A LARGE-DEVIATIONS APPROACH TO GELATION

LUISA ANDREIS, WOLFGANG KÖNIG, AND ROBERT I. A. PATTERSON

(07 January, 2019)

Abstract. A large-deviations principle (LDP) is derived for the state at fixed time, of the multiplicative coalescent in the large particle number limit. The rate function is explicit and describes each of the three parts of the state: microscopic, mesoscopic and macroscopic. In particular, it clearly captures the well known gelation phase transition given by the formation of a particle containing a positive fraction of the system mass. Via a standard map of the multiplicative coalescent onto a time-dependent version of the Erdős-Rényi random graph, our results can also be rephrased as an LDP for the component sizes in that graph. The proofs rely on estimates and asymptotics for the probability that smaller Erdős-Rényi graphs are connected.

MSC 2010: 05C80, 60F10, 60K35, 82B26.

Keywords and phrases. Coagulation process, multiplicative coalescent, gelation, phase transition, large-deviations, Erdős-Rényi random graph.

1. Introduction

Smoluchowski introduced a (deterministic) ODE model for the concentrations of coagulating particles in the course of his work on Brownian motion [vS16]. This was motivated by an underlying concept of microscopic dynamics similar to the mean-field limit. Flory [Flo41a] investigated the distributions of polymer sizes and connectivity structures in order to understand the physical properties of these (at that time relatively new) materials. As is now well known the particle and random graph models are almost equivalent, but unlike Smoluchowski, Flory was interested in the phase transition that occurs when a giant particle/connected component of the graph forms, which he called a *gel*, terminology which we shall adopt.

In this paper, we study one of the simplest stochastic coagulation models and give a complete description by means of a powerful mathematical theory, the large-deviations theory. This turns out to be equivalent to the description, in the large-deviations framework of the statistics of connected components in the sparse regime for the Erdős-Rényi random graph, i.e. $\mathcal{G}(N,p)$ when $p \sim \frac{1}{N}$. We provide a new analysis to the gelation phase transition via the large-deviations rate function. This then provides an alternative proof of the uniqueness of the giant component in sparse Erdős-Rényi random graphs.

1.1. **A non-spatial mean-field coagulation model.** In this paper, we study a stochastic coagulation process, called the *Marcus–Lushnikov process*, see [Mar68, Gil72, Lus78]. This is a continuous-time Markov process of vectors of particle masses $M_i^{(N)}(t) \in \mathbb{N}$ at time $t \in [0,\infty)$, arranged in descending order:

$$M_1^{(N)}(t) \geq M_2^{(N)}(t) \geq M_3^{(N)}(t) \geq \cdots \geq M_{n(t)}^{(N)}(t) \geq 1, \qquad \sum_{i=1}^{n(t)} M_i^{(N)}(t) = N, \tag{1.1}$$

for some parameter $N \in \mathbb{N}$. This process is specified by the initial configuration, which we take in the monodisperse case, i.e. $M_i^{(N)}(0) = 1$ for all $i = 1, \dots, N = n(0)$, and by the transition mechanism, which is given in terms of a symmetric, non-negative *coagulation kernel* $K_N\colon \mathbb{N}\times\mathbb{N}\to[0,\infty)$. That is, we start with N particles of unit mass at time 0, and in the course of the process, each (unordered) pair of particles with respective masses $m,\widetilde{m}\in\mathbb{N}$ coagulate to a particle of mass $m+\widetilde{m}$ with rate $K_N(m,\widetilde{m})$, independently of all the other pairs of particles.

If the dependence of the coagulation kernel K_N on N is chosen so that $K_N(m,\widetilde{m})\asymp 1/N$ for fixed $m,\widetilde{m}$ as $N\to\infty$, then any given pair of finitely sized particles coagulates after a time $\asymp N$ and so each particle undergoes a coagulation with some other particle after an order 1 time, since it is in contact with $\asymp N$ other particles. This is a mean field interaction in which every particle in a large population has contact with every other on an equal basis. In this limit it is reasonable to write $l_k(t)=\lim_{N\to\infty}\frac{1}{N}\#\{\text{particles of size } k \text{ at time } t\}$, then under suitable conditions these limits satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}l_k(t)=\frac{1}{2}\sum_{\substack{m,\widetilde{m}:\\ m+\widetilde{m}=k}} l_m(t)l_{\widetilde{m}}(t)K(m,\widetilde{m})-l_k(t)\sum_m l_m(t)K(k,m)\qquad\forall k\in\mathbb{N}\tag{1.2}$$

where $K(m,\widetilde{m})=\lim_{N\to\infty}NK_N(m,\widetilde{m})$. This is the *Smoluchowski equation* [vS16] referred to above. In this paper, we exclusively study the case of the *multiplicative kernel*, $K_N(m,\widetilde{m})=m\widetilde{m}/N$. This choice has the two interesting features: (1) it can be mapped onto a natural time-dependent version of the Erdős-Rényi random graph [ER61], and (2) it exhibits an interesting phase transition in the limit $N\to\infty$ at time $t=1$, because a *gel*, i.e., a particle of macroscopic size, appears. Our main goal in this paper is to recover the gelation phenomenon in rather explicit terms through a *large-deviations principle (LDP)*. Moreover, we will be able to describe the large-deviations of all parts of the particle model, the microscopic, mesoscopic and macroscopic parts. As a consequence, the above mentioned phase transition as well as the solution of the Smoluchowski ODE will be clear from our formulas and will be given a new interpretation in terms of combinatorial structures. Indeed, we will analyse the joint distribution of the *microscopic* and the *macroscopic* empirical measure of the particle sizes, and we will also deduce interesting information about the *mesoscopic* part of the configuration. We will keep the time $t\in(0,\infty)$ fixed and consider only the limit as $N\to\infty$.

There is a well-known description of the distribution of the coagulation process that we study in the present work at time t in terms of the well-known Erdős-Rényi random graph on N nodes with edge probability $1-\mathrm{e}^{-t/N}$, see the review [Ald99]. More precisely, the joint distribution of all the cluster sizes in this graph is equal to the distribution of the particle sizes, $(M_i^{(N)}(t))_i$. Hence, the formation of a gel in the coagulation process is equivalent to occurrence of percolation in the graph, i.e., formation of a giant component, see the classic reference [Bél01]. The connection between the two models will be the starting point of our analysis and will be recalled at the beginning of Section 2.1. Because of this connection, our results give a new contribution to the theory of the Erdős-Rényi graph in terms of an LDP for connected component statistics in the sparse regime. These kind of asymptotic results were not previously available in random graph theory, even though there are large-deviations results of various types, see Section 1.5 for an overview.

1.2. **Our results: Large-deviations principles.** In this section, we present all our results on the exponential behaviour of distributions of the main characteristics of the Marcus–Lushnikov model. In Section 1.3 we will draw conclusions about the gelation phase transition from that.

For $N\in\mathbb{N}$ we consider the state space

$$S_N=\Big\{(m_i)_{i=1,\dots,n}\in\mathbb{N}^n\colon n\in\mathbb{N}, n\le N, m_1\ge m_2\ge m_3\ge\cdots\ge m_n\ge 1,\sum_{i=1}^n m_i=N\Big\}\tag{1.3}$$

of tuples of positive integers summing to N, ordered in a decreasing way. Starting from the initial configuration $M^{(N)}(0)=(1,\dots,1)\in S_N$, the Markov process $(M^{(N)}(t))_{t\in[0,\infty)}$ specified in the previous section has the generator

$$\mathcal{L}_N f\big((m_i)_{i=1,\dots,n}\big) = \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{n} K_N(m_i,m_j)\Big[f(\widetilde{m}^{(i,j)}) - f\big((m_i)_{i=1,\dots,n}\big)\Big],$$

where $\widetilde{m}^{(i,j)}$ is the collection of the $n-1$ numbers m_k with $k\neq i,j$ and m_i+m_j, properly ordered. Here we restrict to the multiplicative kernel, i.e. $K_N(m,\widetilde{m}) = m\widetilde{m}/N$. That is, each pair of distinct particles coagulates after an exponential time with mean $N/m\widetilde{m}$ where m and $\widetilde{m}$ are the particle masses. Coagulation means that the two particles are replaced by one particle of mass $m+\widetilde{m}$. All the exponential waiting times are independent. Then the number $n(t)$ of particles at time t is a decreasing function of t, and with probability one, it reaches the value one in finite time.

We denote the probability and expectation by $\mathbb{P}_N$ and $\mathbb{E}_N$, respectively. We fix a time horizon $t\in(0,\infty)$ and describe the distribution of the particle masses $M^{(N)}(t)$ in (1.1) in the limit $N\to\infty$ in terms of a large-deviations principle. It will be convenient to work with the *empirical measures* of the particle masses in the microscopic and macroscopic size ranges:

$$\mathrm{Mi}^{(N)}(t) = \frac{1}{N}\sum_{i=1}^{n(t)} \delta_{M_i^{(N)}(t)} \qquad \text{and} \qquad \mathrm{Ma}^{(N)}(t) = \sum_{i=1}^{n(t)} \delta_{\frac{1}{N}M_i^{(N)}(t)}. \tag{1.4}$$

Intuitively, while $\mathrm{Mi}^{(N)}(t)$ registers the numbers of particles of “microscopic” sizes 1,2,3,... on the scale N, in contrast $\mathrm{Ma}^{(N)}(t)$ registers the numbers of particles of “macroscopic” sizes of order N. Even though each of the two measures admits a one-to-one map onto the vector $(M_i^{(N)}(t))_i$ for fixed $N\in\mathbb{N}$, in the limit $N\to\infty$, for topological reasons, they will be able to describe only the statistics of the microscopic, respectively macroscopic, part of the particle configuration. For a full description, a kind of mesoscopic part has to be considered, but this is a more complicated issue, which we defer.

$\mathrm{Mi}^{(N)}(t)$ is a random element of the set $\mathcal{N} = \bigcup_{c\in[0,1]}\mathcal{N}(c)$ of measures on $\mathbb{N}$ that have an integral against the identity not larger than one, where

$$\mathcal{N}(c) = \Big\{\lambda\in[0,\infty)^{\mathbb{N}}\colon \sum_{k\in\mathbb{N}} k\lambda_k = c\Big\}, \qquad c>0. \tag{1.5}$$

We equip $\mathcal{N}$ with the topology of coordinate-wise convergence, which is compact by the Bolzano-Weierstrass theorem combined with Fatou's lemma.

$\mathrm{Ma}^{(N)}(t)$ is a random element of the set $\mathcal{M}_{\mathbb{N}_0} = \bigcup_{c\in[0,1]}\mathcal{M}_{\mathbb{N}_0}(c)$, where

$$\mathcal{M}_{\mathbb{N}_0}(c) = \Big\{\alpha\in\mathcal{M}_{\mathbb{N}_0}((0,1])\colon \int_{(0,1]} x\,\alpha(\mathrm{d}x) = c\Big\}, \tag{1.6}$$

and $\mathcal{M}_{\mathbb{N}_0}((0,1])$ is the set of all measures on $(0,1]$ with values in $\mathbb{N}_0=\{0\}\cup\mathbb{N}$. We equip $\mathcal{M}_{\mathbb{N}_0}$ with the topology that is induced by functionals of the form $\mu\mapsto\int_{(0,1]} f(x)\,\mu(\mathrm{d}x)$ where $f\colon(0,1]\to\mathbb{R}$ is continuous and compactly supported. We always write the elements of $\mathcal{M}_{\mathbb{N}_0}(c)$ as $\alpha=\sum_j\delta_{\alpha_j}$ with $1\geq\alpha_1\geq\alpha_2\geq\dots>0$ and $\sum_j\alpha_j = c$, where j extends over a finite subset of $\mathbb{N}$ or over $\mathbb{N}$. Then convergence is equivalent with the pointwise convergence of each of the atoms. By similar arguments as for $\mathcal{N}$, also $\mathcal{M}_{\mathbb{N}_0}$ is compact.

Note that the microscopic and the macroscopic total masses $\sum_k k\mathrm{Mi}_k^{(N)}(t)$ and $\int_{(0,1]} x\,\mathrm{Ma}^{(N)}(t)(\mathrm{d}x)$ are each equal to one, hence indeed $\mathrm{Mi}^{(N)}(t)\in\mathcal{N}(1)$ and $\mathrm{Ma}^{(N)}(t)\in\mathcal{M}_{\mathbb{N}_0}(1)$. However the functions

$$\lambda\mapsto c_\lambda := \sum_{k\in\mathbb{N}} k\lambda_k \qquad \text{and} \qquad \alpha\mapsto c_\alpha := \int_{(0,1]} x\,\alpha(\mathrm{d}x)$$

are not continuous, but only lower semicontinuous in the respective topologies.

We equip the product of $\mathcal{N}$ and $\mathcal{M}_{\mathbb{N}_0}$ with the product topology, so that it is also compact.

