_n^2, \mu_u \rangle E_u^\xi \right] \right) du; \tag{220}
\]

\[
E_t^\xi - E_s^\xi = \int_s^t \left( \kappa \langle \pi_n \pi_e, \mu_u \rangle M_u^\xi + 2\gamma \left[ \langle \pi_e^2, \mu_u \rangle M_u^\xi + \langle \pi_n \pi_e, \mu_u \rangle E_u^\xi \right] \right) du. \tag{221}
\]

By the monotonicity $\mu_u^\xi \leq \mu_u$, and local boundedness in Lemma 7.8, $\langle \varphi^2, \mu_u^\xi \rangle$ is bounded, uniformly in $\xi < \infty$ and $u \in [s, t]$. It is also straightforward to see that the truncated gel data are bounded by $M_u^\xi \leq 1; \quad E_u^\xi \leq \langle \pi_e, \mu_u \rangle$. Together, these imply that the integrands are bounded. Using (219) and bounded convergence, we take the limit $\xi \to \infty$ to obtain

\[
M_t - M_s = \int_s^t \left( \kappa \langle \pi_n^2, \mu_u \rangle M_u^\xi + 2\gamma \left[ \langle \pi_n \pi_e, \mu_u \rangle M_u + \langle \pi_n^2, \mu_u \rangle E_u \right] \right) du; \tag{222}
\]

\[
E_t - E_s = \int_s^t \left( \kappa \langle \pi_n \pi_e, \mu_u \rangle M_u + 2\gamma \left[ \langle \pi_e^2, \mu_u \rangle M_u + \langle \pi_n \pi_e, \mu_u \rangle E_u \right] \right) du. \tag{223}
\]

Taking $s \downarrow t_\xi$, and using the continuity $(M_s, E_s) \downarrow (0, 0)$ established in Lemma 8.1, we obtain the claimed result. \hfill \Box

9 Uniform Convergence of the Stochastic Coagulant

We now show how previous results, describing the dynamics of $E_t, M_t$, imply convergence to the limiting values

\[
M_t \to 1; \quad E_t \to \frac{1}{2} \sigma_2(m) \tag{224}
\]

at an exponential rate. From this, we will be able to upgrade the previous result, Lemma 3.1, on the convergence of the stochastic coagulant to uniform convergence.
Lemma 9.1. Let $\mu_0$ be an initial measure satisfying (A1-4.) for a probability measure $m$, and let $M_t, E_t$ be the mass and energy of the associated gel. As $t \uparrow \infty$, we have

$$ M_t \uparrow 1; \quad E_t \uparrow \frac{1}{2} \sigma_2(m) $$

(225)

where $m$ is given by (A2.).

Proof. We recall that at least one of the rate parameters $\kappa, \gamma$ is strictly positive. Therefore, it is sufficient to prove the result in the cases where $\gamma > 0$ and $\kappa \geq 0$ is arbitrary, or where $\kappa > 0$ and $\gamma \geq 0$ is arbitrary.

1. $\gamma > 0$. We deal first with the case where $\gamma > 0$ and $\kappa \geq 0$ is arbitrary. Choose $t_0 > t_{gel}$, and let $\lambda = E_{t_0}$; it follows from Lemma 8.1 that $\lambda > 0$, and since $E_t$ is increasing, $E_t \geq \lambda > 0$ for all $t \geq t_0$. By Lemma 8.2, for all $t \geq t_0$, we have

$$ M_t \geq M_{t_0} + \int_{t_0}^t 2\gamma \langle \pi_n, \mu_s \rangle E_s ds \geq M_{t_0} + 2\lambda \gamma \int_{t_0}^t \langle \pi_n, \mu_s \rangle ds $$

$$ = M_{t_0} + 2\lambda \gamma \int_{t_0}^t (1 - M_s) ds $$

(226)

where the second inequality uses the fact that $\pi_n \geq 1$ everywhere, and the final equality is the definition of $M_t$. It follows that, for all $t \geq t_0$,

$$ M_t \geq 1 - e^{-2\lambda \gamma (t-t_0)} (1 - M_{t_0}). $$

(227)

From the definition, it is immediate that $M_t \leq 1$, which implies that $M_t \to 1$ as claimed. A similar argument for the energy shows that

$$ E_t \geq E_{t_0} + 2\lambda \gamma \int_{t_0}^t (\langle \pi_e, \mu_0 \rangle - \langle \pi_e, \mu_s \rangle) ds $$

(228)

from which it follows that, for $t \geq t_0$,

$$ E_t \geq \langle \pi_e, \mu_0 \rangle - e^{-2\lambda \gamma (t-t_0)} (\langle \pi_e, \mu_0 \rangle - E_{t_0}) $$

(229)

which, as above, implies the claim.

2. $\kappa > 0$. The case $\kappa > 0$ is essentially identical. Let $t_0 > t_{gel}$, and let $\lambda'$ be given instead by $\lambda' = M_{t_0}$. Then the same argument as above gives the inequalities, for all $t \geq t_0$,

$$ M_t \geq M_{t_0} + \kappa \lambda' \int_{t_0}^t (1 - M_s) ds; $$

(230)

$$ E_t \geq E_{t_0} + \kappa \lambda' \int_{t_0}^t (\langle \pi_e, \mu_0 \rangle - E_s) ds $$

(231)

which yield the claimed convergences as above. \hfill \Box

Lemma 9.2. Let $\mu^N_t$ be the stochastic coagulants constructed in the introduction, with $N$ particles sampled independently from a probability measure $m$, and let $\mu_t$ be the corresponding solution to the Smoluchowski equation $(E+G)$, with conditions (A1-4.). Then we have the uniform convergence

$$ \sup_{t \geq 0} d_0(\mu^N_t, \mu_t) \to 0 $$

(232)

in probability.
Proof. From the definition of the vague topology, it is sufficient to prove that, for any \( f \in C_c(S) \) with \( 0 \leq f \leq 1 \), we have the uniform convergence \( \sup_{t \geq 0} \langle f, \mu^N_t - \mu_t \rangle \to 0 \) in probability.

