Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

A large-deviations principle for all the components in a sparse inhomogeneous random graph

Luisa Andreis[1], Wolfgang König[2], Heide Langhammer[3], Robert I. A. Patterson[3]

submitted: December 3, 2021

[1] Universitá degli Studi di Firenze
Dipartimento Sistemi e Informatica
Viale Morgagni, 65
50134 Firenze
Italy
E-Mail: luisa.andreis@unifi.it

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
and
Technische Universität Berlin
Institut für Mathematik
Str. des 17. Juni 136
10623 Berlin
Germany
E-Mail: wolfgang.koenig@wias-berlin.de

[3] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: heide.langhammer@wias-berlin.de
robert.i.a.patterson@wias-berlin.de

No. 2898
Berlin 2021

2020 *Mathematics Subject Classification.* 05C30, 05C80, 60F10, 60G57, 60J80.

Key words and phrases. Sparse random graph, empirical measures of components, large deviations, projective limits, giant cluster phase transition, asymptotics for connection probabilities spatial coagulation model, Flory equation, stochastic block model.

This research has been funded by the Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 “Scaling Cascades in Complex Systems”, Project C08, and grant SPP2265 “Random Geometric Systems”, Project P01.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

A large-deviations principle for all the components in a sparse inhomogeneous random graph

Luisa Andreis, Wolfgang König, Heide Langhammer, Robert I. A. Patterson

Abstract

We study an inhomogeneous sparse random graph, $\mathcal{G}_N$, on $[N] = \{1, \ldots, N\}$ as introduced in a seminal paper [BJR07] by Bollobás, Janson and Riordan (2007): vertices have a type (here in a compact metric space $\mathcal{S}$), and edges between different vertices occur randomly and independently over all vertex pairs, with a probability depending on the two vertex types. In the limit $N \to \infty$, we consider the sparse regime, where the average degree is $O(1)$. We prove a large-deviations principle with explicit rate function for the statistics of the collection of all the connected components, registered according to their vertex type sets, and distinguished according to being microscopic (of finite size) or macroscopic (of size $\asymp N$). In doing so, we derive explicit logarithmic asymptotics for the probability that $\mathcal{G}_N$ is connected. We present a full analysis of the rate function including its minimizers. From this analysis we deduce a number of limit laws, conditional and unconditional, which provide comprehensive information about all the microscopic and macroscopic components of $\mathcal{G}_N$. In particular, we recover the criterion for the existence of the phase transition given in [BJR07].

1 Introduction

In this paper, we study the *inhomogeneous random graph* model as introduced in the seminal paper [BJR07], that is, an Erdős–Rényi graph whose vertices have types. We consider the limit of a large number of vertices and concentrate on the *sparse setting*, where each vertex has a number of edges that is of order one. This setting is famous for the emergence of a *giant cluster*. This phase transition was detected and characterized in [BJR07] with the help of a branching process. We consider the case in which the type set is a compact metric space, but our analysis builds on the proof for the type set being any finite set.

In the present paper, we analyze the model from the view point of *large deviations* in a detailed way. We go beyond existing results by (1) considering the joint statistics of all the clusters, both microscopic and macroscopic, (2) registering the types within the clusters (not only their sizes), and (3) giving a joint large-deviations principle (LDP) for all this information. In particular, we recover the limiting quantities and the resulting phase transition in great detail, giving a lot of additional information. Our main results are Theorems 1 and 29 (the LDPs for the type set being a compact metric space and a finite type set, respectively) and Theorems 9 and 7, where we deduce consequences for the phase transition. A building block for our study is Theorem 15, which gives explicit logarithmic asymptotics for the probability of a macroscopic subgraph being connected and which is of independent interest.

The remainder of this section is organized as follows. In Section 1.1 we introduce the inhomogeneous random graph, in Section 1.2 we present our first main result, the large-deviations principle, and in Section 1.3 we give an interpretation of the result. Asymptotic results on the connectivity of graphs are highlighted in Section 1.4.