Our main result is the following description of the two empirical measures in terms of a joint large-deviations principle (LDP).

Theorem 1.1 (LDP for the empirical measures)**.** *Fix $t \in [0,\infty)$. Then, as $N \to \infty$, the pair $(\mathrm{Mi}^{(N)}(t), \mathrm{Ma}^{(N)}(t))$ satisfies a large-deviations principle with speed N and rate function*

$$(\lambda,\alpha) \mapsto I(\lambda,\alpha;t) = \begin{cases} I_{\mathrm{Mi}}(\lambda;t) + I_{\mathrm{Ma}}(\alpha;t) + (1-c_\lambda-c_\alpha)\Big(\frac{t}{2}-\log t\Big), & \text{if } c_\lambda + c_\alpha \le 1,\\ \infty & \text{otherwise,}\end{cases}$$

where

$$I_{\mathrm{Mi}}(\lambda;t) = \sum_{k=1}^\infty \lambda_k \log \frac{k!t\lambda_k}{\mathrm{e}\,k^{k-2}} + c_\lambda\Big(1+\frac{t}{2}-\log t\Big), \qquad c_\lambda = \sum_{k=1}^\infty k\lambda_k, \tag{1.7}$$

$$I_{\mathrm{Ma}}(\alpha;t) = \int_0^1 \Big[x\log\frac{x}{1-\mathrm{e}^{-tx}} + \frac{t}{2}x(1-x)\Big]\,\alpha(\mathrm{d}x), \qquad c_\alpha = \int_{(0,1]} x\,\alpha(\mathrm{d}x). \tag{1.8}$$

The proof of this theorem is in Section 3. Let us recall the notion of an LDP: Theorem 1.1 says that, for any open set $G \subset \mathcal{N}\times\mathcal{M}_{\mathbb{N}_0}$ respectively closed set $F \subset \mathcal{N}\times\mathcal{M}_{\mathbb{N}_0}$,

$$\liminf_{N\to\infty} \frac{1}{N}\log\mathbb{P}_N((\mathrm{Mi}^{(N)}(t),\mathrm{Ma}^{(N)}(t))\in G) \ge -\inf_G I(\cdot;t),$$
$$\limsup_{N\to\infty} \frac{1}{N}\log\mathbb{P}_N((\mathrm{Mi}^{(N)}(t),\mathrm{Ma}^{(N)}(t))\in F) \le -\inf_F I(\cdot;t).$$

For the theory of large-deviations, see e.g. [DZ10]. It is not difficult to see that the rate function $I(\cdot,\cdot;t)$ is lower semicontinuous. Since $\mathcal{N}\times\mathcal{M}_{\mathbb{N}_0}$ is compact, it is even a good rate function, i.e., its level sets $\{(\lambda,\alpha)\colon I(\lambda,\alpha;t)\le r\}$ are compact for any r.

From our main result, the LDP in Theorem 1.1, a number of other LDPs follow via the contraction principle (which says that if a random variable satisfies an LDP, so does its image under a continuous transformation). Let us begin with the particle size distribution of the microscopic part.

Corollary 1.2 (LDP for particle size statistics)**.** *Fix $t\in[0,\infty)$. Then, as $N\to\infty$, $\mathrm{Mi}^{(N)}(t)$ satisfies an LDP with rate function $\mathcal{I}_{\mathrm{Mi}}(\cdot;t)\colon \mathcal{N}\to[0,\infty]$, given by*

$$\mathcal{I}_{\mathrm{Mi}}(\lambda;t) = \inf_{\alpha\in\mathcal{M}_{\mathbb{N}}} I(\lambda,\alpha;t) = I_{\mathrm{Mi}}(\lambda;t) - (1-c_\lambda)\left(\log\frac{1-\mathrm{e}^{(c_\lambda-1)t}}{1-c_\lambda} - \frac{c_\lambda t}{2}\right). \tag{1.9}$$

The first equality is the contraction principle [DZ10]; the second equality is checked in Lemma 4.1.

In the same way one can investigate the macroscopic part of the system.

Corollary 1.3 (LDP for macroscopic particles)**.** *Fix $t\in[0,\infty)$. Then, as $N\to\infty$, $\mathrm{Ma}^{(N)}(t)$ satisfies an LDP with rate function $\mathcal{I}_{\mathrm{Ma}}(\cdot;t)\colon\mathcal{M}_{\mathbb{N}_0}\to[0,\infty]$, given by*

$$\begin{aligned}\mathcal{I}_{\mathrm{Ma}}(\alpha;t) &= \inf_{\lambda\in\mathcal{N}} I(\lambda,\alpha;t)\\ &= I_{\mathrm{Ma}}(\alpha;t) + (1-c_\alpha)\Big(\frac{t}{2}-\log t\Big) + C_{\alpha,t}\Big(\log(tC_{\alpha,t}) - \frac{t}{2}C_{\alpha,t}\Big),\end{aligned} \tag{1.10}$$

where $C_{\alpha,t} = (1-c_\alpha)\wedge\frac{1}{t}$ (recall $c_\alpha = \int_0^1 x\,\alpha(\mathrm{d}x)$).

Only the second equality has to be checked; this is done in Lemma 4.2.

Hence, we can hope to derive a phase transition from non-existence to existence of a gel, i.e., to a non-trivial macroscopic part, at $t=1$, from the rate functions. However, even though it seems as if this phenomenon is present only in the macroscopic rate function $\mathcal{I}_{\mathrm{Ma}}$, actually it has its origin in the discussion of the existence of minimizers λ for the microscopic rate function $\mathcal{I}_{\mathrm{Mi}}$ in Corollary 1.2; this is the content of Theorem 1.5.

Now we come to the mesoscopic part of the particle configuration. Since this part comprises particle sizes on all the scales between finite and $O(N)$, it makes no sense to consider an empirical measure. Instead, we consider only the total mass of the mesoscopic part. Let $\varepsilon > 0$ and $R \in \mathbb{N}$, we define the (R, ε)-mesoscopic total mass as

$$\overline{\mathrm{Me}}^{(N)}_{R,\varepsilon}(t) = \frac{1}{N} \sum_{i\colon R < M_i^{(N)}(t) < \varepsilon N} M_i^{(N)}(t). \tag{1.11}$$

The mesoscopic total mass in a strict sense arises after taking the limits $N \to \infty$, followed by $\varepsilon \downarrow 0$ and $R \to \infty$, but this does not define a random variable. However, it is possible to calculate an LDP in the $N \to \infty$ limit and then to study the rate function, $\mathcal{J}_{\mathrm{Me}}^{(R,\varepsilon)}$, as $\varepsilon \downarrow 0$ and $R \to \infty$. On the other hand, the proof of Theorem 1.1 shows that it is possible to define a coupled mesoscopic total mass $\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t)$, for any diverging sequence R_N and vanishing sequence ε_N. This is a well-defined random variable and it satisfies an LDP.

Corollary 1.4 (LDP for mesoscopic mass)**.** *Fix $t \in [0, \infty)$.*

(1) Then, for any $R \in \mathbb{N}$ and $\varepsilon \in (0, 1)$, as $N \to \infty$, $\overline{\mathrm{Me}}^{(N)}_{R,\varepsilon}(t)$ satisfies an LDP with rate function $c \mapsto \mathcal{J}_{\mathrm{Me}}^{(R,\varepsilon)}(c; t)$, where

$$\mathcal{J}_{\mathrm{Me}}^{(R,\varepsilon)}(c;t) = \inf\Big\{ I(\lambda, \alpha; t) \colon \sum_{k=1}^{R} k\lambda_k + c + \int_\varepsilon^1 x\,\alpha(\mathrm{d}x) = 1 \Big\}.$$

(2) For any $R_N \in \mathbb{N}$ and $\varepsilon_N \in (0, 1)$ such that $1 \ll R_N < \varepsilon_N N \ll N$, the coupled mesoscopic total mass $\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t)$ satisfies an LDP with rate function

$$\mathcal{J}_{\mathrm{Me}}(c;t) = (1-c)\Big(\log(1-c)t - \frac{(1-c)t}{2}\Big) + \frac{t}{2} - \log t. \tag{1.12}$$

The function $\mathcal{J}_{\mathrm{Me}}(c; t)$ is strictly increasing in c, its minimum over $[0, 1]$ is $\mathcal{J}_{\mathrm{Me}}(0; t) = 0$.

Hence, $\mathcal{J}_{\mathrm{Me}}(\cdot; t)$ can rightfully be called the rate function for the mesoscopic total mass. The probability to have a non-trivial mesoscopic part decays exponentially towards zero. Interestingly, taking $R_N + 1 = \varepsilon_n N \in \mathbb{N}$, we see that already just one mesoscopic particle alone satisfies the same LDP as the entire (R, ε)-mesoscopic total mass in the limit $R \to \infty$, $\varepsilon \downarrow 0$.

Corollary 1.4 part (1) is a simple consequence of the contraction principle, as the maps $\lambda \mapsto \sum_{k=1}^R k\lambda_k$ and $\alpha \mapsto \int_\varepsilon^1 x\,\alpha(\mathrm{d}x)$ are continuous. Assertion (2) follows as a byproduct of our proof of Theorem 1.1 in Section 3.

1.3. **Our results: Phase transitions.** Now we proceed with the main phenomenon in the Marcus–Lushnikov model: the gelation phase transition. We will deduce it from our large-deviations rate functions from Section 1.2. The LDPs and the identification of their strict minimiser(s) lead to laws of large numbers for a number of random quantities.

Consider the following functions of the total masses of the microscopic and macroscopic particles respectively:

$$\mathcal{J}_{\mathrm{Mi}}(c;t) = \inf_{\lambda\in\mathcal{N}(c)} \mathcal{I}_{\mathrm{Mi}}(\lambda;t) \qquad \text{and} \qquad \mathcal{J}_{\mathrm{Ma}}(c;t) = \inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(c)} \mathcal{I}_{\mathrm{Ma}}(\alpha;t),$$

where $c \in [0, 1]$. Clearly, $\mathcal{J}_{\mathrm{Mi}}(c; t) = \mathcal{J}_{\mathrm{Ma}}(1 - c; t)$. These two functions are not entirely analogous to $\mathcal{J}_{\mathrm{Me}}(c; t)$ as rate functions for the total masses of the micro and the macro part, because the total masses both of $\mathrm{Mi}^{(N)}(t)$ and $\mathrm{Ma}^{(N)}(t)$ are equal to one. This is consistent with the fact that the contraction principle cannot be applied to total masses, as they are not continuous functions of the measures. However, they contain rather interesting information about the gelation phase transition.

Theorem 1.5 (Microscopic total mass phase transition)**.** *Fix $t \in [0,\infty)$.*

(1) For any $c \in [0,1]$,

$$\mathcal{J}_{\mathrm{Mi}}(c;t) = tc + (1-c)\log\frac{1-c}{1-\mathrm{e}^{t(c-1)}} + \begin{cases} c\log c - tc^2 & \text{for } c < \frac{1}{t}, \\ -\frac{1}{2t} - \frac{t}{2}c^2 - c\log t & \text{for } c \geq \frac{1}{t}. \end{cases} \tag{1.13}$$

(2) For $c \in (0,1]$, the minimum of $\mathcal{N}(c) \ni \lambda \mapsto \mathcal{I}_{\mathrm{Mi}}(\lambda;t)$ is attained precisely at $\lambda^*(c;t) \in \mathcal{N}(c)$ given by

$$\lambda^*_k(c;t) = \frac{k^{k-2}c^k t^{k-1}\mathrm{e}^{-ctk}}{k!}, \qquad k \in \mathbb{N}, \tag{1.14}$$

and the minimum of the function $c \mapsto \mathcal{J}_{\mathrm{Mi}}(c;t)$ is attained precisely at $c = 1$ with value $\mathcal{J}_{\mathrm{Mi}}(1;t) = 0$. Therefore the infimum

$$\inf_{(\lambda,\alpha)\in\mathcal{N}\times\mathcal{M}_{\mathbb{N}_0}} I(\lambda,\alpha;t) \tag{1.15}$$

is attained at $(\lambda,\alpha) = (\lambda^*(1;t),\mathbf{0})$, where $\mathbf{0} = (0,0,\dots)$.

(3) For $t \in (1,\infty)$, the minimum of the function $c \mapsto \mathcal{J}_{\mathrm{Mi}}(c;t)$ is attained at $c = \beta_t$ where $\beta_t \in (0,t)$ is the smallest positive solution to

$$\log\beta_t = t\beta_t - t. \tag{1.16}$$

The infimum in (1.15) is attained precisely at $(\lambda,\alpha) = (\lambda^*(\beta_t;t),(1-\beta_t,0,0,\dots))$.

The proof is found in Section 4.2.