Fix \( \epsilon > 0 \). By Lemma 9.1, we can find \( t_+ \in (t_g, \infty) \) such that \( M_{t_+} > 1 - \frac{\epsilon}{2} \). Let \( A^1_N \) be the event \( A^1_N = \{ M^N_{t_+} > 1 - \frac{\epsilon}{2} \} \); by Lemma 6.3, it follows that \( \mathbb{P}(A^1_N) \to 1 \). On this event, we have

\[
\sup_{t \geq 0} \langle f, \mu^N_t - \mu_t \rangle \leq \sup_{0 \leq t \leq t_+} \langle f, \mu^N_t - \mu_t \rangle + \sup_{t > t_+} \langle f, \mu^N_t - \mu_t \rangle
\]

\[
\leq \sup_{0 \leq t \leq t_+} \langle f, \mu^N_t - \mu_t \rangle + \left( 1 - M^N_{t_+} \right) + \left( 1 - M_{t_+} \right)
\]

\[
\leq \sup_{0 \leq t \leq t_+} \langle f, \mu^N_t - \mu_t \rangle + \epsilon
\]

and the first term converges to 0 in probability by Lemma 3.1.

\[ \square \]

10 Behaviour Near the Critical Point

We now prove item 5 of Theorem 1.1, concerning the phase transition: we will show that the gel data \((M_t, E_t)\) have strictly positive right-derivatives at the gelation time \( t_{gel} \). We start from the nonlinear fixed point equation (22), which we rewrite as

\[
c_t = tF(c_t); \quad F \left( \frac{a}{b} \right) = 2 \int_{\mathbb{R}^d} (1 - e^{-a-b|v|^2}) \left( \frac{\kappa + \gamma |v|^2}{\gamma} \right) m(dv).
\]

(234)

The following proof is a modification of the arguments in [7, Theorem 3.17], which itself generalises an analogous, well-known result for the phase transition of Erdős-Rényi graphs.

Lemma 10.1. Suppose that \( \mu_0 \) satisfies (A1-4.) for a probability measure \( m \), and let \( c_t = (a_t, b_t) \) be as in Lemma 5.5. Then \( c_t \) is right-differentiable at \( t_g \), \( a_{t_g}' > 0 \), and

\[
\lambda = \frac{b_{t_g}'}{a_{t_g}'} = \frac{\sqrt{\kappa^2 + 4\gamma(\kappa\sigma_2(m) + \gamma\sigma_4(m)) - \kappa}}{2(\kappa\sigma_2(m) + \gamma\sigma_4(m))}.
\]

(235)

Proof. We first assume that \(|v|\) is not constant \( m \)-almost everywhere, and equip \( \mathbb{R}^2 \) with the inner product

\[
((a, b), (a', b'))_m = \int_{\mathbb{R}^d} (a + b|v|^2)(a' + b'|v|^2)m(dv)
\]

(236)

and write \(|.|_m\) for the associated norm. Differentiating under the integral sign twice, and using (A3.), we write

\[
F \left( \frac{a}{b} \right) = \Lambda \left( \frac{a}{b} \right) - \Sigma \left( \frac{a}{b} \right) + R \left( \frac{a}{b} \right)
\]

(237)

where \( \Lambda(\cdot), \Sigma(\cdot) \) are the linear and quadratic terms, and \( R \) is a remainder term:

\[
\Lambda \left( \frac{a}{b} \right) = 2 \left( \begin{array}{c} \kappa + \gamma\sigma_2(m) \\ \gamma \sigma_2(m) \end{array} \right) \left( \begin{array}{c} a \\ b \end{array} \right);
\]

(238)

\[
\Sigma \left( \frac{a}{b} \right) = \int_{\mathbb{R}^d} (a + b|v|^2)^2 \left( \begin{array}{c} \kappa + \gamma |v|^2 \\ \gamma \end{array} \right) m(dv);
\]

(239)

\[
|R(c)|_m = o \left( |c|^2_m \right) \text{ as } |c| \to 0.
\]

(240)
The signs here are chosen to guarantee that, if \( c > 0 \), then \( \Delta c, B(c) > 0 \), and it is straightforward to verify that \( \Lambda \) is self-adjoint with respect to \((\cdot, \cdot)_m\). We also note that we have already found the spectrum of \( \Lambda \) in the computation in Lemma 5.3: \( \Lambda \) has exactly two eigenvalues, of which the larger is \( t_g^{-1} \), and the corresponding \(| \cdot |_m\)-unit eigenvector is given by

\[
\psi = (t_g^{-2}/4 + \gamma^2(\sigma_4(m) - \sigma_2(m)^2))^{-1/2} \left( t_g^{-1}/2 - \gamma \sigma_2(m) \right).
\] (241)

From Lemma 5.5, Theorem 4.2 and Theorem 4.5, we know that \( c(t_g) = 0 \), that \( c(t_g + \epsilon) \in [0, \infty) \setminus \{(0, 0)\} \) for all \( \epsilon > 0 \), and that \( t \mapsto c(t) \) is continuous at \( t_g \).

It is straightforward to see that \( \psi \) is an eigenvector of \( \Lambda \) of eigenvalue \( t_g^{-1} \), and \(|\psi|_m = 1\). Writing \( \psi^\perp \) for the orthogonal compliment of \( \text{Span}(\psi) \) with respect to \((\cdot, \cdot)_m\), it follows from the self-adjointness of \( \Lambda \) that \( \Lambda \) maps \( \psi^\perp \) into itself. Moreover, for \( t > t_g \) small enough, \((t\Lambda - 1)|\psi^\perp\) is invertible, and that the operator norm \( \|(t\Lambda - 1)|_{\psi^\perp}\|_{m \to m} \) is bounded as \( t \downarrow t_g \).

Let \( Q : \mathbb{R}^2 \to \mathbb{R}^2 \) be the orthogonal projection onto \( \psi^\perp \) with respect to \((\cdot, \cdot)_m\), and write \( c^*_t = Qc_t \) so that we have the orthogonal decomposition

\[
c_t = \alpha_t \psi + c^*_t
\] (242)

for some \( \alpha_t \in \mathbb{R} \). Noting that \( \Lambda Q = Q \Lambda \), it follows from (234, 242) that

\[
c^*_t = Q(tF(c_t)) = t\Lambda c^*_t + tQ(\Sigma(c_t) + R(c_t)).
\] (243)

The function \( -\Sigma(c) + R(c) \) is of quadratic growth as \(|c|_m \to 0\), and using the invertibility of \((t\Lambda - 1)|_{\psi^\perp}\) described above, it follows that there exists \( \beta > 0 \) such that \( |c^*_t|_m \leq \beta |c|_m^2 \) whenever \(|c|_m \leq 1\). In turn, it follows that \(|c_t|_m \sim \alpha_t\) as \( t \downarrow t_g \). Now, using (234) and the self-adjointness of \( \Lambda \), we obtain

\[
\alpha_t = (t_g \Lambda \psi, c_t)_m = t_g(\psi, \Lambda c_t)_m = t_g(\psi, \Lambda c_t)_m
\]

\[
= \frac{t_g}{t}(\psi, c_t)_m - t_g(\psi, -\Sigma(c_t) + R(c_t))_m
\]

\[
= \frac{t_g}{t}\alpha_t - t_g(\psi, -\Sigma(c_t) + R(c_t))_m.
\] (244)