DOI 10.20347/WIAS.PREPRINT.2898 Berlin 2021

The structure of the rest of the paper is as follows. Our second main result concerns consequences of the large-deviations principle for the limiting behaviour of the model, i.e., conditional and unconditional laws of large numbers. This relies on explicit variational analysis of the rate function of Theorem 1 and it is explained in Section 2, together with the giant-cluster phase transition and a comparison to the results of [BJR07]. Additionally, we explain the deep connection with an important inhomogeneous coagulation process and derive a solution to a spatial version of the Flory equation. In Section 3 we focus on the finite type setting and prove asymptotics for connection probabilities and derive results on multivariate power series (and their link with branching processes). These will play a key role in the remaining sections. The proof of our main result, the LDP, for a finite type set is in Section 4. In Section 5, we derive the LDP for compact metric type spaces via a discrete approximation in the spirit of the Dawson–Gärtner theorem. In Section 6 we derive a full characterization of the minimizers of the microscopic part of the rate function, and in Section 7 we analyze all the other parts of the rate function.

Let us make some few remarks on the literature. In our earlier work [AKP19], we derived an LDP for the sizes of all the cluster sizes of the Erdős–Rényi random graph with one type in the sparse setting. In [AKP19, Section 1.4], we gave a broad survey on known results on such LDPs; summarizing, there are indeed some results on particular statistics of the graph $\mathcal{G}_N$, many of which are *a posteriori* contained in [AKP19, Theorem 1.1] as special cases. However all these results concern only the Erdős–Rényi graph with one type. There is some recent literature on LDPs for various statistics of different types of random graphs outside the sparse regime, which are by their nature so different from the setting of the present paper that we do not go into this here. For inhomogeneous graphs in the sparse regime, there are only a few results in the literature, starting with the seminal paper [BJR07], which introduced the model and investigated the giant-cluster phase transition in detail. Furthermore, clusters of critical sizes of order N^α with some $\alpha \in (0,1)$ around the phase transition have been studied for some types of inhomogeneous random graphs in [BvdHvL10], [BvdHvL12] and [vdH13] under certain moment assumptions on the (scalar) types. Let us finally mention some results just outside the sparse setting: [CCH20] analyzes the eigenvalues of the adjacency matrix of an inhomogeneous Erdős–Rényi random graph with vanishing edge probabilities of order $\gg \frac{1}{N}$ and [DF14] studies the probability of the graph to be connected when the edge probabilities are of order $\frac{\log N}{N}$.

1.1 Inhomogeneous random graphs

We are going to define the random graph model that we study in this paper. It is called an *inhomogeneous random graph* in [BJR07], while at full length it is sometimes named *inhomogeneous random Erdős–Rényi graph*.

Let $N \in \mathbb{N}$; we consider a random graph on the vertex set $[N] = \{1, \dots, N\}$ and fix a non-void set $\mathcal{S}$, the *type set*. We take a type vector $\mathbf{x} = \mathbf{x}^{(N)} = (x_1, \dots, x_N) \in \mathcal{S}^N$ and interpret x_i as the type of i. The edges of this graph are undirected and randomly drawn; self and multiple edges are excluded. The $\binom{N}{2}$ possible edges are sampled independently. The probability to draw an edge between two vertices with types r and s, is called $p_{r,s}$; this defines a map $p\colon \mathcal{S} \times \mathcal{S} \to [0,1]$. The resulting random graph is denoted $\mathcal{G}(N, \mathbf{x}, p)$ and is called the inhomogeneous random graph on $[N]$ with type vector $\mathbf{x}$ and function of probabilities p.

In this paper, we are interested in the limit as $N \to \infty$ in the *sparse case*, i.e., in the case where the number of edges per vertex is of finite order. This is the case if the probabilities $p_{r,s}$ are of order $1/N$.

Actually, we now impose that they are given by

$$p_{r,s} = 1 \wedge \frac{1}{N}\kappa(r,s), \qquad r,s \in \mathcal{S},$$

where $\kappa\colon \mathcal{S}\times\mathcal{S} \to [0,\infty)$ is a symmetric non-negative bounded function, called a *kernel*. Without loss of generality, we are from now on assuming that $\frac{1}{N}\kappa \le 1$, and hence $\kappa(r,s)/N$ is a probability for any $N \in \mathbb{N}$ and any $r,s \in \mathcal{S}$. We will study the graph $\mathcal{G}_N = \mathcal{G}(N,\mathbf{x},\frac{1}{N}\kappa)$ in the limit $N \to \infty$.