Theorem 1.5 implies the well-known phase transition at $t = 1$ because the derivative of the minimiser jumps at this point. Combining Theorem 1.5 with the LDP in Theorem 1.1 one has the following law of large numbers:

$$\left(\mathrm{Mi}^{(N)}(t),\mathrm{Ma}^{(N)}(t)\right) \overset{N\to\infty}{\Longrightarrow} \begin{cases} (\lambda^*(1;t),\mathbf{0}) & \text{if } t \leq 1, \\ (\lambda^*(\beta_t;t),(1-\beta_t,0,\dots)) & \text{if } t \geq 1, \end{cases}$$

and can check for $t \leq 1$ that $\lambda^*(1;t)$ is the exact solution of (1.2), the Smoluchowski equation, also given in [Ald99, Table 2]. One also sees that the cut-off versions of the total masses, $\sum_{k=1}^R k\mathrm{Mi}_k^{(N)}(t)$ and $\int_{[\varepsilon,1]} x\,\mathrm{Ma}^{(N)}(t)(\mathrm{d}x)$, converge towards the respective cut-off versions of the limits, and their limits as $R \to \infty$ and $\varepsilon \downarrow 0$ are $(1,0)$ for $t \leq 1$ and $(\beta_t, 1-\beta_t)$ for $t \geq 1$.

1.4. **Literature remarks.** It was expected for a long time that the empirical measure of the masses from the Marcus–Lushnikov process converge in a weak sense to a solution of the Smoluchowski ODEs. The first rigorous convergence result of this kind is due to Lang and Nguyen [LN80], but Lushnikov [Lus78] provided a more informal justification. In the case of a multiplicative kernel, $K(m,\widetilde{m}) = cm\widetilde{m}$ for all $m,\widetilde{m}$ and a proportionality factor c, Smoluchowski's ODEs exhibit mass loss after a critical time (which is equal to $1/c$); a feature that cannot be reproduced by any Marcus–Lushnikov process for finite N. Lushnikov however realised that at large enough times $M_1^{(N)}(t) \asymp N$, that is: a macroscopic particle or gel forms and the Smoluchowski ODEs are not able to describe it. On the other hand the Smouchowski ODEs can be augmented by an equation for the size of the gel, this takes into account the gelation phase transition and convergence has been later proved for the empirical measure plus the rescaled gel mass by Norris [Nor00].

In [HSES85], it is noted that the representation of the solutions of the master equation[1] from Lushnikov [Lus78] can be written in product form as

$$\mathbb{P}\left(M_1^{(N)}(t)=m_1, M_2^{(N)}(t)=m_2,\dots\right)=\frac{1}{Z_N}\prod_i \varphi_N(m_i), \qquad (m_i)_i\in S_N, \tag{1.17}$$

for some $Z_N>0$ and positive function φ_N. Actually, this is true for any coagulation kernel that can be written as $K(m,\widetilde{m})=mf(\widetilde{m})+f(m)\widetilde{m}$ for any $m,\widetilde{m}$, for some positive function f; see [BP90]. Necessary conditions for (1.17) to hold are given by Granovsky and Kryvoshaev [GK12].

Building on the product structure in (1.17), Buffet and Pulé [BP90, BP91] make precise and rigorous the insight of Lushnikov regarding the formation of a gel in the limit $N\to\infty$. Their main result is the existence of the gelation phase transition, for the kernel $K_N(m,\widetilde{m})=m\widetilde{m}/N$, somewhere in the time interval $[\log 2,1]$, by exclusively looking at the macroscopic particles and deriving estimate for the expected values of their sizes. In arriving at these estimates they use bounds on the solution of the master equation derived from the recursive representation going back to Lushnikov [Lus78], which are equivalent to our limits derived via random graph arguments in (2.5) below. They do not prove large-deviations upper bounds for the state of the Marcus–Lushnikov process. In a later work, providing more precise characterisation of the gel, but still at a formal level, Lushnikov [Lus04] picks up on this idea talking about a particular quantity playing the role of a “free energy”. We also exploit this product structure in the present work along with connections to random graph theory.

1.5. **Large-deviations for Erdős-Rényi random graphs.** As noted in the introduction, the distribution at time t of the Markus–Lushnikov process with multiplicative kernel is closely related to that of the connected component sizes for the Erdős-Rényi random graph $\mathcal{G}(N,1-\mathrm{e}^{-t/N})$. This correspondence was not mentioned in [BP90], but was discussed one year later in [BP91], which highlights the connection between gelation in the coagulation process and the phase transition given by the formation of a giant connected component in the Erdős-Rényi random graph [ER60].

Our analysis and results are both closely connected to the study of Erdős-Rényi random graph $\mathcal{G}(N,1-\mathrm{e}^{-t/N})$, more precisely with the large-deviation properties of the sizes of all the components as $N\to\infty$. The literature does not contain many results in this respect for sparse graphs, that is $\mathcal{G}(N,p)$ with $p\sim 1/N$. An LDP for the size of the largest component has been found [O'C98], some results dealing with the macroscopic components [Puh05] and the degree distribution are available [BC15]. Under assumptions that imply at a minimum $p\gg N^{-\frac{1}{2}}$ recent progress has been made on the upper tails of sub-graph counts [CD16, BCCL18, Aug18, CD18]. In the case of dense graphs (fixed $p\in(0,1)$) there is a complete treatment thanks to Chatterjee and Varadhan [CV11], see [Cha16] for an overview.

1.6. **Pathwise large-deviations.** In this work we focus on large-deviations for the coagulation process at a single time t starting with an initial state composed entirely of monomers, that is, with no prior coagulation. One could seek to derive an LDP for the paths of the coagulation process on a compact time interval. For this, one would have to extend the combinatorial and asymptotical work of the present paper to general starting configurations, and use the Markov property to prove LDPs for the finite dimensional distributions. One could then use a projective limit argument augmented by a path space exponential tightness result (see for example [FK06, Chapter 4]). We consider this programme doable, but cumbersome, and therefore decided to defer it to future work.

Such an LDP would be in the spirit of the well-known Wentsel–Freidlin theory, and there are already a number of results of this type in the literature. For the coalescent process an LDP has been derived formally [MPPR17, Thm 3.4] following the non-linear semi-group approach of [FK06]. This type of

[1] The master equation is the Kolmogorov forward equation, that is, the ODE for the time marginals $\nu_t^{(N)}$ of the law of $M^{(N)}$. It is given by $\frac{\mathrm{d}}{\mathrm{d}t}\nu_t^{(N)}=\mathcal{L}_N^\dagger\nu_t^{(N)}$, where $\mathcal{L}_N^\dagger$ is the adjoint of $\mathcal{L}_N$.

results yield a formula for the rate function that is much less explicit and rather different from those that an extension of the present work would yield. On the other hand the less explicit approach is applicable to a wide range of Markov processes.

As in the classic Wentsel–Freidlin results, the formal path space calculations from [MPPR17] yield a rate function for the entire path of the microscopic part of the system $[0,T] \ni t \mapsto \lambda(t)$ that can be written as

$$\mathcal{I}_{\rm Mi}(\lambda) = \int_0^T L\big(\lambda(t), \dot\lambda(t)\big)\,\mathrm{d}t,$$

where L is the solution of variational problem. In order to derive, via the contraction principle, a formula for the rate function for $\lambda(t)$ at a fixed time t, one observes that all paths where the rate function is finite are continuous (in fact a.e. differentiable), and obtains, for configurations λ,

$$\mathcal{I}_{\rm Mi}(\lambda;t) = \inf_{c\colon c(t)=\lambda} \int_0^T L\big(c(s), \dot c(s)\big)\,\mathrm{d}s.$$

Solving this multilayered optimisation problem seems to require major work, and identifying the right-hand side with our formula in (1.9) remains an intriguing problem.

1.7. **Comparison to Bose-Einstein condensation without interaction.** Our large-deviations approach to the Marcus–Lushnikov models shows remarkable similarities to another well-known phase transition in a non-spatial model, the non-interacting Bose gas. Here the situation is similar in that the gas can be conceived as a joint distribution of N particles that are randomly grouped into smaller units, called *cycles*, which can become arbitrarily large. The natural question is then, under what circumstances do macroscopic cycles arise. An explicit answer in terms of a large-deviations analysis has been given in [Ada08], where the transition, the famous *Bose-Einstein Condensation (BEC)* in dimensions $d \geq 3$, is derived from the minimization of the rate function, in a way analogous to that in our Theorem 1.5. The two phase transitions differ in that the BEC transition is of *saturation type*, while the gelation transition is not.

For the non-interacting Bose gas in the thermodynamic limit at temperature $1/\beta \in (0,\infty)$ with particle density $\rho \in (0,\infty)$ the partition function is given by

$$Z_{\Lambda_N}^{(\beta)} = \sum_{(\ell_k)_{k\in\mathbb{N}}\in\mathbb{N}_0^{\mathbb{N}}\colon \sum_k k\ell_k=N} \prod_k \frac{N^{\ell_k}}{\ell_k!\,k^{\ell_k}} [\rho(4\pi\beta k)^{\frac d2}]^{-\ell_k},$$

where Λ_N is the centred box in $\mathbb{R}^d$ with volume N/ρ. The free energy per particle is then

$$f(\beta,\rho) = \lim_{N\to\infty}\frac1N \log Z_{\Lambda_N}^{(\beta)} = -\inf_{\lambda\in\mathcal{N}(\rho)} I(\lambda), \qquad \text{where} \qquad I(\lambda) = \sum_k \lambda_k \log\frac{\lambda_k k}{(4\pi\beta k)^{\frac d2}\,\mathrm{e}}.$$

For the Marcus–Lushnikov model the equivalent quantity is the rate function $\mathcal{I}_{\rm Mi}$ from (1.9). The key difference between the rate functions is that only $\mathcal{I}_{\rm Mi}$ contains terms in the total mass of microscopic particles, c_λ. This reflects the fact that the giant particle makes a significant contribution to the rate function in the Marcus–Lushnikov model, but the condensate in the non-interacting Bose gas does not.

The respective minimisers of $I_{\rm Mi}$ and I are

$$k\lambda_k^{(\rm ML)}(c;t) = \frac1t\,\frac{(ct\mathrm{e}^{-ct})^k}{k^{1-k}\,k!} \sim \frac{1}{\sqrt{2\pi}t}\frac{\big(ct\mathrm{e}^{-ct+1}\big)^k}{k^{3/2}} \quad \text{and} \quad k\lambda_k^{(\rm BEC)}(\alpha;\beta) = \frac{1}{\rho(4\pi\beta)^{\frac d2}}\frac{\mathrm{e}^{-\alpha k}}{k^{\frac d2}},$$

where c and α control the values of $\sum_k k\lambda_k$.

The crucial parameters are the time t for the Marcus–Lushnikov model and the inverse temperature β for the Bose gas. Both models have a trivial upper bound for the total microscopic mass, $\sum_k k\lambda_k$, namely one. One additional upper bound arises in each model from the optimisation of the rate function with respect to the λ_k, but these are not relevant, until t respectively β rises to its critical value. For the Marcus–Lushnikov model this bound is $1/t$, because $\sum_k \frac{(ct\mathrm{e}^{-ct})^k}{k^{1-k}\,k!} \leq 1$ for all $ct \in (0,\infty)$, and the summands take their maxima at $ct = 1$, when they correspond to the Borel probability distribution with parameter 1. For $\lambda^{(\mathrm{BEC})}$ this bound is $\rho^{-1}(4\pi\beta)^{-d/2}\sum_k k^{-\frac{d}{2}}$. At this point we see a difference between the two models, because the total microscopic mass in the Bose gas remains on this bound as β rises further, while for the Marcus–Lushnikov model it immediately drops strictly below the bound. This explains why BEC is known as a saturation phase transition, but this description cannot be applied to gelation.

2. The distribution of the particle sizes

We fix the parameter $t \in (0,\infty)$. In Section 2.1, we derive, for fixed $N \in \mathbb{N}$, an explicit formula for the distribution of the empirical measure of the particle sizes $M_i^{(N)}(t)$ in terms of connectivity probabilities for Erdős-Rényi random graphs. Furthermore, we prepare in Section 2.2 for the asymptotic analysis by giving some estimates and asymptotics for the most crucial object, the probability that a graph is connected.

2.1. **The connection with random graphs.**

Let us explain the connection between the Marcus–Lushnikov coagulation model with multiplicative kernel and a time-dependent version of the well-known Erdős-Rényi graph, see [Ald99]. This will be our starting point for the identification of the joint distribution of $\mathrm{Mi}^{(N)}(t)$ and $\mathrm{Ma}^{(N)}(t)$.