We now expand to second order in \( \alpha_t \); for clarity, we will number the error terms \( T^i_t \). Since \(|c_t| \sim \alpha_t\), it follows that that \(|c^*_t|_m = O(\alpha_t^2)\) and that \( R(c_t) = o(\alpha_t^2) \). Expanding \( \Sigma(c_t) \) using (242),

\[
-\Sigma(c_t) + R(c_t) = -\alpha_t^2 \Sigma(\psi) + T^1_1;
\]

\[
|T^1_1|_m = o(\alpha_t^2).
\] (245)

It therefore follows that

\[
\alpha_t = t_g \left( \frac{\alpha_t}{t} + \alpha_t^2 (\psi, \Sigma(\psi))_m \right) + T^2_1;
\]

\[
T^2_1 = o(\alpha_t^2).
\] (246)

For \( t > t_g \) small enough, \( \alpha_t > 0 \), and we may rearrange to find

\[
t - t_g = t t_g \alpha_t (\psi, \Sigma(\psi)) + T^3_1;
\]

\[
T^3_1 = o(\alpha_t)
\] (247)

and in particular \( \alpha_t = \Theta(t - t_g) \) as \( t \downarrow t_g \), since \((\psi, \Sigma(\psi))_m > 0\). Finally, we obtain

\[
\frac{\alpha_t}{t - t_g} \to \frac{1}{t_g^2(\psi, \Sigma(\psi))} \text{ as } t \downarrow t_g.
\] (248)

The calculations above show that \(|c_t - \alpha_t \psi| = O((t - t_g)^2)\), and the claimed right-differentiability now follows. Observing that \( \alpha^*_t > 0 \) and that the first component of \( \psi \) is strictly positive, it follows that \( \alpha^*_t > 0 \) as claimed. Finally, the expression (235) for the ratio of the right-derivatives follows from the definition of \( \psi \) with some...
elementary algebra, and using the computation of \( t_g \) in Lemma 5.3.

We now deal with the degenerate case in which \( |v| \) is constant almost everywhere; say, \( |v| = v \) for \( m \)-almost all \( v \). In this case, let \( \tilde{a}_t = a_t + b_t v^2 \); taking \( v = (v, 0, ..., 0) \) in (234), we obtain

\[
\tilde{a}_t = 2 \int_{\mathbb{R}^d} (1 - e^{-a_t - b_t |w|^2})(\kappa + \gamma |w - (v, 0, ..., 0)|^2) m(dw)
  = 2(\kappa + 2\gamma v^2)(1 - e^{-\tilde{a}_t}).
\]

This may be rearranged into the nonlinear fixed point equation for Erdős-Rényi random graphs, and analysed using the standard argument, or viewed as a special case of (234) with

\[
\tilde{\kappa} = \kappa + 2\gamma v^2; \quad \tilde{\gamma} = 0.
\]

Either argument shows that \( \tilde{a}_t \) is right-differentiable at \( t_g \), with right-derivative

\[
\frac{\tilde{a}'_t}{t_g} = \frac{4}{\kappa}.
\]

To go from \( \tilde{a}_t \) to the original parameters \((a_t, b_t)\), we observe that from (234), we have

\[
\begin{pmatrix} a_t \\ b_t \end{pmatrix} = 2(1 - e^{-\tilde{a}_t}) \begin{pmatrix} \kappa + \gamma v^2 \\ \gamma \end{pmatrix}
\]

and so the right-differentiability of \( \tilde{a}_t \) implies the right-differentiability of \( c_t = (a_t, b_t) \) at \( t_g \). Since \( v > 0 \) by (A4.) and at least one of \( \kappa, \gamma \) is strictly positive, it follows that \( a'_t > 0 \). Finally, (252) implies that

\[
\frac{b'_t}{a'_t} = \frac{\gamma}{\kappa + \gamma v^2}
\]

which may be seen to coincide with claimed expression (235).

We now show how this implies item 5 of Theorem 1.1. From Lemmas 5.5, 8.1, we have that

\[
M_t = \int_{\mathbb{R}^d} (1 - e^{-a_t - b_t |v|^2}) m(dv); \quad E_t = \int_{\mathbb{R}^d} |v|^2 (1 - e^{-a_t - b_t |v|^2}) m(dv).
\]

Differentiating under the integral sign using hypothesis (A3.), we obtain

\[
M_t = a_t + b_t \sigma_2(m) + o(a_t + b_t); \quad E_t = \frac{1}{2} a_t \sigma_2(m) + \frac{1}{2} b_t \sigma_4(m) + o(a_t + b_t).
\]

From the previous result, we see that for \( t > t_g \),

\[
M_t = (t - t_g)(a'_t + b'_t \sigma_2(m)) + o(t - t_g);
\]

\[
E_t = \frac{1}{2}(t - t_g)(a'_t \sigma_2(m) + b'_t \sigma_4(m)) + o(t - t_g)
\]

which proves the desired right-differentiability, and the positivity of the right-derivatives. Since \( M'_t > 0 \), we may take a quotient and let \( t \downarrow t_g \) to obtain the claimed size-biasing effect.
11 Convergence of the Gel

Finally, we prove the remaining part of Theorem 1.2, concerning the uniform convergence of the stochastic gel, drawing on other results we have proven. We recall that $g_t^N = (M_t^N, P_t^N, E_t^N)$ are the normalised mass, momentum and energy of the cluster with the largest mass, corresponding to the largest component $C_1(G_t^N)$. To conclude the proof of Theorem 1.2, we must extend Lemma 6.3, to show that $g_t^N \to g_t$ uniformly in time, in probability.

This subsection is structured as follows. We recall that, in the proof of Lemma 6.3, we used the result on mesoscopic clusters from [7]: if $\xi_N \to \infty$ and $\frac{C_0}{N} \to 0$, then for all $t \geq 0$,

$$
\frac{1}{N} \sum_{j \geq 2; C_j(G_t^N) \geq \xi_N} C_j(G_t^N) \to 0
$$

(258)

in probability. In the first subsection, we improve this to uniform convergence in probability, and collect some corollaries. In the second subsection, we then use these to prove the claimed convergence.

11.1 Bounds on the Mesoscopic Clusters

The main result of this section is the following.

Lemma 11.1. Let $G_t^N$ be the random graph process constructed in Section 5, with $N$ particles sampled independently according to a probability measure $m$ satisfying (A1-4.). For any sequence $\xi_N \to \infty$ with $\xi_N \ll N$, we have the convergence

$$
\sup_{t \geq t_g} \left[ \frac{1}{N} \sum_{j \geq 2; C_j(G_t^N) \geq \xi_N} C_j(G_t^N) \right] \to 0 \quad \text{in probability.}
$$

(259)

We prove this lemma as follows. First, we prove uniform convergence on compact subsets $I \subset (t_g, \infty)$ in Lemma 11.2. We will then show how this may be extended to the whole interval $[t_g, \infty)$. The subcritical case is also true (Lemma 11.3), and we will then deduce analogous results for the momentum and energy.