We are interested in the structure of all the components[1] of $\mathcal{G}_N$, depending on the types, but not the indices of the vertices. Hence it is sufficient to consider the *empirical measure* $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ of the type vector $\mathbf{x}$ and $N\mu(R)$ is the number of vertices with type in $R \subset \mathcal{S}$. We will assume that, as $N \to \infty$, μ_N converges weakly to a given probability measure μ on $\mathcal{S}$. One can conceive $\mathcal{G}_N$ as a graph on $\mathcal{S}$ where in each point $x \in \mathcal{S}$ there sit precisely $N\mu_N(\{x\}) \in \mathbb{N}$ vertices, which are all distinguished and labelled.

We denote by $\{\mathcal{C}_j\}_j$ the collection of all the vertex sets of the connected components of $\mathcal{G}(N,\mathbf{x},\frac{1}{N}\kappa)$. This collection is a random decomposition of $[N]$. Since we are only interested in the statistics of these components, the labeling of the vertices in a component is irrelevant. Actually, we are even only interested in the statistics of the types of all the vertices, counted with multiplicity. To this end, we introduce the *type-registering empirical-measure function*

$$\eta_{\mathbf{x}}\colon \mathcal{P}([N]) \to \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}), \qquad \eta_{\mathbf{x}}(A) = \sum_{i\in A}\delta_{x_i}, \tag{1.1}$$

where $\mathcal{P}([N])$ is the set of subsets of $[N]$, and $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ is the space of finite measures on $\mathcal{S}$ with values in $\mathbb{N}_0$. We call every element of $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ a *type-configuration* and will denote it by k. In words, $\eta_{\mathbf{x}}(A)(R)$ is, for any set $R \subset \mathcal{S}$, the number of vertices in $A \subset [N]$ with type in R. In particular, $\eta_{\mathbf{x}}([N]) = N\mu_N$. We write $\mathcal{M}(\mathcal{X})$ for the set of finite measures on a set $\mathcal{X}$; the measurable structure on $\mathcal{X}$ will be clear from the context.

We will study the empirical measure of the collection $(\eta_{\mathbf{x}}(\mathcal{C}_j))_j$, i.e., the statistics of how many times a given type-configuration appears as the type-configuration of a component of $\mathcal{G}_N$. We will pay particular attention to the scale of the size of the component, more precisely, whether it is finite or it has a size $\asymp N$. We will call the first scale *microscopic* and the second scale *macroscopic* (macroscopic components are usually called *giant components*). Hence, the quantities of interest in our study are the *microscopic* and the *macroscopic empirical measures* of the type-configurations of the components, which we define as follows:

$$\mathrm{Mi}_N = \mathrm{Mi}_N(\mathbf{x}) = \frac{1}{N}\sum_j \delta_{\eta_{\mathbf{x}}(\mathcal{C}_j)} \qquad \text{and} \qquad \mathrm{Ma}_N = \mathrm{Ma}_N(\mathbf{x}) = \sum_j \delta_{\frac{1}{N}\eta_{\mathbf{x}}(\mathcal{C}_j)}. \tag{1.2}$$

It is clear that both Mi_N and Ma_N only depend on $\mathbf{x}$ through its empirical measure μ_N. Both Mi_N and Ma_N are random measures on $\mathcal{M}(\mathcal{S})$, i.e., they are elements of $\mathcal{M}(\mathcal{M}(\mathcal{S}))$. More precisely, Mi_N is a measure on the set $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ of measures on $\mathcal{S}$ with values in $\mathbb{N}_0$, and Ma_N is a measure on the set $\mathcal{M}(\mathcal{S})$ of measures on $\mathcal{S}$ with values in $\mathbb{N}_0$. Both Mi_N and Ma_N contain precisely the same information for fixed N, but in the limit $N \to \infty$, Mi_N asymptotically registers only the microscopic components and Ma_N only the macroscopic ones. Indeed, the non-microscopic components leave the state space $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ via the set of measures with unbounded total mass, and the non-macroscopic ones (with prefactor $1/N$) leave $\mathcal{M}(\mathcal{S})$ via the measures with vanishing total mass, i.e., via $\{0\}$. The topologies that we will introduce below reflect this effect; it lies at the heart of the phase

[1] We use "cluster" and "component" synonymously.

transition of the emergence of a giant cluster, which we are also interested here. On the other side, they are very natural, as they reduce to the pointwise respectively to the usual vague topology for a finite set $\mathcal{S}$.