Equip each unordered pair $\{i,j\}$ of distinct numbers in $\{1,\dots,N\}$ with an exponentially distributed random time $\mathrm{e}_{i,j}$ with parameter N, i.e., with expected value $1/N$. All these $N(N-1)/2$ random times are assumed to be independent. At time $\mathrm{e}_{i,j}$ a bond is created between i and j so the probability of a bond between i and j forming by time t is $1-\mathrm{e}^{-t/N}$. By independence, an Erdős-Rényi graph $\mathcal{G}(N,p)$ with parameter $p_N(t) = 1-\mathrm{e}^{-t/N}$ arises. If now $m_i^{(N)}(t)$ denotes the size of the i-th largest connected component (cluster) of this graph, then we have that $(m_i^{(N)}(t))_i$ and $(M_i^{(N)}(t))_i$ are identical in distribution. (This equality is even true process-wise in t, but we are here not interested in that.) In this way, we can see the coagulation process as a function of $\mathcal{G}(N,p_N(t))$. The fact that this description of the distribution is correct comes from the two characteristic properties of the exponential distribution: (1) it has no memory, and (2) the minimum of two independent exponential times is an exponential random time with parameter equal to the sum of the two parameters. Two particles in the coagulation process of cardinalities m and $\widetilde{m}$ at a given time have precisely $m\,\widetilde{m}$ independent exponential times with parameter N that have not yet elapsed; any elapsure of any of them would connect the two particles. Altogether, this means that these two particles coagulate with rate $m\widetilde{m}/N = K_N(m,\widetilde{m})$.

Hence, an important quantity is

$$\mu_t^{(N)}(k) = \mathtt{P}\big(\mathcal{G}(k, 1-\mathrm{e}^{-\frac{t}{N}}) \text{ is connected}\big), \tag{2.1}$$

where we wrote $\mathtt{P}$ for the probability on the graphs.

We denote by $\mathcal{P}_N$ the set of all partitions of $\{1,\dots,N\}$. We write $B_i(\pi)$ for the number of sets in $\pi \in \mathcal{P}_N$ with cardinality i. Then we can describe the distribution of the coagulation process at time t as follows.

Lemma 2.1. *For any $N \in \mathbb{N}$ and every $(m_i)_i \in S_N$,*

$$\mathbb{P}_N\big((M_i^{(N)}(t))_i = (m_i)_i\big) = \#\{\pi \in \mathcal{P}_N \colon B_i(\pi) = m_i\,\forall i\} \times \Big(\prod_i \mu_t^{(N)}(m_i)\Big) \times \Big(\prod_{i\neq j} \mathrm{e}^{-\frac{t}{2N}m_i\,m_j}\Big). \tag{2.2}$$

Proof. A set $A \subset \{1, \dots, N\}$ of indices is a connected component in the graph $\mathcal{G}(N, p_N(t))$ if and only if (1) no bond between any index in A and any index outside has been connected by time t, and (2) the subgraph formed out of the vertices in A and all the bonds between any two vertices in A that have been created by time t is connected. This is the case precisely if any only if $\mathrm{e}_{i,j} > t$ for all $i \in A$ and $j \in A^{\mathrm{c}} = \{1, \dots, N\} \setminus A$, and A is connected. This has probability $\mathrm{e}^{-|A|\,|A^{\mathrm{c}}|t/N} \times \mu_t^{(N)}(|A|)$. Applying this reasoning to A^{c} and describing the next cluster, and iterating this argument, shows that the product of the two products on the right-hand side of (2.2) is equal to the probability, for a given partition π with m_i sets of size i for any i, that the clusters of $\mathcal{G}(N, p_N(t))$ are precisely the sets of π. Since this probability depends only on the cardinalities, the counting term completes the formula. □

Now we rewrite the right-hand side of (2.2) in terms of the empirical measure of $(m_i)_i$, i.e., of the numbers ℓ_k of indices i such that $m_i = k$. Introduce the event

$$A_{N,t}(\ell) = \bigcap_{k \in \mathbb{N}} \{\#\{i \colon M_i^{(N)}(t) = k\} = \ell_k\}, \qquad \ell = (\ell_k)_{k \in \mathbb{N}} \in \mathbb{N}_0^{\mathbb{N}}, \tag{2.3}$$

Corollary 2.2. *For any N and any $\ell = (\ell_k)_k \in \mathbb{N}_0^{\mathbb{N}}$ satisfying $\sum_k k\ell_k = N$,*

$$\mathbb{P}_N(A_{N,t}(\ell)) = N! \prod_k \frac{\mu_t^{(N)}(k)^{\ell_k} \mathrm{e}^{-\frac{t}{2N} k(N-k)\ell_k}}{k!^{\ell_k}\, \ell_k!}. \tag{2.4}$$

Proof. Note that the last product on the right-hand side of (2.2) can also be written as $\prod_i \mathrm{e}^{-\frac{t}{2} m_i(N-m_i)}$. Hence, if ℓ_k is equal to the number of i such that $m_i = k$ for any k, then the product of the last two product can be written as

$$\prod_k \Big(\mu_t^{(N)}(k)^{\ell_k} \mathrm{e}^{-\frac{t}{2N} k(N-k)\ell_k} \Big).$$

The counting term is easily identified as

$$\#\{\pi \in \mathcal{P}_N \colon \#\{A \in \pi \colon |A| = k\} = \ell_k \,\forall k\} = \frac{N!}{\prod_k k!^{\ell_k}\, \ell_k!}.$$

Substituting ends the proof. □

(To avoid confusion, we note that there is a typographical error in Section 4.5 of [Ald99], where the formula (2.4) appears with $\mathrm{e}^{-\frac{t}{2}}$ replaced by e^{-t}.)

2.2. **The probability of being connected.** Our analysis of (2.4) will depend crucially on an analysis of $\mu_t^{(N)}(k)$. The next two lemmas collect results from [Ste70, Lemma1&2, Theorem 1].

Lemma 2.3 (Bounds and asymptotics for $\mu_t^{(N)}$, [Ste70])**.** *For any $N \in \mathbb{N}$ and any $k \le N$,*

$$\mathrm{e}^{-\frac{t}{2N}(k-1)(k-2)} \le \frac{\mu_t^{(N)}(k)}{k^{k-2}(1 - \mathrm{e}^{-\frac{t}{N}})^{k-1}} \le 1. \tag{2.5}$$

In particular, if $k = o(\sqrt{N})$,

$$\mu_t^{(N)}(k) = k^{k-2} \Big(\frac{t}{N} \Big)^{k-1} (1 + o(1)), \qquad N \to \infty.$$

The expression for the upper bound in (2.5) appears to be present (using somewhat applied chemical language) in [Flo41b, equation (5)]. The following is an alternative upper bound to $\mu_t^{(N)}(k)$, which will be useful in the macroscopic setting, together with an asymptotic result for the connection probability in the so-called sparse case, where the bond probability is proportional to the inverse of the size of the graph.

Lemma 2.4 ([Ste70]). *For all $t > 0$ and $k \in \mathbb{N}$*

$$\mu_t^{(N)}(k) \le \big(1 - \mathrm{e}^{\frac{-kt}{N}}\big)^{k-1}.$$

Fix $\alpha \in (0,1)$. Then, for $N \to \infty$,

$$\mu_t^{(N)}(\lfloor \alpha N \rfloor) = \Big(1 - \frac{\alpha t}{\mathrm{e}^{\alpha t} - 1}\Big)\big(1 - \mathrm{e}^{-t\alpha}\big)^{\alpha N}(1 + o(1)).$$

The assertion remains true if the bond probability $1 - \mathrm{e}^{-t/N}$ is replaced by any sequence $t_N = \frac{t}{N}(1 + o(1))$.

3. Proof of the LDP

In this section we prove the main result of this paper, the large-deviations principle in Theorem 1.1.

Recall the topological remarks on the two state spaces $\mathcal{N}$ and $\mathcal{M}_{\mathbb{N}_0}$ at the beginning of Section 1.2. The metrics d on $\mathcal{N}$ and D on $\mathcal{M}_{\mathbb{N}}((0,1])$, defined by

$$d(\lambda, \widetilde{\lambda}) = \sum_{k=1}^{\infty} 2^{-k} \, |\lambda_k - \widetilde{\lambda}_k| \qquad \text{and} \qquad D(\alpha, \widetilde{\alpha}) = \sum_{i=1}^{\infty} 2^{-i} |\alpha_i - \tilde{\alpha}_i|, \tag{3.1}$$

induce the respective topologies of pointwise and vague convergence. We write $B_\delta(\lambda)$ respectively $B_\rho(\alpha)$ for the δ-ball around λ respectively for the ρ-ball around α. Our main result, the LDP in Theorem 1.1, follows from the following.

Proposition 3.1. *Fix $t \in [0, \infty)$. Then, for any $\lambda \in \mathcal{N}$ and $\alpha \in \mathcal{M}_{\mathbb{N}}((0,1])$,*

$$\lim_{\delta,\rho \downarrow 0} \lim_{N \to \infty} \frac{1}{N} \log \mathbb{P}_N\big(\mathrm{Mi}^{(N)}(t) \in B_\delta(\lambda),\, \mathrm{Ma}^{(N)}(t) \in B_\rho(\alpha)\big) = -I(\lambda, \alpha; t). \tag{3.2}$$

Proof. To each element $(m_i)_i$ of the state space S_N defined in (1.3), we associate a unique element of the space

$$\mathcal{N}_N = \Big\{\ell = (\ell_k)_k \in \mathbb{N}_0^{\mathbb{N}} \colon \sum_k k\ell_k = N\Big\}, \tag{3.3}$$

where for each k, ℓ_k is the number of indices i such that $m_i = k$. The map $(m_i)_i \mapsto \ell$ is a bijection and in the following we refer to configurations equally in terms of $(m_i)_i$ or ℓ.

Fix $\delta, \rho > 0$ and $N \in \mathbb{N}$ and recall the definition of $A_{N,t}(\ell)$ in (2.3), then we see that

$$\begin{aligned} &\mathbb{P}_N\big(\mathrm{Mi}^{(N)}(t) \in B_\delta(\lambda),\, \mathrm{Ma}^{(N)}(t) \in B_\rho(\alpha)\big) \\ &\quad = \sum_{\ell \in \mathcal{N}_N} \mathbb{1}\{d(\tfrac{1}{N}\ell, \lambda) < \delta\}\, \mathbb{1}\{D(\ell_{\lfloor \cdot N \rfloor}, \alpha) < \rho\}\, \mathbb{P}_N(A_{N,t}(\ell)). \end{aligned} \tag{3.4}$$

Step 1: Cardinality of $\mathcal{N}_N$: First we note that $|\mathcal{N}_N| = \mathrm{e}^{o(N)}$ because the following argument (which is due to an argument in [Ada08]). For any $\ell \in \mathcal{N}_N$, the set $H(\ell) = \{k \in \mathbb{N} \colon \ell_k > 0\}$ has no more than $2\sqrt{N}$ elements, since

$$N = \sum_{k \in H(\ell)} k\ell_k \ge \sum_{k \in H(\ell)} k \ge \sum_{k=1}^{|H(\ell)|} k = |H(\ell)| \frac{1}{2}(|H(\ell)| - 1).$$

Hence,

$$
\begin{aligned}
|\mathcal{N}_N| &\le \Big|\Big\{(\ell_k)_k \in \mathbb{N}_0^{\mathbb{N}} \colon \sum_k k\ell_k = N, |H(\ell)| \le 2\sqrt{N}\Big\}\Big| \le \sum_{H\subset\mathbb{N}\colon |H|\le 2\sqrt{N}} \Big|\{(\ell_k)_{k\in H} \in \mathbb{N}^H \colon \sum_{k\in H} k\ell_k = N\Big\}\Big| \\
&\le \sum_{H\subset\mathbb{N}\colon |H|\le 2\sqrt{N}} \Big|\{(L_k)_{k\in H} \in \mathbb{N}^H \colon \sum_{k\in H} L_k = N\Big\}\Big| \le \binom{N}{\lfloor 2\sqrt{N}\rfloor}\binom{N+\lfloor 2\sqrt{N}\rfloor}{\lfloor 2\sqrt{N}\rfloor} \\
&= \mathrm{e}^{o(N)}.
\end{aligned}
$$

Hence, we only have to give asymptotic estimates on the single summands on the right-hand side of (3.4).