Lemma 11.2. Let $G_t^N$ and $\xi_N$ be as above. Fix a compact subset $I \subset (t_g, \infty)$. Then we have the convergence

$$
\sup_{t \in I} \left[ \frac{1}{N} \sum_{j \geq 2; C_j(G_t^N) \geq \xi_N} C_j(G_t^N) \right] \to 0 \quad \text{in probability.}
$$

(260)

Proof of Lemma 11.2. It is sufficient to show that for every $t > t_g$ the claim holds for some $I$ of the form $I = (t_-, t_+) \subset (t_g, \infty)$ containing $t$. As in Theorem 7.8, let $\hat{m}^t$ be the measure on $\mathbb{R}^d$ given by $\hat{m}^t(dv) = (1 - \rho_t(v))m(dv)$. We also let $\hat{\mu}_0^t$ be the pushforward of $\hat{m}^t$ by $v \mapsto (1, v, \frac{1}{2}v^2)$, and $\hat{t}_g(t)$ the gelation time of the solution $(\hat{\mu}_s^t)_{s \geq 0}$ to (E+G) starting at $\hat{\mu}_0^t$. We showed in the proof of Theorem 7.8 that, for all $t > t_g$, $\hat{t}_g(t) > t$, and the map $t \mapsto \hat{t}_g(t)$ is continuous. Therefore, for any $t > t_g$, we can choose $t_\pm$ such that

$$
t_g < t_- < t < t_+ < \hat{t}_g(t_-).
$$

(261)

We form $\tilde{G}_{t_-}^N$ from $G_{t_-}^N$ by deleting all vertexes of the giant component of $C_1(G_{t_-}^N)$. We now form a new graph, $\tilde{G}_{t_+}^N$, by including all edges between vertexes of $\tilde{G}_{t_-}^N$ which are present in the graph $G_{t_+}$.

From Theorem 4.6, we can construct a generalised vertex space $\tilde{V}$ and graphs $\tilde{G}_{t_-}^N \sim \tilde{\mathcal{G}}(N, t_- K)$, with the same vertex data as $\tilde{G}_{t_-}^N$, such that

$$
P \left( \tilde{G}_{t_-}^N = \tilde{G}_{t_-}^N \right) \to 1.
$$

(262)
We now form $\tilde{G}_{t_-, t_+}^N$ by adding those edges present in $G_{t_-, t_+}^N$. By the Markov property of the graph process $(G_s^N)_{t \geq 0}$, these edges are independent of the construction of $G_{t_-, t_+}^N$, and so $\tilde{G}_{t_-, t_+}^N \sim \mathcal{G}(N, t_+K)$. By the choices of $t_\pm$, $\tilde{G}_{t_-, t_+}^N$ is still subcritical, and by construction,

$$\mathbb{P}\left( \tilde{G}_{t_-, t_+}^N = \tilde{G}_{t_-, t_+}^N \right) \to 1.$$ (263)

For $s \in [t_-, t_+]$, let $C_1(G_s^N)$ be the connected component of $G_s^N$ which contains $C_1(G_{t_-}^N)$. By considering the cases $C_1'(G_s^N) = C_1(G_s^N)$ and $C_1'(G_s^N) \neq C_1(G_s^N)$ separately, for all $s \in [t_-, t_+]$ we bound

$$\sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N) \leq \sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N)$$ (264)

Observe that we can rewrite the sum as

$$\sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N) = \sum_{i=1}^N 1[C(i; G_s^N) \geq \xi N; i \notin C_1'(G_s^N)].$$ (265)

Fet $s \in [t_-, t_+]$, and $i \in \{1, 2, \ldots, N\}$. Suppose that $C(i; G_s^N) \geq \xi N$ and that $i \notin C_1'(G_s^N)$. It follows that $C(i; G_s^N)$ is disjoint from $C_1(G_{t_-}^N)$, and so all vertexes of $C(i; G_s^N)$ are present in $\tilde{G}_{t_-, t_+}^N$. Moreover, all edges in $C(i; G_s^N)$ are also present in $\tilde{G}_{t_-, t_+}^N$, and so $C(i; G_s^N) \subset C(i; \tilde{G}_{t_-, t_+}^N)$. Therefore, for all $i$, and all $s \in [t_-, t_+]$

$$1[C(i; G_s^N) \geq \xi N; i \notin C_1'(G_s^N)] \leq 1[i \in V(\tilde{G}_{t_-, t_+}^N); C(i; \tilde{G}_{t_-, t_+}^N) \geq \xi N].$$ (266)

Summing, we have the bound

$$\frac{1}{N} \sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N) \leq \frac{1}{N} \sum_{j \geq 1: C_j(\tilde{G}_{t_-, t_+}^N) \geq \xi N} C_j(\tilde{G}_{t_-, t_+}^N)$$ (267)

Therefore,

$$\sup_{s \in [t_-, t_+]} \left[ \frac{1}{N} \sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N) \right] \leq \frac{1}{N} C_1(\tilde{G}_{t_-, t_+}^N) + \frac{1}{N} \sum_{j \geq 2: C_j(\tilde{G}_{t_-, t_+}^N) \geq \xi N} C_j(\tilde{G}_{t_-, t_+}^N).$$ (268)

Since $\tilde{G}_{t_-, t_+}^N = \tilde{G}_{t_-, t_+}^N$ with high probability, we obtain the bound

$$\sup_{s \in [t_-, t_+]} \left[ \frac{1}{N} \sum_{j \geq 1: C_j(G_s^N) \geq \xi N} C_j(G_s^N) \right] \leq \frac{1}{N} C_1(\tilde{G}_{t_-, t_+}^N) + \frac{1}{N} \sum_{j \geq 2: C_j(\tilde{G}_{t_-, t_+}^N) \geq \xi N} C_j(\tilde{G}_{t_-, t_+}^N)$$ (269)

with high probability. The first term of the right-hand side converges to 0 in probability because $\tilde{G}_{t_-, t_+}^N$ is subcritical, and the second term converges to 0 in probability by Theorem 4.4. \qed

45
Proof of Lemma 11.1. Fix \( \epsilon > 0 \); without loss of generality, assume that \( \epsilon < 1 \). By continuity from Lemma 8.1 and Lemma 9.1, we can choose \( t_\pm \in (t_g, \infty) \) such that
\[
M_{t_-} < \frac{\epsilon}{3}; \quad M_{t_+} > 1 - \frac{\epsilon}{3}.
\] (270)

Consider now the events
\[
A^1_N = \left\{ M^{N\downarrow}_{t_-} < \frac{2\epsilon}{3}; \quad M^{N\uparrow}_{t_+} > 1 - \epsilon \right\} ;
\] (271)
\[
A^2_N = \left\{ \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_{t_-}) \geq \xi_N} C_j(G^N_{t_-}) < \frac{\epsilon}{3} \right\} .
\] (272)