The effect of a diverging or vanishing total mass can also be observed in terms of integrated versions of Mi_N respectively Ma_N. Indeed, observe that for any measurable $R \subset \mathcal{S}$ and any $N \in \mathbb{N}$

$$\int k(R)\,\mathrm{Mi}_N(\mathrm{d}k) = \frac{1}{N}\sum_j \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} k(R)\,\delta_{\eta_{\mathbf{x}}(\mathcal{C}_j)}(\mathrm{d}k) = \frac{1}{N}\sum_j \eta_{\mathbf{x}}(\mathcal{C}_j)(R) = \mu_N(R). \tag{1.3}$$

According to our assumption that μ_N weakly converges towards μ, the right-hand side of (1.3) converges to $\mu(R)$, as $N \to \infty$, if R is a μ-continuity set, i.e., if $\mu(\partial R) = 0$. However, the topology on the state space for Mi_N will be chosen as the vague one, and in this topology any accumulation point λ of $(\mathrm{Mi}_N)_{N\in\mathbb{N}}$ satisfies *a priori* only $\int k(R)\,\lambda(\mathrm{d}k) \le \mu(R)$, since the map $k \mapsto k(R)$ is unbounded in general. The same holds for $(\mathrm{Ma}_N)_{N\in\mathbb{N}}$.

To take into account the possibility of a loss of mass we introduce for any $\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}))$ and any $\alpha \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{M}(\mathcal{S}) \setminus \{0\})$ the measures on $\mathcal{S}$

$$c_\lambda(R) = \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} k(R)\,\lambda(\mathrm{d}k), \qquad R \subset \mathcal{S} \text{ measurable}, \tag{1.4}$$

$$c_\alpha(R) = \int_{\mathcal{M}(\mathcal{S})\setminus\{0\}} y(R)\,\alpha(\mathrm{d}y), \qquad R \subset \mathcal{S} \text{ measurable}. \tag{1.5}$$

We call c_λ and c_α the *integrated type-configurations* of λ, respectively of α. If one sees λ as a (not normalized) 'distribution' of finite subsets of $\mathcal{S}$, then c_λ registers the 'expected' total mass of particles in a given subset of $\mathcal{S}$ that appear in this distribution λ; an analogous statement holds for α. The total masses of c_λ and c_α are equal to the integrals of $k \mapsto k(\mathcal{S})$ under λ respectively under α; they are ≤ 1 for any accumulation point of $(\mathrm{Mi}_N)_{N\in\mathbb{N}}$ respectively of $(\mathrm{Ma}_N)_{N\in\mathbb{N}}$, according to the above.

The natural state space containing Mi_N for any $N \in \mathbb{N}$, is the set

$$\mathcal{L} = \big\{\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})) \colon c_\lambda \le \mu \text{ or } c_\lambda \le \mu_N \text{ for some } N \in \mathbb{N},\ \lambda(\{0\}) = 0\big\}. \tag{1.6}$$

The condition $\lambda(\{0\}) = 0$ is clearly satisfied by any empirical measure Mi_N, we could have identified it indeed as an element of $\mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \setminus \{0\})$. However for later notational convenience we do not exclude $\{0\}$ but we add the constraint $\lambda(\{0\}) = 0$. Notice that with this constraint and the conditions on $c(\lambda)$ any $\lambda \in \mathcal{L}$ is a sub-probability measure. Any k in the support of λ is a finite and non-zero point process $k = \sum_i \delta_{z_i}$ with $z_i \in \mathcal{S}$ (with possible repetitions) and stands for the empirical measure of the types of a vertex set of a component, its type-configuration. Informally, the event $\{\mathrm{Mi}_N = \lambda\}$ is the event that, for every $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$, $\mathcal{G}_N$ has $N\lambda(\{k\})$ components with type-configuration k.