Step 2: Splitting $\mathbb{P}_N(A_{N,t}(\ell))$. The strategy is to divide the terms in the product representation from Corollary 2.2 into three groups, which we call micro-, meso-, and macroscopic. We fix two increasing sequences R_N and $\varepsilon_N N$ in $\mathbb{N}$ such that $R_N \nearrow \infty$, $\varepsilon_N \downarrow 0$ and $R_N < \varepsilon_N N$. We write

$$
\mathbb{P}_N(A_{N,t}(\ell)) = N! \times F_{\mathrm{Mi}}(\ell) \times F_{\mathrm{Me}}(\ell) \times F_{\mathrm{Ma}}(\ell), \tag{3.5}
$$

where

$$
F_{\mathrm{Mi}}(\ell) = \prod_{k=1}^{R_N} z_k(\ell), \qquad F_{\mathrm{Me}}(\ell) = \prod_{R_N<k\le\varepsilon_N N} z_k(\ell), \qquad F_{\mathrm{Ma}}(\ell) = \prod_{\varepsilon_N N<k\le N} z_k(\ell),
$$

and

$$
z_k(\ell) = \frac{\mu_t^N(k)^{\ell_k}\mathrm{e}^{-\frac{t}{2N}k(N-k)\ell_k}}{k!^{\ell_k}\,\ell_k!}.
$$

Let us set

$$
c_{\mathrm{Mi}}(\ell/N) = \frac{1}{N}\sum_{k=1}^{R_N} k\ell_k, \qquad c_{\mathrm{Me}}(\ell/N) = \frac{1}{N}\sum_{R_N<k\le\varepsilon_N N} k\ell_k, \qquad c_{\mathrm{Ma}}(\ell/N) = \frac{1}{N}\sum_{\varepsilon_N N<k\le N} k\ell_k. \tag{3.6}
$$

Note that the sum of these three terms is equal to one. For the factor $N!$, we use Stirling's formula $N! = (\frac{N}{\mathrm{e}})^N\mathrm{e}^{o(N)}$ so that uniformly in $\ell\in\mathcal{N}_N$

$$
N! = \Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Mi}}(\ell/N)}\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Me}}(\ell/N)}\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Ma}}(\ell/N)}\mathrm{e}^{o(N)}, \qquad N\to\infty. \tag{3.7}
$$

Step 3: Upper bound in the case $c_\lambda + c_\alpha \le 1$. We start by looking at the first term on the right-hand side, i.e., the 'microscopic' term. We use the upper bound in (2.5), to obtain

$$
z_k(\ell) \le \frac{k^{(k-2)\ell_k}t^{(k-1)\ell_k}\mathrm{e}^{-\frac{t}{2N}k(N-k)\ell_k}}{k!^{\ell_k}N^{(k-1)\ell_k}\,(\frac{1}{\mathrm{e}}\ell_k)^{\ell_k}}.
$$

Using this in the first term of (3.5) (together with the first term in (3.7)), we obtain, uniformly for $\ell\in\mathcal{N}_N$, also using that $\sum_{k=1}^{R_N}\frac{t}{2N}k^2\ell_k \le \frac{t}{2}R_N c_{\mathrm{Mi}}(\ell/N)$,

$$
\begin{aligned}
\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Mi}}(\ell/N)}F_{\mathrm{Mi}}(\ell) &\le \prod_{k=1}^{R_N}\Big[\Big(\frac{N}{\mathrm{e}}\Big)^{k\ell_k}\frac{k^{(k-2)\ell_k}t^{(k-1)\ell_k}\mathrm{e}^{\ell_k}\mathrm{e}^{-\frac{t}{2}k\ell_k}\,\mathrm{e}^{\frac{t}{2N}k^2\ell_k}}{k!^{\ell_k}N^{(k-1)\ell_k}(\ell_k)^{\ell_k}}\Big] \\
&= \exp\Big(-N\sum_{k=1}^{R_N}\frac{1}{N}\ell_k\log\frac{k!\mathrm{e}^k\frac{1}{N}\ell_k}{k^{k-2}t^{k-1}\mathrm{e}^{1-\frac{t}{2}k}}\Big)\mathrm{e}^{o(N)} \\
&= \exp\Big(-NI_{\mathrm{Mi}}^{(R_N)}(\tfrac{1}{N}\ell;t)\Big)\,\mathrm{e}^{o(N)},
\end{aligned}
$$

where

$$I_{\rm Mi}^{(R_N)}(\widetilde{\lambda};t)=f^{(R_N)}(\widetilde{\lambda};t)+\sum_{k=1}^{R_N}k\widetilde{\lambda}_k\Big(\frac{t}{2}-\log t\Big)\qquad\text{with } f^{(R_N)}(\widetilde{\lambda};t)\colon=\sum_{k=1}^{R_N}\widetilde{\lambda}_k\log\frac{k!t{\rm e}^{k-1}\widetilde{\lambda}_k}{k^{k-2}},$$

is the cut-off version of the rate function defined in (1.7). Recall that $d(\frac{1}{N}\ell,\lambda)<\delta$ and that $c_\lambda=\sum_{k\in\mathbb{N}}k\lambda_k\in[0,1]$ and observe that $\lim_{R\to\infty}I_{\rm Mi}^{(R)}(\lambda;t)=I_{\rm Mi}(\lambda;t)$. Therefore we see that, for any $R\in\mathbb{N}$,

$$\Big(\frac{N}{\rm e}\Big)^{Nc_{\rm Mi}(\ell/N)}F_{\rm Mi}(\ell)\le\exp\left(-NI_{\rm Mi}(\lambda;t)\right)\,{\rm e}^{N(C_R(\delta)+\gamma_R)+o(N)}\,{\rm e}^{-N(\frac{t}{2}-\log t)(c_{\rm Mi}(\ell/N)-c_\lambda)},\qquad N\to\infty,\tag{3.8}$$

where $\lim_{R\to\infty}\gamma_R=0$ and $\lim_{\delta\downarrow0}C_R(\delta)=0$. Indeed, since $f^{(R)}(\cdot;t)$ is continuous, it is clear that $\sup_{\ell\colon{\rm d}(\frac{1}{N}\ell,\lambda)<\delta}|f^{(R)}(\frac{1}{N}\ell;t)-f^{(R)}(\lambda;t)|$ vanishes as $\delta\downarrow0$ and can therefore be estimated against such a $C_R(\delta)$. Moreover, we estimate (substituting $\frac{1}{N}\ell$ by $\widetilde{\lambda}$), for any N such that $R_N>R$, with the help of the Stirling bound $k!{\rm e}^kk^{-k}\ge1$ and Jensen's inequality for $\varphi(x)=x\log x$, as follows:

$$\begin{aligned}f^{(R_N)}(\widetilde{\lambda};t)-f^{(R)}(\widetilde{\lambda};t)&=\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\log\frac{k!t{\rm e}^{k-1}\widetilde{\lambda}_k}{k^{k-2}}\ge\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\log\frac{k^2t\widetilde{\lambda}_k}{\rm e}\\&\ge\sum_{k=R+1}^{R_N}\frac{\rm e}{tk^2}\varphi\Big(\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\Big/\sum_{k=R+1}^{R_N}\frac{\rm e}{tk^2}\Big)\\&=\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\log\Big(\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\Big/\sum_{k=R+1}^{R_N}\frac{\rm e}{tk^2}\Big)\ge\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\log\Big(cR\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\Big)\\&\ge-\gamma_R,\end{aligned}\tag{3.9}$$

for some $c>0$, where we used that the remainder sum $\sum_{k>R}\frac{1}{k^2}$ is of order $1/R$ as $R\to\infty$ and that $\sum_{k=R+1}^{R_N}\widetilde{\lambda}_k\le1/R$ since $\sum_kk\widetilde{\lambda}_k\le1$ and that the map $x\mapsto x\log(cRx)$ is decreasing in $(0,1/{\rm e}Rc)$, introducing some $-\gamma_R$ that vanishes as $R\to\infty$. In this way, we arrived at the estimate in (3.8). Notice that the last term on the right-hand side cannot be further estimated with the help of continuity (since $\lambda\mapsto c_\lambda$ is not continuous), but will be jointly handled together with the correspondent macroscopic and mesoscopic terms.

For handling the third term in (3.5), we proceed analogously, but use the upper bound in Lemma 2.4, to obtain, for $k\in\{\varepsilon_NN,\dots,N\}$,

$$z_k(\ell)\le\frac{(1-{\rm e}^{-t\frac{k}{N}})^{(k-1)\ell_k}{\rm e}^{-\frac{t}{2N}k(N-k)\ell_k}}{k!^{\ell_k}\,(\frac{1}{\rm e}\ell_k)^{\ell_k}}$$

and consequently

$$\begin{aligned}\Big(\frac{N}{\rm e}\Big)^{Nc_{\rm Ma}(\ell/N)}F_{\rm Ma}(\ell)&\le\prod_{\varepsilon_NN\le k\le N}\Big[\Big(\frac{N}{\rm e}\Big)^{k\ell_k}\frac{(1-{\rm e}^{-t\frac{k}{N}})^{k\ell_k}{\rm e}^{-\frac{t}{2}k\ell_k}{\rm e}^{\frac{t}{2}k^2\ell_k/N}}{k!^{\ell_k}\,\ell_k!}\Big]\\&\le\prod_{\varepsilon_NN\le k\le N}\Big[\Big(\frac{N}{k}\Big)^{k\ell_k}(1-{\rm e}^{-t\frac{k}{N}})^{k\ell_k}{\rm e}^{-\frac{t}{2}k\ell_k}{\rm e}^{\frac{t}{2}k^2\ell_k/N}\Big]\\&=\exp\big(-NI_{\rm Ma}^{(\varepsilon_N)}(\ell_{\lfloor\cdot N\rfloor};t)\big)\end{aligned}\tag{3.10}$$

where

$$I_{\rm Ma}^{(\varepsilon_N)}(\widetilde{\alpha};t)=g^{(\varepsilon_N)}(\widetilde{\alpha};t)+\int_{[\varepsilon_N,1]}x\,\widetilde{\alpha}({\rm d}x)\Big(\frac{t}{2}-\log t\Big)$$

with

$$g^{(\varepsilon_N)}(\widetilde{\alpha};t)=\int_{[\varepsilon_N,1]}\Big[x\log\frac{tx}{1-\mathrm{e}^{-tx}}-\frac{t}{2}x^2\Big]\,\widetilde{\alpha}(\mathrm{d}x),$$

denotes the cut-off version of the rate function I_{Ma} defined in (1.7). Recall that $D(\ell_{\lfloor\cdot N\rfloor},\alpha)<\rho$ and $c_\alpha=\int_{(0,1]}x\,\alpha(\mathrm{d}x)\leq 1$ and observe that $\lim_{\varepsilon\downarrow 0}I^{(\varepsilon)}_{\mathrm{Ma}}(\alpha;t)=I_{\mathrm{Ma}}(\alpha;t)$, then we see that, for any $\varepsilon>0$,

$$\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Ma}}(\ell/N)}F_{\mathrm{Ma}}(\ell)\leq\exp\big(-NI_{\mathrm{Ma}}(\alpha;t)\big)\,\mathrm{e}^{N(C_\varepsilon(\rho)+\gamma_\varepsilon+\frac{t}{2}\varepsilon)+o(N)}\,\mathrm{e}^{-N(\frac{t}{2}-\log t)(c_{\mathrm{Ma}}(\ell/N)-c_\alpha)},$$
$$N\to\infty,\quad(3.11)$$

for some $C_\varepsilon(\rho)$ and γ_ε that satisfy $\lim_{\varepsilon\downarrow 0}\gamma_\varepsilon=0$ and $\lim_{\rho\downarrow 0}C_\varepsilon(\rho)=0$. Indeed, first observe that $g^{(\varepsilon)}(\cdot;t)$ is continuous and hence $|g^{(\varepsilon)}(\ell_{\lceil\cdot N\rceil};t)-g^{(\varepsilon)}(\alpha;t)|$ can be estimated against such a $C_\varepsilon(\rho)$, uniformly in $N\in\mathbb{N}$ and ℓ such that $D(\ell_{\lfloor\cdot N\rfloor},\alpha)<\rho$. Furthermore, for any $\varepsilon>0$ and any $N\in\mathbb{N}$ such that $\varepsilon_N<\varepsilon$,

$$g^{(\varepsilon_N)}(\ell_{\lceil\cdot N\rceil};t)-g^{(\varepsilon)}(\ell_{\lceil\cdot N\rceil};t)=\sum_{k=\varepsilon_N N}^{\varepsilon N}\ell_k\frac{k}{N}\left(\log\frac{\frac{k}{N}t}{1-\mathrm{e}^{-t\frac{k}{N}}}-\frac{t}{2}\frac{k}{N}\right)\geq-\frac{t}{2}\varepsilon,$$

since $\log\frac{x}{1-\mathrm{e}^{-x}}\geq 0$ for all $x>0$. Hence, we arrived at the bound in (3.11). Notice that again we refrain from estimating the term $\mathrm{e}^{-N(\frac{t}{2}-\log t)(c_{\mathrm{Ma}}(\ell/N)-c_\alpha)}$, which needs to be coupled with the microscopic and the mesoscopic part.