We know that \( \mathbb{P}(A^1_N) \to 1 \) from Lemma 6.3, and that \( \mathbb{P}(A^2_N) \to 1 \) from Theorem 4.4. On the event \( A^1_N \cap A^2_N \), we bound as follows.

i). For the initial interval \( [t_g, t_-] \), an argument similar to that of Lemma 11.2 shows that, on this event,
\[
\sup_{t \in [t_g, t_-]} \left[ \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_t) \geq \xi_N} C_j(G^N_t) \right] \leq \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_{t_-}) \geq \xi_N} C_j(G^N_{t_-})
\]
\[
= M^{N\downarrow}_{t_-} + \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_{t_-}) \geq \xi_N} C_j(G^N_{t_-})
\]
\[
< \epsilon.
\] (273)

ii). For late times \( t \in [t_+, \infty) \), the largest cluster \( C_1(G^N_t) \) is at least the size of the cluster containing \( C_1(G^N_{t_+}) \). Therefore,
\[
\inf_{t \geq t_+} \frac{1}{N} C_1(G^N_t) \geq M^{N\uparrow}_{t_+} > 1 - \epsilon
\] (274)

and so
\[
\sup_{t \geq t_+} \left[ \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_t) \geq \xi_N} C_j(G^N_t) \right] \leq \sup_{t \geq t_+} \left[ \frac{1}{N} \sum_{j \geq 2} C_j(G^N_t) \right] < \epsilon.
\] (275)

Now, consider the events
\[
A^3_N = \left\{ \sup_{t \in [t_g, t_-]} \left[ \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_t) \geq \xi_N} C_j(G^N_t) \right] < \epsilon \right\} ;
\] (276)
\[
A_N = A^1_N \cap A^2_N \cap A^3_N.
\] (277)

By Lemma 11.2, \( \mathbb{P}(A^3_N) \to 1 \), and so \( \mathbb{P}(A_N) \to 1 \). On the event \( A_N \), we have
\[
\sup_{t \geq t_g} \left[ \frac{1}{N} \sum_{j : 2 \leq C_j(G^N_t) \geq \xi_N} C_j(G^N_t) \right] < \epsilon
\] (278)
which proves the claimed convergence in probability.

We also remark that this is also true, and much simpler, in the subcritical and critical regimes \( t \leq t_g \).
Lemma 11.3. Let $G_t^N$ and $\xi_N$ be as above. Then
\[
\sup_{t \leq t_0} \left[ \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right] \rightarrow 0 \quad \text{in probability.} \tag{279}
\]

Proof. Following a similar argument as in the supercritical case, we bound for $t \leq t_0$,
\[
\frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \leq \frac{1}{N} \sum_{j \geq 1} C_j(G_t^N).
\tag{280}
\]
Therefore
\[
\sup_{t \leq t_0} \left[ \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right] \leq \frac{1}{N} C_1(G_t^N) + \frac{1}{N} \sum_{j \geq 1} C_j(G_t^N). \tag{281}
\]
By Theorems 4.2, 4.4, both terms converge to 0 in probability.

We now use these results to deduce an analogous result for the momentum and energy of these ‘mesoscopic clusters’.

Corollary 11.4. Let $G_t^N$ be as above. In the notation of (126), the momentum and energy of the anomalous clusters satisfy
\[
\sup_{t \geq 0} \left[ \frac{1}{N} \sum_{j \geq 2} P(C_j(G_t^N)) \right] \rightarrow 0 \quad \text{in probability;} \tag{282}
\]
\[
\sup_{t \geq 0} \left[ \frac{1}{N} \sum_{j \geq 2} E(C_j(G_t^N)) \right] \rightarrow 0 \quad \text{in probability.} \tag{283}
\]

Proof. By the law of large numbers and (A3.), we can find a constant $\eta < \infty$ such that the events
\[
A_N = \left\{ \frac{1}{N} \sum_{i=1}^{N} |v_i|^2 \leq \eta; \quad \frac{1}{N} \sum_{i=1}^{N} \frac{1}{4} |v_i|^4 \leq \eta \right\}
\tag{284}
\]
have probability $\mathbb{P}(A_N) \rightarrow 1$. On this event, we use Cauchy-Schwarz to estimate, for any $t \geq 0$,
\[
\left\{ \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right\} \leq \left( \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right)^{\frac{1}{2}} \left( \frac{1}{N} \sum_{i=1}^{N} |v_i|^2 \right)^{\frac{1}{2}} \tag{285}
\]
\[
\leq \sqrt{\eta} \left( \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right)^{\frac{1}{2}},
\]
and similarly
\[
\sum_{j \geq 2} C_j(G_t^N) \leq \sqrt{\eta} \left( \frac{1}{N} \sum_{j \geq 2} C_j(G_t^N) \right)^{\frac{1}{2}}. \tag{286}
\]
In each case, the right-hand side converges uniformly to 0 in probability by Lemmas 11.1, 11.3, which proves the claimed convergence.
11.2 Proof of Theorem 1.2

We can now prove the claimed convergence of the gel. The result is broken up into several Lemmas, which together prove the remainder of Theorem 1.2.

**Lemma 11.5.** Let $\mu_0$ be a measure satisfying (A1-4.) for a probability measure $m$, and let $g_t = (M_t, P_t, E_t)$ be the mass, momentum, and energy of the gel in the solution $(\mu_t)_{t \geq 0}$ to $(E+G)$. Let $(G_t^N)_{t \geq 0}$ be the graph process constructed in Section 5, with $N$ particles sampled independently for the same base measure $m$, and write $g_t^N = (M_t^N, P_t^N, E_t^N)$ for the data of the stochastic gel:

$$g_t^N = \frac{1}{N}(M(C_1(G_t^N)), P(C_1(G_t^N)), E(C_1(G_t^N))) = \frac{1}{N} \sum_{i \in C_1(G_t^N)} (1, v_i, \frac{1}{2}|v_i|^2).$$

Then we have uniform convergence in probability:

$$\sup_{t \geq 0} |g_t^N - g_t| \to 0 \quad \text{ in probability.} \quad (287)$$

Combined with Lemma 9.2, this proves the first assertion of Theorem 1.2.

**Proof.** The following proof is very similar to that of Lemma 6.3, with suitable modifications to work in a uniform setting; for ease of readability, we will recall all necessary constructions. Throughout, we let $(\mu_t^N)_{t \geq 0}$ be a stochastic coagulant coupled to a random graphs process $(G_t^N)_{t \geq 0}$, as described in Section 5. We write $(v_i)^N_{i=1}$ for the velocities associated to the graph vertexes.

Throughout, we will deal with the mass $M_t^N$; we will discuss the necessary modifications for momentum $P_t^N$ and energy $E_t^N$ at the end of the proof.