The natural state space containing Ma_N for any $N \in \mathbb{N}$, is the set

$$\mathcal{A} = \big\{\alpha \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{M}(\mathcal{S}) \setminus \{0\}) \colon c_\alpha \le \mu \text{ or } c_\alpha \le \mu_N \text{ for some } N \in \mathbb{N}\big\}. \tag{1.7}$$

One can write $\alpha \in \mathcal{A}$ as a finite or at most countable point measure $\alpha = \sum_n \delta_{y_n}$ on measures y_n on $\mathcal{S}$ (with possible repetitions). The total masses $y_n(\mathcal{S})$ of the measures y_n can accumulate only at zero and every $\alpha \in \mathcal{A}$ is concentrated on the set of sub-probability measures. For each y_n we interpret Ny_n as the type-configuration of a giant component. Informally, the event $\{\mathrm{Ma}_N = \alpha\}$ is the event that $\mathcal{G}_N$ has, for any n, a macroscopic component with type-configuration Ny_n.

We will be working only with two particular choices of the type set: $\mathcal{S}$ as a finite set (equipped with the power set as topology and as sigma-field) and $\mathcal{S}$ equal to a compact metric space (equipped with

the topology induced by the metric and the corresponding Borel sigma-field). We equip $\mathcal{M}(\mathcal{S})$ with the weak topology that is generated by all the test integrals against continuous bounded functions $\mathcal{S} \to \mathbb{R}$. However, on the sets $\mathcal{L}$ and $\mathcal{A}$, the appropriate topologies for our purposes are the *vague topologies*, the ones that are induced by all the test integrals against compactly supported continuous test functions $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \to \mathbb{R}$, respectively $\mathcal{M}(\mathcal{S}) \setminus \{0\} \to \mathbb{R}$. If $|\mathcal{S}| < \infty$, then $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ can be identified with $\mathbb{N}_0^{\mathcal{S}}$ and vague convergence in $\mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}))$ is the same as pointwise convergence on $\mathcal{M}(\mathbb{N}_0^{\mathcal{S}})$. On $\mathcal{L} \times \mathcal{A}$ we use the product topology. We will show in Lemma 38 that both $\mathcal{L}$ and $\mathcal{A}$ are compact, hence also $\mathcal{L} \times \mathcal{A}$ is.

The convergence in this topology is the natural one that reflects the possible loss of mass that we are interested in in view of the phase transition. The crucial point is that mass of Mi_N can leak out only via the unboundedness of $k \mapsto k(\mathcal{S})$, i.e., via having larger and larger connected components, while mass of Ma_N can leak out due to the fact that $y \mapsto y(\mathcal{S})$ is not bounded away from zero. With other words, mass of Ma_N leaks out only via the zero measure, where every non-giant component leaves. See Section 5 for details. By the definitions of $\mathcal{L}$ respectively $\mathcal{A}$, the integrated type-configurations c_λ and c_α for $\lambda \in \mathcal{L}$ and $\alpha \in \mathcal{A}$ are sub-probability measures, while for fixed $N \in \mathbb{N}$ both c_{Mi_N} and c_{Ma_N} are even probability measures. The total mass one of c_{Mi_N} can partially go lost and c_{Ma_N} may not lose all its mass in the limit $N \to \infty$. This is precisely the phase transition that we are after.

1.2 The large-deviations principle for the cluster statistics

In this section we formulate the main result of this paper. We assume that $\mathcal{S}$ is a compact metric space.

We need some notation. For any measure ν on $\mathcal{S}$ and any function $\kappa\colon \mathcal{S} \times \mathcal{S} \to [0, \infty)$, we write $\kappa\nu(r) = \int_{\mathcal{S}} \kappa(r, s)\, \nu(\mathrm{d}s)$. The total mass of a measure ν on a measure space $\mathcal{X}$ is denoted by $|\nu| = \nu(\mathcal{X})$. The relative entropy of two (possibly non-normalized) finite measures $\nu, \widetilde{\nu}$ on $\mathcal{X}$ is denoted by

$$\mathbb{H}(\nu|\widetilde{\nu}) = \begin{cases} |\widetilde{\nu}| - |\nu| + \int_{\mathcal{X}} \nu(\mathrm{d}x)\, \log \frac{\mathrm{d}\nu}{\mathrm{d}\widetilde{\nu}}(x), & \text{if } \frac{\mathrm{d}\nu}{\mathrm{d}\widetilde{\nu}} \text{ exists,} \\ +\infty & \text{otherwise.} \end{cases} \tag{1.8}$$

We write $\langle \nu, f \rangle$ for the integral of a function f with respect to a measure ν, and we write $f\,\nu$ for the measure that has the density f with respect to a measure ν.