Then we are left to handle the middle term in (3.5), for which we use again the upper bound in (2.5) and Stirling's formula, to see that

$$\begin{aligned}\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Me}}(\ell/N)}F_{\mathrm{Me}}(\ell)&\leq\prod_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}\Big[\Big(\frac{N}{\mathrm{e}}\Big)^{k\ell_k}\frac{k^{(k-2)\ell_k}(1-\mathrm{e}^{-t/N})^{(k-1)\ell_k}\mathrm{e}^{-\frac{t}{2N}k(N-k)\ell_k}}{k!^{\ell_k}\,\ell_k!}\Big]\\&\leq\Big(\prod_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}\Big[\frac{N\mathrm{e}}{k^2\ell_k t}\Big]^{\ell_k}\Big)\Big(\prod_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}\mathrm{e}^{\frac{t}{2N}k^2\ell_k}\Big)\big(t\mathrm{e}^{-t/2}\big)^{Nc_{\mathrm{Me}}(\ell/N)}.\end{aligned}$$

We claim that the right-hand side is equal to $(t\mathrm{e}^{-t/2})^{Nc_{\mathrm{Me}}(\ell/N)}\mathrm{e}^{NL_N(\ell)}$ for some $L_N(\ell)$ that vanishes, uniformly in ℓ, as $N\to\infty$. First note that the one-but-last term is such a term, since $\frac{t}{2N}\sum_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}k^2\ell_k\leq\frac{t}{2}\varepsilon_N Nc_{\mathrm{Me}}(\ell/N)$. Furthermore, $\sum_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}\ell_k\leq N/R_N$, which shows that the terms containing t and e in the first product are as small. With the same approach as in (3.9), we see the lower bound

$$\liminf_{N\to\infty}\sum_{k=R_N+1}^{\lfloor\varepsilon_N N\rfloor}\frac{\ell_k}{N}\log\frac{k^2\ell_k}{N}\geq 0.$$

Therefore, uniformly in ℓ such that $D(\ell_{\lceil\cdot N\rceil},\alpha)<\rho$, we have arrived at the estimate

$$\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Me}}(\ell/N)}F_{\mathrm{Me}}(\ell)\leq\big(t\mathrm{e}^{-t/2}\big)^{Nc_{\mathrm{Me}}(\ell/N))+o(N)}=\exp\Big(-N\Big(\frac{t}{2}-\log t\Big)c_{\mathrm{Me}}(\ell/N)\Big)\,\mathrm{e}^{o(N)},\qquad N\to\infty.\quad(3.12)$$

Now we collect (3.8), (3.11) and (3.12) and substitute them in (3.5), also using (3.7), then we obtain, uniformly in ℓ such that $d(\frac{1}{N}\ell,\lambda)<\delta)$ and $D(\ell_{\lfloor\cdot N\rfloor},\alpha)<\rho$, for any $R\in\mathbb{N}$ and any $\varepsilon>0$, as $N\to\infty$,

$$\begin{aligned}\frac{1}{N}\log\mathbb{P}_N(A_{N,t}(\ell))&\leq-I_{\mathrm{Mi}}(\lambda;t)-I_{\mathrm{Ma}}(\alpha;t)+C_R(\delta)+\gamma_R+C_\varepsilon(\rho)+\gamma_\varepsilon+\frac{t}{2}\varepsilon\\&\quad-\Big(\frac{t}{2}-\log t\Big)(1-c_\lambda-c_\alpha)+o(1)\\&=-I(\lambda,\alpha;t)+K_{R,\varepsilon}(\delta,\rho)+o(1),\end{aligned}$$

where $K_{R,\varepsilon}(\delta,\rho)$ vanishes as $\delta \downarrow 0$ and $\rho \downarrow 0$, followed by $R \to \infty$ and $\varepsilon \downarrow 0$, and we recall that $c_{\mathrm{Me}}(\ell/N) = 1 - c_{\mathrm{Mi}}(\ell/N) - c_{\mathrm{Ma}}(\ell/N)$. This implies the upper bound in (3.2) in the case where $c_\lambda + c_\alpha \le 1$.

Step 4: Upper bound in the case $c_\lambda + c_\alpha > 1$. In this case, we implicitly use the lower semicontinuity of the maps $\lambda \mapsto c_\lambda$ and $\alpha \mapsto c_\alpha$ to show that the event $A_{N,t}(\ell)$ is empty for any ℓ such that $d(\frac{1}{N}\ell,\lambda) < \delta$ and $D(\ell_{\lfloor \cdot N\rfloor},\alpha) < \rho$, if δ and ρ are small enough. This will give the right super-exponential upper bound for $\mathbb{P}_N(A_{N,t}(\ell))$, since $I(\lambda,\alpha;t) = \infty$.

Indeed, first pick $R \in \mathbb{N}$ so large and $\varepsilon \in (0,1)$ so small that $\sum_{k=1}^R k\lambda_k + \int_{[\varepsilon,1]} x\,\alpha(\mathrm{d}x)$ are larger than one, say equal to $1+\eta$ for some $\eta > 0$. Then choose δ and ρ in $(0,1)$ so small that, for any ℓ such that $d(\frac{1}{N}\ell,\lambda) < \delta)$ and $D(\ell_{\lfloor \cdot N\rfloor},\alpha) < \rho$, we have

$$\Big|\frac{1}{N}\sum_{k=1}^R k\ell_k + \frac{1}{N}\sum_{\varepsilon N \le k \le N} k\ell_k - \sum_{k=1}^R k\lambda_k + \int_{[\varepsilon,1]} x\,\alpha(\mathrm{d}x)\Big| < \frac{\eta}{2}.$$

For such ℓ, we then have that, using the notation in (3.6),

$$c_{\mathrm{Mi}}(\ell/N) + c_{\mathrm{Ma}}(\ell/N) \ge 1 + \frac{\eta}{2},$$

which contradicts the fact that $c_{\mathrm{Mi}}(\ell/N)$, $c_{\mathrm{Me}}(\ell/N)$ and $c_{\mathrm{Ma}}(\ell/N)$ sum up to one.

Step 5: Logarithmic asymptotics of the lower bound. For the lower bound we only need to consider the case $I(\lambda,\alpha;t) < \infty$, and here we construct a "recovery sequence", that is, a sequence $(\ell^{(N)})_{N\in\mathbb{N}}$ such that

$$\lim_{N\to\infty} d(\tfrac{1}{N}\ell^{(N)},\lambda) = 0, \tag{3.13}$$

$$\lim_{N\to\infty} D(\ell^{(N)}_{\lfloor \cdot N\rfloor},\alpha) = 0, \tag{3.14}$$

$$\liminf_{N\to\infty} \frac{1}{N}\log \mathbb{P}_N(A_{N,t}(\ell^{(N)})) \ge -I(\lambda,\alpha;t). \tag{3.15}$$

For N large enough define $\ell^{(N)} \in \mathcal{N}_N$ by

$$\ell_k^{(N)} = \begin{cases} \lfloor \lambda_k N\rfloor & \text{for } k = 2,\dots,R_N;\\ \lfloor \frac{(1-c_\lambda-c_\alpha)N}{R_N+1}\rfloor & \text{for } k = R_N+1;\\ \alpha\big(\frac{k-1}{N},\frac{k}{N}\big] & \text{for } k = R_N+2,\dots,N;\\ N - \sum_{j\ge 2}^N j\ell_j^{(N)} & \text{for } k = 1,\end{cases} \tag{3.16}$$

where R_N is an arbitrary diverging sequence in $\mathbb{N}$ such that $\log N \ll R_N$. It is clear by construction that (3.13) and (3.14) hold.

Using Lemma 2.3, a calculation similar to that for the upper bound shows that

$$\Big(\frac{N}{\mathrm{e}}\Big)^{Nc_{\mathrm{Mi}}(\ell^{(N)}/N)} F_{\mathrm{Mi}}(\ell^{(N)}) \ge \exp\Big(-NI_{\mathrm{Mi}}^{(R_N)}(\lambda;t)\Big)\,\mathrm{e}^{o(N)},$$

and in the same way, one checks that for $k = R_N+1$

$$\liminf_{N\to\infty}\frac{1}{N}\left[\log z_k\left(\ell_k^{(N)}\right) + k\ell_k^{(N)}\frac{\log N!}{N}\right] \ge (1-c_\lambda-c_\alpha)\left(\log t - \frac{t}{2}\right).$$

Now, fix $\delta \in (0,1)$. One can see that

$$\alpha\big(\tfrac{R_N+2}{N},\delta\big] \le \frac{N}{R_N+2}\int_{\frac{R_N+2}{N}}^{\delta} x\,\alpha(\mathrm{d}x) \le \frac{N}{R_N+2}c_\alpha(\delta), \tag{3.17}$$

where $c_\alpha(\delta)$ vanishes as $\delta \downarrow 0$ whenever $\alpha \in \mathcal{M}_{\mathbb{N}_0}$. Then we use Lemma 2.3 and Stirling's upper bound to see that, for N large

$$\begin{aligned}
&\frac{1}{N}\sum_{k=R_N+2}^{\lfloor \delta N\rfloor}\left[\log z_k\big(\ell^{(N)}\big) + k\ell_k^{(N)}\frac{\log N!}{N}\right] \\
&\quad\geq \big(\log t - \tfrac{t}{2}\big)\sum_{k=R_N+2}^{\delta N}\frac{k}{N}\alpha\big(\tfrac{k}{N},\tfrac{k+1}{N}]\big)! - \frac{1}{N}\sum_{k=R_N+2}^{\delta N}\log\alpha\big((\tfrac{k}{N},\tfrac{k+1}{N}]\big) \\
&\qquad - \sum_{k=R_N+2}^{\delta N}\frac{2\log k + \log\sqrt{2\pi k}}{N}\alpha\big((\tfrac{k}{N},\tfrac{k+1}{N}]\big) + o(1).
\end{aligned}$$

By using (3.17), $\log k \leq k$ and that

$$\prod_{k=R_N+2}^{\delta N}\alpha\big((\tfrac{k}{N},\tfrac{k+1}{N}]\big)! \leq \Big(\sum_{k=R_N+2}^{\delta N}\alpha\big(\tfrac{k}{N},\tfrac{k+1}{N}]\big)\Big)! \leq \Big(\frac{N}{R_N+2}\Big)!$$

we see that, whenever $\log N \ll R_N$, there exist a finite ε_δ, vanishing as $\delta \searrow 0$ such that

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{k=R_N+2}^{\lfloor \delta N\rfloor}\left[\log z_k\big(\ell^{(N)}\big) + k\ell_k^{(N)}\frac{\log N!}{N}\right] \geq \varepsilon_\delta.$$

For the remaining terms one calculates that

$$\begin{aligned}
&\frac{1}{N}\sum_{\lfloor \delta N\rfloor+1}^{N}\left[\log z_k\left(\ell_k^{(N)}\right) + k\ell_k^{(N)}\frac{\log N!}{N}\right] \\
&\qquad = \frac{1}{N}\sum_{\lfloor \delta N\rfloor+1}^{N}k\ell_k^{(N)}\left[\log\left(\frac{\mu_t^{(N)}(k)^{\frac{1}{k}}}{k/N}\right) - \frac{t}{2N}(N-k)\right] + o(1) \\
&\qquad\qquad = \int_{(\delta,1]}x\left[\log\left(\frac{\mu_t^{(N)}\left(\lfloor Nx\rfloor\right)^{\frac{1}{\lfloor Nx\rfloor}}}{x}\right) - \frac{t}{2}(1-x)\right]\alpha(\mathrm{d}x) + o(1). \qquad (3.18)
\end{aligned}$$

Now by Lemma 2.4 the integrand converges pointwise to

$$x\left[\log\left(\frac{1-\mathrm{e}^{-tx}}{x}\right) - \frac{t}{2}(1-x)\right].$$

Since, for N large enough,

$$x\log\left(\frac{x}{\mu_t^{(N)}\left(\lfloor Nx\rfloor\right)^{\frac{1}{\lfloor Nx\rfloor}}}\right) \leq x\left[\log\left(\frac{x}{1-\mathrm{e}^{-tx}}\right) + 1\right],$$

which is clearly integrable over $x \in (\delta, 1]$ with respect to α, we can apply the dominated convergence theorem.

Combining the above estimates we see that for any $\delta \in (0,1)$ we have a_δ with $\lim_{\delta\searrow 0} a_\delta = 0$ and

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}_N(A_{N,t}(\ell^{(N)})) \geq -I(\lambda,\alpha;t) - a_\delta \qquad (3.19)$$

so (3.15) follows on taking the limit $\delta \searrow 0$.

□

4. Corollaries and study of the rate functions

In this section we analyse, for fixed $t \in [0,\infty)$, the minima of the rate function, $I(\lambda,\alpha;t)$, over the configurations λ respectively α, and afterwards the minimima of the rate functions for the total masses, $\mathcal{J}_{\mathrm{Mi}}$, $\mathcal{J}_{\mathrm{Me}}$ and $\mathcal{J}_{\mathrm{Ma}}$. In particular, we find analytical characterisations for the gelation phase transition at $1/t$ if $t > 1$.

4.1. Rate functions for the microscopic part.

We start by minimizing $I(\lambda,\alpha;t)$, for a fixed $\lambda \in \mathcal{N}$ over all compatible $\alpha \in \mathcal{M}_{\mathbb{N}_0}$. We will obtain the rate function for the microscopic part, and we will see that this minimum is attained for α of the form $\alpha = \delta_{c_\alpha}$. Informally speaking, the following in particular implies that, with probability tending to one, there is at most one macroscopic particle.