We will deal first with the critical and supercritical regimes $t \geq t_g$. Following the proof of Lemma 6.3, let $\xi_N$ be a sequence, to be constructed later, such that

$$\xi_N \to \infty; \quad \frac{\xi_N}{N} \to 0; \quad \mathbb{P}(C_1(G_{t_g}^N) \geq \xi_N) \to 1. \quad (289)$$

We recall the following construction of ‘bump functions’ from Lemma 6.3. First, using Lemma 6.5, we construct a sequence $\eta_r \to \infty$ such that

$$\beta(r, \eta_r) := \sup_{N \geq 1} \mathbb{E} \left[ \sup_{t \geq 0} \langle \varphi_1[\pi_e(x) > \eta_r, \pi_n(x) < r], \mu_t^N \rangle \right] \to 0; \quad (290)$$

We let $S_r$ be the set

$$\{x : \pi_n(x) < r, |\pi_p(x)| \leq \sqrt{2r}\eta_r, \pi_e(x) \leq \eta_r\}. \quad (291)$$

Let $\bar{g}_r$ be the indicator $\bar{g}_r = 1[\pi_n(x) < r]$, and construct a continuous, compactly supported function $\bar{f}_r$ such that

$$0 \leq \bar{f}_r \leq 1; \quad \bar{f}_r = 1 \text{ on } S_r; \quad \bar{f}_r(x) = 0 \text{ if } \pi_n(x) \geq r. \quad (292)$$

By symmetrising if necessary, we also assume that $f_r(Rx) = f_r(x)$, for all $x \in S$. We define $f_N = \bar{f}_\xi_N$ and $g_N = \bar{g}_\xi_N$.

As in Lemma 6.3, we can decompose the difference $M_t^N - M_t$ as

$$M_t^N - M_t = (1 - M_t - \langle \pi_n f_N, \mu_t \rangle) + \langle \pi_n f_N, \mu_t^N - \mu_t \rangle + \langle \pi_n (f_N - g_N), \mu_t^N \rangle + \langle \pi_N g_N, \mu_t^N - (1 - M_t^N) \rangle. \quad (293)$$

We now modify the argument of Lemma 6.3 to estimate the errors $\mathcal{T}_N^i$, $i = 1, 3, 4$, uniformly in $t \geq t_g$. As in Lemma 6.3, we ensure uniform convergence of $\mathcal{T}_N^2$ later with a particular construction of the sequence $\xi_N$. 

48
1. Estimate on $T^1_N$. Let $h_N = 1_{S(t_N)}$, so that $h_N \leq f_N \leq 1$. As $N \to \infty$, $\pi_N h_N \uparrow \pi$, and so by monotone convergence,

$$\langle \pi_N h_N, \mu_t \rangle \uparrow \langle \pi, \mu_t \rangle = 1 - M_t.$$  
(294)

Moreover, each function $t \mapsto \langle \pi_N h_N, \mu_t \rangle$ is continuous on $[0, \infty)$ by the definition (E+G) of the Smoluchowski dynamics, and the limit function $t \mapsto 1 - M_t$ is continuous on $[0, \infty)$ by Lemma 8.1. We extend the maps $t \mapsto \langle \pi_N h_N, \mu_t \rangle$ to the compactification $[0, \infty]$, by defining both to be equal to 0 at infinity. By Lemma 9.1, it follows that $\langle \pi_N, \mu_t \rangle, \langle \pi_N h_N, \mu_t \rangle \to 0$ as $t \to \infty$, so both of these extensions are continuous on the whole compactification $[0, \infty]$. Therefore, by Dini’s Theorem, the convergence is uniform in $t \geq t_g$, and since

$$\langle \pi_N h_N, \mu_t \rangle \leq \langle \pi_N f_N, \mu_t \rangle \leq \langle \pi, \mu_t \rangle = 1 - M_t,$$  
(295)

it follows that $T^1_N(t) \to 0$, uniformly in $t \geq t_g$.

2. Estimate on $T^3_N$. From the definitions of $f_N, g_N$, we observe that

$$|T^3_N(t)| = \langle \pi_n(g_N - f_N), \mu^N_t \rangle \leq \langle \pi_n 1[\pi_n(x) < \xi_N], \pi(x) > \eta_N], \mu^N_t \rangle.$$  
(296)

Therefore, in the notation of (290),

$$E \left[ \sup_{t \geq t_g} |T^3_N(t)| \right] \leq \beta(\xi_N, \eta_N).$$  
(297)

By construction of $\eta_r$, and since $\xi_N \to \infty$, the right hand side converges to 0. Therefore, $T^3_N$ converges to 0, uniformly in probability.

3. Estimate on $T^4_N$. By the choice (289) of $\xi_N$, we have that $P(\forall t \geq t_g, C_1(G^N_t) \geq \xi_N) \to 1$. On this event, we have the equality

$$\langle \pi_n g_N, \mu^N_t \rangle = \langle \pi_n, \mu^N_t \rangle - \langle \pi_n 1[\pi_n \geq \xi_N], \mu^N_t \rangle = 1 - M^N_t - \frac{1}{N} \sum_{j \geq 2} C_j(G^N_t).$$  
(298)

where $(G^N_t)_{t \geq 0}$ is the random graph process coupled to the stochastic coagulant. Therefore, with high probability, for all $t \geq t_g$,

$$T^4_N(t) = \frac{1}{N} \sum_{j \geq 2} C_j(G^N_t)$$  
(299)

which converges to 0, uniformly in probability on $t \geq t_g$, by Lemma 11.1.

4. Construction of $\xi_N$, and convergence of $T^2_N$. To conclude the proof of the supercritical case, it remains to show how a sequence $\xi_N$ can be constructed such that $T^2_N \to 0$ uniformly, in probability. Let $A^1_{r,N}, A^2_{r,N}$ be the events

$$A^1_{r,N} = \left\{ \sup_{t \geq 0} |\langle \pi_n f_r, \mu^N_t - \mu_t | < \frac{1}{r} \right\}; \quad A^2_{r,N} = \left\{ C_1(G^N_{t_g}) \geq r \right\}.$$  
(300)

Then, as $N \to \infty$, both $P(A^1_{r,N}), P(A^2_{r,N}) \to 1$, by Lemmas 9.2, 6.4. We now define $N_r$ inductively for $r \geq 1$ inductively, as in Lemma 6.3, by setting $N_1 = 1$ and letting $N_{r+1}$ be the minimal $N > N_r$ such that, for all $N' \geq N$,

$$N \geq \max((r + 1)^2, N_r + 1); \quad P(A^1_{r,1,N}) > \frac{r}{r + 1}; \quad P(A^2_{r+1,N}) > \frac{r}{r + 1}.$$  
(301)

Now, we set $\xi_N = r$ for $N \in [N_r, N_{r+1}) \cap N$. It follows that $\xi_N \to \infty$ and $\xi_N \leq \sqrt{N} \ll N$, and

$$P \left( C_1(G^N_{t_g}) \geq \xi_N \right) \geq 1 - \frac{1}{\xi_N} \to 1.$$  
(302)

49
Therefore, $\xi_N$ satisfies the requirements (289) above. Moreover,

$$\mathbb{P}\left(\sup_{t \geq t_k} |T^2_N| < \frac{1}{\xi_N}\right) \geq \mathbb{P}\left(A^1_{\xi_N,N}\right) > 1 - \frac{1}{\xi_N} \to 1$$

(303)

and so, with this choice of $\xi_N$, $T^2_N \to 0$ uniformly in probability on $t \geq t_k$.