An important reference measure is the distribution $\mathbb{Q}_\nu$ of a Poisson point process $\mathbb{X}$ on $\mathcal{S}$ with intensity measure $\nu \in \mathcal{M}(\mathcal{S})$, then $\mathbb{Q}_\nu$ is a measure on $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$. Note that we do not assume that ν has a density, hence $\mathbb{X}$ is not necessarily simple.

We define a function τ by

$$\tau(k) := \sum_{T \in \mathcal{T}(k)} \prod_{\{i,j\} \in E(T)} \kappa\left(x_i, x_j\right), \qquad k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}), \tag{1.9}$$

where $(x_1, \ldots, x_{|k|}) \in \mathcal{S}^{|k|}$ is any vector that is *compatible* with k, i.e., $k = \sum_{i=1}^{|k|} \delta_{x_i}$, and $\mathcal{T}(k)$ is the set of spanning trees on $[|k|]$. Notice that τ depends on $(x_1, \ldots, x_{|k|})$ only through k. We use the convention that $\tau(0) = 0$, since the sum is empty. As we will see in Lemma 14, up to a factor of $N^{-|k|+1}$, $\tau(k)$ is equal to the large-N asymptotics of the probability that a random subgraph of $\mathcal{G}(N, \mathbf{x}^{(N)}, \frac{1}{N}\kappa)$ with a given type-configuration k is connected.

We say that $\kappa\colon \mathcal{S} \times \mathcal{S} \to [0, \infty)$ is *irreducible* with respect to a measure $\mu \in \mathcal{M}(\mathcal{S})$ if

$$A \subset \mathcal{S} \text{ and } \kappa = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A) \qquad \Longrightarrow \qquad \mu(A) = 0 \text{ or } \mu(\mathcal{S} \setminus A) = 0. \tag{1.10}$$

Otherwise κ is called reducible.

Here is the main result of this paper.

Theorem 1 (LDP for $(\mathrm{Mi}_N, \mathrm{Ma}_N)$)**.** *Fix a probability measure μ on a compact metric space $\mathcal{S}$ and a kernel κ on $\mathcal{S} \times \mathcal{S}$ that is nonnegative, continuous and irreducible with respect to μ. Assume that $\mathbf{x} = \mathbf{x}^{(N)} \in \mathcal{S}^N$ is such that its empirical measure μ_N converges weakly towards μ as $N \to \infty$. Assume that κ_N is a nonnegative and continuous kernel for any $N \in \mathbb{N}$ such that κ_N converges uniformly towards κ as $N \to \infty$. Let Mi_N and Ma_N be, respectively, the microscopic and macroscopic empirical measure of the connected components of $\mathcal{G}_N = \mathcal{G}(N, \mathbf{x}^{(N)}, \frac{1}{N}\kappa_N)$, for any $N \in \mathbb{N}$, as defined in* (1.2)*.*

Then $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ satisfies a large-deviations principle with speed N and rate function I defined by

$$I(\lambda, \alpha) = \begin{cases} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Ma}}(\alpha) + I_{\mathrm{Me}}(\mu - c_\lambda - c_\alpha), & \text{if } c_\lambda + c_\alpha \le \mu, \\ +\infty & \text{otherwise,} \end{cases}$$

where, for $\lambda \in \mathcal{L}$, $\alpha \in \mathcal{A}$ and $\nu \in \mathcal{M}(\mathcal{S})$,

$$I_{\mathrm{Mi}}(\lambda) = \mathbb{H}(\lambda|\mathbb{Q}_\mu) - 1 - \langle \lambda, \log \tau \rangle + |c_\lambda| - |\lambda| + \frac{1}{2}\langle c_\lambda, \kappa\mu \rangle, \tag{1.11}$$