Lemma 4.1 (Analysis of the microscopic rate function)**.** *Fix $\lambda \in \mathcal{N}$ and recall that $c_\lambda = \sum_{k\in\mathbb{N}} k\lambda_k \in [0,1]$, then*

$$\inf_{\alpha\in\mathcal{M}_{\mathbb{N}}} I(\lambda,\alpha;t) = \sum_{k=1}^{\infty} \lambda_k \log \frac{k!\mathrm{e}^{k-1}t\lambda_k}{k^{k-2}} - (1-c_\lambda)\left(\log\frac{1-\mathrm{e}^{(c_\lambda-1)t}}{(1-c_\lambda)} - \frac{c_\lambda t}{2}\right) + c_\lambda\Big(\frac{t}{2} - \log t\Big).$$

Moreover, the minimum is attained precisely at $\alpha = \delta_{1-c_\lambda}$.

Proof. Clearly

$$\begin{aligned}\inf_{\alpha\in\mathcal{M}_{\mathbb{N}}} I(\lambda,\alpha;t) &= \inf_{c\in[0,1-c_\lambda]}\ \inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(c)} I(\lambda,\alpha;t)\\ &= I_{\mathrm{Mi}}(\lambda;t) + \inf_{c\in[0,1-c_\lambda]}\Big(\inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(c)} I_{\mathrm{Ma}}(\alpha;t) + (1-c_\lambda-c)\Big(\frac{t}{2}-\log t\Big)\Big).\end{aligned}$$

Fix $c \in [0,1]$ and $\alpha \in \mathcal{M}_{\mathbb{N}_0}(c)$. Note that $\alpha((c,1]) = 0$ since α is a point measure with $\int_{(0,1]} x\,\alpha(\mathrm{d}x) = c$. We have, denoting $f_t(x) = \log\frac{x}{1-\mathrm{e}^{-tx}} + \frac{t}{2}(1-x)$,

$$I_{\mathrm{Ma}}(\alpha;t) = \int_{(0,c]} x f_t(x)\,\alpha(\mathrm{d}x) \ge \int x f_t(c)\,\alpha(\mathrm{d}x) = c f_t(c) = c\log\frac{c}{1-\mathrm{e}^{-tc}} + \frac{t}{2}c(1-c) = I_{\mathrm{Ma}}(\delta_c;t), \tag{4.1}$$

since f_t is strictly decreasing in $[0,\infty)$. Indeed,

$$f_t'(x) = \frac{1}{x} - \frac{t\mathrm{e}^{-tx}}{1-\mathrm{e}^{-tx}} - \frac{t}{2} = \frac{t(1+y)}{2y(1-\mathrm{e}^{-2y})}\Big[\frac{1-y}{1+y} - \mathrm{e}^{-2y}\Big], \qquad y = \frac{tx}{2}.$$

We want to prove that $f'(x) < 0$ for $x \in [0,\infty)$. For $y \ge 1$, this is obvious from above, and for $y \in [0,1)$, this is easily seen as follows.

$$\mathrm{e}^{2y} = 1 + \sum_{k=1}^{\infty}\frac{(2y)^k}{k!} < 1 + \sum_{k=1}^{\infty} 2y^k = (1+y)\sum_{k=0}^{\infty} y^k = \frac{1+y}{1-y},$$

since $\frac{2^k}{k!} < 2$ for all $k \ge 3$. Hence, we see that $f_t'(x) \le 0$ for $x \in [0,\infty)$, and (4.1) follows.

Furthermore, it is immediate that $c\log\frac{c}{1-\mathrm{e}^{-tc}} + \frac{t}{2}c(1-c) + (1-c_\lambda-c)(\frac{t}{2}-\log t)$ is decreasing in c, and hence the optimal value of c is $c = 1-c_\lambda$. □

Now the proof of Corollary 1.2 directly follows from Theorem 1.1, Lemma 4.1 and the contraction principle since the projection $(\lambda,\alpha) \mapsto \lambda$ is continuous in the product topology.

Let us analyse the minimising statistics of the macroscopic part.

Lemma 4.2 (Analysis of the macroscopic rate function)**.** *Fix $\alpha \in \mathcal{M}_{\mathbb{N}_0}$ and recall that $c_\alpha = \int_{(0,1]} x\,\alpha(\mathrm{d}x) \in [0,1]$, then*

$$\inf_{\lambda\in\mathcal{N}} I(\lambda,\alpha;t) = I_{\mathrm{Ma}}(\alpha;t) + C_{\alpha,t}\Big(\log(tC_{\alpha,t}) - \frac{t}{2}C_{\alpha,t}\Big) + (1-c_\alpha)\Big(\frac{t}{2} - \log t\Big), \tag{4.2}$$

where $C_{\alpha,t} = (1-c_\alpha)\wedge\frac{1}{t}$. Furthermore, for $1-c_\alpha \le \frac{1}{t}$, the unique minimizer is equal to $\lambda^*(1-c_\alpha;t)$ defined in (1.14)*, and for $1-c_\alpha > \frac{1}{t}$, there is no minimizer, but there are approximating sequences that approach $\lambda^*(\frac{1}{t};t)$.*

Proof. As in the proof of Lemma 4.1, we see that

$$\begin{aligned}\inf_{\lambda\in\mathcal{N}} I(\lambda,\alpha;t) &= \inf_{c\in[0,1-c_\alpha]}\ \inf_{\lambda\in\mathcal{N}(c)} I(\lambda,\alpha;t)\\ &= I_{\mathrm{Ma}}(\alpha;t) + \inf_{c\in[0,1-c_\alpha]}\Big(\inf_{\lambda\in\mathcal{N}(c)} I_{\mathrm{Mi}}(\lambda;t) + (1-c_\alpha-c)\Big(\frac{t}{2}-\log t\Big)\Big)\\ &= I_{\mathrm{Ma}}(\alpha;t) + (1-c_\alpha)\Big(\frac{t}{2}-\log t\Big) + \inf_{c\in[0,1-c_\alpha]}\ \inf_{\lambda\in\mathcal{N}(c)} \widehat{I}(\lambda),\end{aligned} \tag{4.3}$$

where

$$\widehat{I}(\lambda) = \sum_{k=1}^\infty \lambda_k \log\frac{k!t\mathrm{e}^{k-1}\lambda_k}{k^{k-2}}.$$

We always interpret $0\log 0$ as 0. Fix $c\in[0,1]$. Since $\widehat{I}$ is strictly convex on the convex set $\mathcal{N}(c)$, we see by evaluating the variational equations that the only candidate for a minimiser in the interior is

$$\lambda_k^*(c;t) = \frac{\mathrm{e}^{k(\rho-1)}k^{k-2}}{k!t}, \qquad k\in\mathbb{N},$$

with $\rho\in\mathbb{R}$ such that $\sum_{k=1}^\infty k\lambda_k^*(c;t) = c$. Interestingly, we can identify $k\lambda_k^*(c;t) = \mathrm{Bo}_{\mu_\rho}(k)\mu_\rho/t$, where μ_ρ is determined by $\mu_\rho - \log\mu_\rho = 1-\rho$, and

$$\mathrm{Bo}_\mu(k) = \frac{\mathrm{e}^{-\mu k}(\mu k)^{k-1}}{k!}, \qquad k\in\mathbb{N},$$

are the probabilities of the Borel distribution with parameter $\mu\in[0,1]$. Note that $\mathrm{Bo}_\mu(k)$ is not summable for $\mu>1$. Hence, ρ must be picked such that $c = \mu_\rho/t$. The largest value c that can be realised in this way is $c = 1/t$ by picking $\rho = 0$. Hence, the preceding is possible at most for $c\in[0,1\wedge\frac{1}{t}]$. By continuity and strict monotonicity of $\sum_{k=1}^\infty k\lambda_k^*(c;t)$ in ρ, indeed, any $c\in[0,1\wedge\frac{1}{t}]$ can be uniquely realized, by picking $\rho = -tc+\log tc+1 \le 0$ such that $\sum_{k=1}^\infty k\lambda_k^*(c;t) = c$... In this case, it is clear that the minimizer of $\widehat{I}$ in the interior of $\mathcal{N}(c)$ is equal to

$$\lambda_k^*(c;t) = \frac{k^{k-2}c^kt^{k-1}\mathrm{e}^{-ctk}}{k!}, \qquad k\in\mathbb{N},$$

as claimed in (1.14), with value

$$\widehat{I}(\lambda^*(c;t)) = c\Big(\log tc - \frac{tc}{2}\Big). \tag{4.4}$$

Now give an argument why $\lambda^*(c;t)$ realises the minimum of $\widehat{I}$ over $\mathcal{N}(c)$. We show that any such minimiser must be positive in every component. Indeed, if $\lambda_{k^*} = 0$ for some $k^*\in\mathbb{N}$, then we consider $\widehat{\lambda}\in\mathcal{N}(c)$, defined by

$$\widehat{\lambda}_k = \begin{cases}\varepsilon, & \text{if } k = k^*,\\ -\varepsilon C, & \text{if } k = \widehat{k},\\ \lambda_k & \text{otherwise,}\end{cases}$$

with $\widehat{k}\in\mathbb{N}\setminus\{k^*\}$ such that $\lambda_{\widehat{k}} > 0$ and $C>0$ such that $\widehat{\lambda}\in\mathcal{N}(c)$ for any sufficiently small $\varepsilon>0$. Now a simple insertion shows that $\widehat{I}(\widehat{\lambda}) < \widehat{I}(\lambda)$, if $\varepsilon>0$ is small enough, since the slope of $\varepsilon\mapsto\varepsilon\log\varepsilon$

at zero is $-\infty$. Hence, λ cannot be a minimizer. On the other hand, $\lambda^*(c;t)$ has the property that all directional derivatives of $\widehat{I}$ in all admissible directions with compact support are zero; hence it is the minimizer of $\widehat{I}$ over $\mathcal{N}(c)$ for $c \in [0, \frac{1}{t}]$.

When $c > \frac{1}{t}$, it is possible to pick a sequence of $\lambda^{(n)} \in \mathcal{N}(c)$ such that $\lim_{n\to\infty} \widehat{I}(\lambda^{(n)}) = -\frac{1}{2t}$ (pick $\lambda_k^{(n)}$ as $\lambda_k^*(\frac{1}{t};t) + \varepsilon_n \delta_n(k)$ for some suitable $\varepsilon_n > 0$). Furthermore, we now show that

$$\inf_{\lambda\in\mathcal{N}} \widehat{I}(\lambda) \geq -\frac{1}{2t}.$$

Minimising in x for each $k \in \mathbb{N}$ independently shows

$$x \log \frac{k! t \mathrm{e}^{k-1} x}{k^{k-2}} \geq -\frac{1}{t}\frac{1}{k}\frac{k^{k-1}\mathrm{e}^{-k}}{k!} = -\frac{1}{t}\frac{1}{k}\mathrm{Bo}_1(k), \tag{4.5}$$

and so

$$\widehat{I}(\lambda) \geq -\frac{1}{t}\mathbb{E}\left[\frac{1}{X}\right],$$

where X is Borel distributed with parameter 1. Now this expectation [AP98, §4.5] is precisely $\frac{1}{2}$, which is equal to the value of the right-hand side of (4.4) for the critical value $c = \frac{1}{t}$. Hence, the infimum of $\widehat{I}$ over $\lambda \in \mathcal{N}(c)$ for $c \geq \frac{1}{t}$ is equal to $-\frac{1}{2t}$. This shows that the infimum over $\lambda \in \mathcal{N}(c)$ in the last line of (4.3) is equal to $(c \wedge \frac{1}{t})(\log(t(c \wedge \frac{1}{t})) - \frac{t}{2}(c \wedge \frac{1}{t})$, and (4.2) follows. $\square$

Then, the proof of Corollary 1.3 directly follows from Theorem 1.1, Lemma 4.2 and the contraction principle, since the projection is continuous.

Finally, let us draw some conclusions regarding the mesoscopic mass. As stated after Corollary 1.4, it is not possible to apply the contraction principle, if we want to derive an LDP for the sequence of random variables $\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t)$, however we can still identify the rate function by minimizing I over all pairs (λ, α) such that $c_\lambda + c_\alpha = 1 - c$. Even if the contraction principle cannot be applied directly, the following lemma proves that the rate function $\mathcal{J}_{\mathrm{Me}}(c;t)$ has exactly the expected form, given by (4.7).