This concludes the argument for the mass $M^N_t$ of the stochastic gel, in the critical and supercritical regime $t \geq t_g$. The cases for the momentum $P^N_t$ and energy $E^N_t$ are essentially identical, with the following minor differences.

i). For the momentum, by symmetry under $R$ we have

$$\langle \pi_p f_N, \mu_t \rangle = 0; \quad \langle \pi_p g_N, \mu_t \rangle = 0.$$  

(304)

Therefore, for the momentum term, $T^3_N$ is identically 0. For the energy, we argue by Dini as above.

ii). For the momentum, we relate the error term $T^3_N$ to the equivalent terms for the mass and energy, using the bound $|\pi_p(x)| \leq \sqrt{2\pi_n(x)}|\pi_e(x)\leq \varphi(x)$ by Cauchy-Schwarz. In the notation of (290), this produces the bound

$$\mathbb{E}\left[\sup_{t \geq t_k} |T^3_N(t)|\right] \leq \beta(\xi_N, \eta_N) \to 0.$$  

(305)

iii). In the decomposition (293) for the momentum (resp. energy) terms, there is an additional error

$$T^3_N = \langle \pi_p, \mu^N_0 - \mu_0 \rangle \quad \text{(resp.} \quad \langle \pi_e, \mu^N_0 - \mu_0 \rangle).$$  

(306)

Each of these converge to 0 in probability, by the law of large numbers.

To conclude the proof for $M^N_{t}$, we show uniform convergence in the subcritical phase $t < t_g$. In this region, we have $M_t = E_t = 0$ and $P_t = 0$. Hence

$$\sup_{t < t_k} |(M^N_t - M_t)| = \sup_{t < t_k} \left[\frac{1}{N}C_1(G^N_t)\right] = \frac{1}{N}C_1(G^N_{t_k}) \to 0$$

(307)

in probability, by Theorem 4.2. The cases for momentum and energy are similar, using Cauchy-Schwarz as in (285, 286).

\[\square\]

In order to complete the proof of Theorem 1.2, we wish to prove a similar result, when $g^N_t$ is replaced by cutting off at a suitable deterministic scale $\xi_N$:

$$\tilde{g}^N_t = (\tilde{M}^N_t, \tilde{P}^N_t, \tilde{E}^N_t)$$

$$= \frac{1}{N} \sum_{j \geq 1: C_j(G^N_t) \geq \xi_N} (M(C_j(G^N_t)), P(C_j(G^N_t)), E(C_j(G^N_t)))$$

(308)

$$= \left(\langle \pi_n \chi[\pi_n \geq \xi_N], \mu^N_t \rangle, \langle \pi_p \chi[\pi_n \geq \xi_N], \mu^N_t \rangle, \langle \pi_e \chi[\pi_n \geq \xi_N], \mu^N_t \rangle\right).$$

Together with the previous lemma, it suffices to prove the following.

**Lemma 11.6.** Let $\xi_N$ be a sequence such that

$$\xi_N \to \infty; \quad \frac{\xi_N}{N} \to 0.$$  

(309)

Let $g^N_t$ be as in Lemma 11.5, and $\tilde{g}^N_t$ be as in (308). Then

$$\sup_{t \geq 0} |g^N_t - \tilde{g}^N_t| \to 0 \quad \text{in probability.}$$

(310)
Proof. Let $K_\ell$ be the symmetric difference of the clusters considered:

$$K_\ell = \{1\} \triangle \{ j : C_j(G_\ell^N) \geq \xi_N \} = \begin{cases} \{1\} & \text{if } C_1(G_\ell^N) < \xi_N; \\ \{ j \geq 2 : C_j(G_\ell^N) \geq \xi_N \} & \text{if } C_1(G_\ell^N) \geq \xi_N \end{cases} \quad (311)$$

By considering the two cases separately, we observe that

$$\left| M_\ell^N - \tilde{M}_\ell^N \right| = \frac{1}{N} \sum_{j \in K_\ell} C_j(G_\ell^N) \quad (312)$$

with similar equalities for the momentum and energy. As usual, we deal with subcritical and supercritical cases separately; in this case, we will see that it is more natural to split at a time $t_+$ slightly larger than $t_g$.

1. **Subcritical Case.** Let $\epsilon > 0$. By continuity in Lemma 8.1, we can choose $t_+ > t_g$ such that $M_{t_+} < \epsilon$. Then, for $t \leq t_+$, we bound

$$\left| M_\ell^N - \tilde{M}_\ell^N \right| \leq M_\ell^N + \frac{1}{N} \sum_{j \geq 2 : C_j(G_\ell^N) \geq \xi_N} C_j(G_\ell^N). \quad (313)$$

By Lemma 11.5, the first term converges to $M_\ell < \epsilon$ uniformly in probability, and by Lemmas 11.1, 11.3, the second term converges in probability to 0, uniformly in $t \leq t_+$. Therefore, the event

$$A_1^N = \left\{ \sup_{t \leq t_+} M_t^N - \tilde{M}_t^N \right\} < \epsilon \quad (314)$$

has $\mathbb{P}(A_N) \rightarrow 1$.

2. **Supercritical Case.** Since $t_+ > t_g$, there is a giant component, and so $C_1(G_{t_+}^N) > \xi_N$ with high probability. On this event, for all $t \geq t_+$, we have

$$\left| M_\ell^N - \tilde{M}_\ell^N \right| = \frac{1}{N} \sum_{j \geq 2 : C_j(G_\ell^N) \geq \xi_N} C_j(G_\ell^N) \quad (315)$$

which converges to 0, uniformly in probability, by Lemma 11.1. Therefore,

$$A_2^N = \left\{ \sup_{t \geq t_+} M_t^N - \tilde{M}_t^N \right\} < \epsilon \quad (316)$$

has probability $\mathbb{P}(A_N^2) \rightarrow 1$.

Combining the two cases above, we have shown that

$$\mathbb{P} \left( \sup_{t \geq 0} \left| M_t^N - \tilde{M}_t^N \right| > \epsilon \right) \rightarrow 1 \quad (317)$$

as desired. For the momentum and energy, we bound

$$\left| P_t^N - \tilde{P}_t^N \right| \leq \frac{1}{N} \sum_{j \in K_\ell} |P(C_j(G_t^N))|; \quad \left| E_t^N - \tilde{E}_t^N \right| \leq \frac{1}{N} \sum_{j \in K_\ell} |E(C_j(G_t^N))|. \quad (318)$$

By using the case for the mass, and using Cauchy-Schwartz as in (285, 286), we see that both right-hand-sides converge to 0, uniformly in probability, as $N \rightarrow \infty$.  \qed
References

[1] Author, Article title, Journal, Volume, page numbers (year)
[2] Author, Book title, page numbers. Publisher, place (year)
[3] Aldous, D., The Incipient Giant Component in Bond Percolation on General Finite Weighted graphs, Electronic Communications in Probability, pp21 (2016).
[4] Aldous, D.J., Deterministic and Stochastic Models for Coalescence (Aggregation and Coagulation): a Review of the Mean-Field Theory for Probabilists, Bernoulli, 5(1), pp.3-48 (1999).
[5] Bollobás, B., The Evolution of Random Graphs, Transactions of the American Mathematical Society, 286(1), pp.257-274 (1984).
[6] Bollobás, B., Random graphs (No. 73), Cambridge University Press, Cambridge (2001).
[7] Bollobás, B., Janson, S. and Riordan, O., The Phase transition in Inhomogeneous Random Graphs, Random Structures & Algorithms, 31(1), pp.3-122 (2007).
[8] Boltzmann, L., Vorlesungen über Gastheorie (Vol. 1).JA Barth, Leipzig (1896).
[9] Buffet, E. and Pulé, J. V., Polymers and Random Graphs, Journal of Statistical Physics, 64, pp.87–110 (1991).
[10] Darling, R.W.R. and Norris, J.R., Differential Equation Approximations for Markov Chains. Probability surveys, 5, pp.37-79 (2008).
[11] Erdős, P. and Rényi, A., On the Evolution of Random Graphs, Magyar Tud. Akad. Mat. Kutató Int. Közl, 5, pp.17-61 (1960).
[12] Ethier, S.N. and Kurtz, T.G., Markov Processes: Characterization and Convergence Vol. 282. John Wiley & Sons, New York (2009).
[13] Flory, P. J., Molecular Size Distribution in Three Dimensional Polymers I. Gelation, Journal of the American Chemical Society, 63, pp.3083-3090 (2008).
[14] Gabrielov, A., Keilis-Borok, V., Sinai, Y. and Zaliapin, I., Statistical Properties of the Cluster Dynamics of the Systems of Statistical Mechanics, Boltzmann’s Legacy, pp.203-215 (2008).
[15] Grünbaum, F.A., Propagation of Chaos for the Boltzmann Equation, Archive for Rational Mechanics and Analysis, 42(5), pp.323-345 (1971).
[16] Jakubowski, A., On the Skorokhod Topology, Ann. Inst. H. Poincaré Probab. Statist, 22(3), pp.263-285 (1986).
[17] Janson, S., Luczak, T. and Rucinski, A., Random Graphs, Vol. 45. John Wiley & Sons, New York (2011).
[18] Jeon, I., Existence of Gelling Solutions for Coagulation-Fragmentation Equations, Communications in Mathematical Physics, 194, pp.541-567 (1998).
[19] Kac, M., Foundations of Kinetic Theory, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 3, pp. 171-197 (1956).
[20] Kallenberg, O., Foundations of Modern Probability. Springer Science & Business Media, New York (2006).
[21] Lanford III, O.E., On a Derivation of the Boltzmann Equation, Nonequilibrium Phenomena I: The Boltzmann Equation, p.1, 1986.
[22] Lanford, O.E., Time Evolution of Large Classical Systems, Dynamical systems, Theory and Applications, pp. 1-111. Springer, Berlin, Heidelberg (1975).

[23] Leontovich, M.A., Basic equations of the Kinetic Gas Theory from the Point of View of the Theory of Random Processes. Zh. Teoret. Ehksper. Fiz, 5, pp.211-231 (1935).

[24] Lu, X. and Mouhot, C., On measure solutions of the Boltzmann Equation, Part I: Moment Production and Stability Estimates. Journal of Differential Equations, 252(4), pp.3305-3363 (2012).

[25] Lushnikov, A.A., Coagulation in Finite Systems, Journal of Colloid and interface science, 65(2), pp.276-285 (1978).

[26] McKean Jr, H.P., An Exponential Formula for Solving Boltzmann’s Equation for a Maxwellian Gas, Journal of Combinatorial Theory, 2(3), pp.358-382 (1967).

[27] Mischler, S. and Mouhot, C., Kac’s Program in Kinetic Theory, Inventiones Mathematicae, 193(1), pp.1-147 (2013).

[28] Normand, R., A Model for Coagulation with Mating, Journal of Statistical Physics, 137, pp.343-371 (2009).

[29] Normand, R. and Zambotti, L., Uniqueness of Post-Gelation Solutions of a Class of Coagulation Equations, Annales de l’Institut Henri Poincaré (C) Non Linear Analysis, 28(2), pp.189-215 (2011).

[30] Norris, J.R., Smoluchowski’s Coagulation Equation: Uniqueness, Nonuniqueness and a Hydrodynamic Limit for the Stochastic Coalescent, Annals of Applied Probability, pp.78-109 (1999).

[31] Norris, J.R., Cluster Coagulation, Communications in Mathematical Physics, 209(2), pp.407-435 (2000).

[32] Norris, J., A Consistency Estimate for Kac’s Model of Elastic Collisions in a Dilute Gas, The Annals of Applied Probability, 26(2), pp.1029-1081 (2016).

[33] Patterson, R.I.A., Simonella, S. and Wagner, W., Kinetic Theory of Cluster Dynamics, Physica D: Non-linear Phenomena, 335, pp.26-32 (2016).

[34] Patterson, R.I.A., Simonella, S. and Wagner, W., A Kinetic Equation for the Distribution of Interaction Clusters in Rarefied Gases, Journal of Statistical Physics, 169(1), pp.126-167 (2017).

[35] Pulvirenti, M. and Simonella, S., The Boltzmann-Grad Limit of a Hard Sphere System: Analysis of the Correlation Error. Inventiones mathematicae, 207(3), pp.1135-1237 (2017).

[36] von Smoluchowski, M., Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Koloidteilchen, Z. Phys., 17, pp.557-585 (1916).

[37] Sznitman, A.S., Équations de type de Boltzmann, Spatialement Homogenes, Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete, 66(4), pp.559-592 (1984).

[38] Villani, C., Cercignani’s Conjecture is Sometimes True and Always Almost True, Communications in Mathematical Physics, 234(3), pp.455-490 (2003).

[39] Ziff, R.M. and Stell, G., Kinetics of Polymer Gelation, The Journal of Chemical Physics, 73(7), pp.3492-3499 (1980).