Lemma 4.3. *Fix $t \in [0, \infty)$. Then, for any $c \in [0, 1]$ and any $R_N \in \mathbb{N}$ and $\varepsilon_N \in (0, 1)$ such that $1 \ll R_N < \varepsilon_N N \ll N$,*

$$\lim_{\delta\downarrow 0}\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\Big(\big|\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t) - c\big| \leq \delta\Big) = -\mathcal{J}_{\mathrm{Me}}(c;t). \tag{4.6}$$

Proof. First we have to prove that, for a fixed $c \in [0, 1]$,

$$\inf_{\substack{\lambda\in\mathcal{N}\\ \alpha\in\mathcal{M}_{\mathbb{N}_0}\\ c_\lambda+c_\alpha=1-c}} I(\lambda,\alpha;t) = \mathcal{J}_{\mathrm{Me}}(c;t)\Big(= (1-c)\Big(\log(1-c)t - \frac{(1-c)t}{2}\Big) + \frac{t}{2} - \log t.\Big) \tag{4.7}$$

Fix $x \in [0, 1-c]$, then for a fixed $\lambda \in \mathcal{N}(x)$

$$\inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(1-c-x)} I(\lambda,\alpha;t) = I_{\mathrm{Mi}}(\lambda;t) + c\Big(\frac{t}{2} - \log t\Big) + \Big[(1-c-x)\log\frac{1-c-x}{1-\mathrm{e}^{-t(1-c-x)}} + \frac{t}{2}(1-c-x)(c+x)\Big],$$

since the infimum is attained in $\alpha = \delta_{1-c-x}$, as proved in Lemma 4.1. Then, with the same procedure of Lemma 4.2, we see that the infimum over $\lambda \in \mathcal{N}(x)$, when $xt \leq 1$ is attained in

$$\lambda^*(x;t) = x\frac{\mathrm{e}^{-xtk}(xt)^{k-1}k^{k-2}}{k!},$$

giving

$$\inf_{\substack{\alpha\in\mathcal{M}_{\mathbb{N}_0}(1-c-x)\\ \lambda\in\mathcal{N}(x)}} I(\lambda,\alpha;t) = x\log(xt\mathrm{e}^{-tx}) - x(1-\frac{xt}{2}) + c\Big(\frac{t}{2}-\log t\Big)$$
$$+\Big[(1-c-x)\log\frac{1-c-x}{1-\mathrm{e}^{-t(1-c-x)}} + \frac{t}{2}(1-c-x)(c+x)\Big],$$

while if $xt>1$, for any $\lambda\in\mathcal{N}(x)$,

$$\inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(1-c-x)} I(\lambda,\alpha;t) \geq -\frac{3}{2t} + c\Big(\frac{t}{2}-\log t\Big) + \Big[(1-c-\frac{1}{t})\log\frac{1-c-\frac{1}{t}}{1-\mathrm{e}^{-t(1-c-\frac{1}{t})}} + \frac{t}{2}(1-c-\frac{1}{t})(c+\frac{1}{t})\Big].$$

Minimizing then for $x\in[0,1-c]$, we see that the infimum is attained in x^* smallest solution to

$$x^* = (1-c)\mathrm{e}^{-t(1-c-x^*)},$$

which is $x^*=1-c$, for all $t\geq\frac{1}{1-c}$ and $x^*<1-c$ otherwise. By substitution, we see that (4.7) holds.

Now, notice that procedure to get the upper bound in the proof of Proposition 3.1 implies in a straightforward way that

$$\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\mathbb{P}_N\Big(\big|\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t)-c\big|\leq\delta\Big)\leq -\mathcal{J}_{\mathrm{Me}}(c;t).$$

In the same way, from the proof of Proposition 3.1, we borrow the strategy of constructing a "recovery sequence", this time using $\lambda^*(x^*;t)$ and $\alpha^*=\delta_{1-c-x^*}$ to construct $\ell^{(N)}$ as in (3.16). This gives

$$\lim_{\delta\downarrow 0}\liminf_{N\to\infty}\mathbb{P}_N\Big(\big|\overline{\mathrm{Me}}^{(N)}_{R_N,\varepsilon_N}(t)-c\big|\leq\delta\Big)\geq -\mathcal{J}_{\mathrm{Me}}(c;t).$$

Notice that in the lower bound part of the proof of Proposition 3.1 we see a restriction on R_N (i.e. $\log N\ll R_N$). However, in this case, we construct the "recovery sequence" in such a way that this is not needed. Indeed, the macroscopic part of the sequence $\ell^{(N)}$ puts all the mass in $k=N(1-c-x^*)$ and the condition on R_N is superfluous. □

The proof of the second point in Corollary 1.4 follows as a direct consequence of Lemma 4.3.

4.2. **Proof of Theorem 1.5.** Item (1) follows by Lemma 4.1 and 4.2. Following the approach of those proofs, one can easily see that

$$\mathcal{J}_{\mathrm{Mi}}(c;t) = \inf_{\lambda\in\mathcal{N}(c)}\mathcal{I}_{\mathrm{Mi}}(\lambda;t) = \inf_{\alpha\in\mathcal{M}_{\mathbb{N}_0}(1-c)}\mathcal{I}_{\mathrm{Ma}}(\alpha;t) = \mathcal{J}_{\mathrm{Ma}}(1-c;t).$$

Let us now prove assertion (2). The form of the minimizing λ follows from Lemma 4.2. Fix $t\in[0,1]$. Then $\mathcal{J}_{\mathrm{Mi}}(c;t) = c\log c - tc^2 + tc + (1-c)\log\frac{1-c}{1-\mathrm{e}^{t(c-1)}}$ is strictly decreasing in $c\in[0,1]$. Indeed

$$\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{J}_{\mathrm{Mi}}(c;t) = \log tc - tc + \frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}} - \log\frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}} = F\Big(\frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}}\Big) - F(tc),$$

where we introduced the function $F(x)=x-\log x$, which is decreasing in $x\in(0,1]$. Hence, monotonicity of $\mathcal{J}_{\mathrm{Mi}}(\cdot;t)$ in $[0,1]$ follows from

$$tc\leq\frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}}\leq 1. \tag{4.8}$$

The first inequality follows by observing that the function $\phi_t(c)=\mathrm{e}^{-t(1-c)}-c$ is nonnegative for all $c\in[0,1\wedge\frac{1}{t}]$, since $\phi_t(0)=\mathrm{e}^{-t}>0$, $\phi_t(1)=0$, and ϕ_t is strictly decreasing in $[0,1]$, since $t\leq 1$. The second inequality follows from the fact that $\psi(z)\colon=1-\mathrm{e}^{-z}-z\mathrm{e}^{-z}\geq 0$ for all $z\in[0,1]$ (substitute $z=t(1-c)$), since $\psi(0)=0$, $\psi(1)=1-2\mathrm{e}^{-1}\geq 0$ and ψ is strictly increasing in $[0,1]$. Therefore, $\mathcal{J}_{\mathrm{Mi}}(\cdot;t)$ is minimized in $c=1$, which implies the conclusion.

Now we turn to assertion (3). For $t \in (1, \infty)$, the derivative of $\mathcal{J}_{\mathrm{Mi}}(c; t)$ writes as follows

$$\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{J}_{\mathrm{Mi}}(c;t) = \frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}} - \log\frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}} + \begin{cases} \log tc - tc & \text{for } c \le \frac{1}{t}, \\ -1 & \text{for } c > \frac{1}{t}. \end{cases}$$

It is clear that $\mathcal{J}_{\mathrm{Mi}}(c;t)$ is strictly increasing in $c \in (\frac{1}{t}, 1]$, while for $c \in [0, \frac{1}{t}]$, we need to go back to (4.8). The right inequality there is still true for any $c < \frac{1}{t}$. Since the quotient in (4.8) is strictly increasing in c and since $F(x) = x - \lg x$ is strictly convex in x, the unique zero of $\frac{\mathrm{d}}{\mathrm{d}c}\mathcal{J}_{\mathrm{Mi}}(c;t)$ is given by the unique solution c of

$$tc = \frac{t(1-c)\mathrm{e}^{-t(1-c)}}{1-\mathrm{e}^{-t(1-c)}},$$

which is precisely the solution $c = \beta_t$ of (1.16). The remaining assertions follow.

Acknowledgements

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project C08.

References

[Ada08] S. Adams. Large deviations for empirical cycle counts of integer partitions and their relation to systems of Bosons. In *Analysis and stochastics of growth processes and interface models*, pages 148–172. Oxford Univ. Press, Oxford, 2008.

[Ald99] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 02 1999.

[AP98] D. Aldous and J. Pitman. Tree-valued Markov chains derived from Galton-Watson processes. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 34(5):637 – 686, 1998.

[Aug18] F. Augeri. Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős-Rényi graphs. Technical Report 1810.01558, arxiv, 2018.

[BC15] C. Bordenave and P. Caputo. Large deviations of empirical neighborhood distribution in sparse random graphs. *Probability Theory and Related Fields*, 163(1-2):149–222, 2015.

[BCCL18] C. Borgs, J. T. Chayes, H. Cohn, and L. M. Lovász. Identifiability for graphexes and the weak kernel metric. Technical Report 1804.03277, arxiv, 2018.

[Bél01] B. Béla. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.

[BP90] E. Buffet and J. V. Pulé. On Lushnikov's model of gelation. *J. Statist. Phys.*, 58(5-6):1041–1058, 1990.

[BP91] E. Buffet and J. V. Pulé. Polymers and random graphs. *J. Statist. Phys.*, 64(1-2):87–110, 1991.

[CD16] S. Chatterjee and A. Dembo. Nonlinear large deviations. *Advances in Mathematics*, 299:396 – 450, 2016.

[CD18] N. A. Cook and A. Dembo. Large deviations of subgraph counts for sparse Erdős-Rényi graphs. *arXiv preprint arXiv:1809.11148*, 2018.

[Cha16] S. Chatterjee. An introduction to large deviations for random graphs. *Bull. Amer. Math. Soc. (N.S.)*, 53(4):617–642, 2016.

[CV11] S. Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European Journal of Combinatorics*, 32(7):1000 – 1017, 2011. Homomorphisms and Limits.

[DZ10] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[ER61] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.

[FK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical surveys and monographs*. American Mathematical Society, Providence, RI, USA, 2006.

[Flo41a] P. J. Flory. Molecular size distribution in three dimensional polymers. I. Gelation. *Journal of the American Chemical Society*, 63:3083–3090, 1941.

[Flo41b] P. J. Flory. Molecular size distribution in three dimensional polymers. III. Tetrafunctional branching units. *Journal of the American Chemical Society*, 63:3096–3100, 1941.

[Gil72] D. T. Gillespie. The stochastic coalescence model for cloud droplet growth. *J. Atmos. Sci.*, 29:1496–1510, 1972.

[GK12] B. L. Granovsky and A. V. Kryvoshaev. Coagulation processes with gibbsian time evolution. *Journal of Applied Probability*, 49(03):612–626, Sep 2012.

[HSES85] E. M. Hendriks, J. L. Spouge, M. Eibl, and M. Schreckenberg. Exact solutions for random coagulation processes. *Zeitschrift für Physik B Condensed Matter*, 58(3):219–227, Sep 1985.

[LN80] R. Lang and X. Nguyen. Smoluchowski's theory of coagulation in colloids holds rigorously in the Boltzmann-Grad-limit. *Z. Wahrsch. Verw. Gebiete*, 54(3):227–280, 1980.

[Lus78] A. A. Lushnikov. Coagulation in finite systems. *Journal of Colloid and Interface Science*, 65(2):276 – 285, 1978.

[Lus04] A. A. Lushnikov. From sol to gel exactly. *Physical review letters*, 93(19):198302, 2004.

[Mar68] A. H. Marcus. Stochastic coalescence. *Technometrics*, 10:133–143, 1968.

[MPPR17] A. Mielke, R. Patterson, M. Peletier, and M. D. Renger. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. *SIAM Journal on Applied Mathematics*, 77(4):1562–1585, 2017.

[Nor00] J. R. Norris. Cluster coagulation. *Communications in Mathematical Physics*, 209(2):407–435, Feb 2000.

[O'C98] N. O'Connell. Some large deviation results for sparse random graphs. *Probability theory and related fields*, 110(3):277–285, 1998.

[Puh05] A. A. Puhalskii. Stochastic processes in random graphs. *The Annals of Probability*, 33(1):337–412, 2005.

[Ste70] V. E. Stepanov. On the probability of connectedness of a random graph $\mathcal{G}_m(t)$. *Theory of Probability & Its Applications*, 15(1):55–67, 1970.

[vS16] M. von Smoluchowski. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Physik. Zeitschr.*, XVII:585–599, 1916. The lectures do not have separate titles; coagulation is treated in lecture III, which starts on page 593.

WIAS, Mohrenstrasse 39, 10117 Berlin

E-mail address: `luisa.andreis@wias-berlin.de`

TU Berlin and WIAS, Mohrenstrasse 39, 10117 Berlin

E-mail address: `wolfgang.koenig@wias-berlin.de`

WIAS, Mohrenstrasse 39, 10117 Berlin

E-mail address: `robert.patterson@wias-berlin.de`