$$I_{\mathrm{Ma}}(\alpha) = \int_{\mathcal{M}(\mathcal{S})\setminus\{0\}} \alpha(\mathrm{d}y) \left[\left\langle y, \log \frac{\mathrm{d}y}{(1 - \mathrm{e}^{-\kappa y})\,\mathrm{d}\mu} \right\rangle + \frac{1}{2}\langle y, \kappa(\mu - y) \rangle \right], \tag{1.12}$$

$$I_{\mathrm{Me}}(\nu) = \left\langle \nu, \log \frac{\mathrm{d}\nu}{\kappa\nu\,\mathrm{d}\mu} \right\rangle + \frac{1}{2}\langle \nu, \kappa\mu \rangle. \tag{1.13}$$

As in the definition of $\mathbb{H}$ in (1.8)*, we define $I_{\mathrm{Ma}}(\alpha) = \infty$ if it is not true that α-almost everywhere $\frac{\mathrm{d}y}{(1-\mathrm{e}^{-\kappa y})\,\mathrm{d}\mu}$ exists. Likewise we define $I_{\mathrm{Me}}(\nu) = \infty$ if $\frac{\mathrm{d}\nu}{\kappa\nu\,\mathrm{d}\mu}$ does not exist. We use the convention that $\log 0 = -\infty$ and $0 \log 0 = 0$.*

Let us recall the notion of an LDP: Theorem 1 says that $I(\cdot)$ is lower semicontinuous and, for any open set $G \subset \mathcal{L} \times \mathcal{A}$ and any closed set $F \subset \mathcal{L} \times \mathcal{A}$,

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N((\mathrm{Mi}_N, \mathrm{Ma}_N) \in G) \ge -\inf_G I(\cdot), \tag{1.14}$$

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N((\mathrm{Mi}_N, \mathrm{Ma}_N) \in F) \le -\inf_F I(\cdot), \tag{1.15}$$

where we wrote $\mathbb{P}_N$ for the probability measure under the random graph $\mathcal{G}_N$. For the theory of large deviations, see e.g. [DZ10].

An intuitive explanation of Theorem 1 is given in Section 1.3. The proof of Theorem 1 is in Section 5. It relies heavily on the special case of Theorem 1 for finite sets $\mathcal{S}$, see Theorem 29, whose proof we present first in Section 4. Our main strategy there is to identify the joint distribution of all the clusters in $\mathcal{G}(N, \mathbf{x}^{(N)}, \frac{1}{N}\kappa_N)$ in a combinatorial way and then to explicitly extract the exponential rates. The proof of Theorem 1 in Section 5 carries out an approximation procedure of $\mathcal{S}$ with finite state spaces in the spirit of the Dawson–Gärtner theorem.

Theorem 1 is an extension of [AKP19, Theorem 1.1] from the special case $\mu = \delta_0$ and constant κ (that is, from the standard Erdős–Rényi graph) to an inhomogeneous Erdős–Rényi graph. Note that this LDP is also highly non-trivial, interesting, and new in the case of an arbitrary μ and constant κ, to the best of our knowledge.

Remark 2. (Quenched and annealed LDPs) *One possible application of Theorem 1 is to the situation where the vertex types $x_1, \dots, x_N$ are themselves random and independent with distribution μ each. Then Theorem 1 can be seen as a conditional LDP given $\mathbf{x}$, sometimes called a* quenched *LDP. The rate function turns out to be not random and depending only on μ. One can then obtain an* annealed *version of the LDP, i.e., when the probabilities are also taken with respect to the vertices $x_1, \dots, x_N$. The annealing follows from a standard mixture argument when $\mathcal{S}$ is a finite set of points; for general $\mathcal{S}$ the construction of a discretization suitable for use in our proof is a delicate matter that we do not explore here. One possible formulation of the annealed result would be that the triple, consisting of the empirical measures of the vertices, and Mi_N and Ma_N satisfies an LDP with rate function equal to $(\nu, \lambda, \alpha) \mapsto \mathbb{H}(\nu|\mu) + I_\nu(\lambda, \alpha)$, where we now wrote I_ν for the rate function I of Theorem 1 with ν the limiting empirical measure of the type vector (instead of μ).*

Remark 3. (Detailedness) *We decided to register any component of the graph $\mathcal{G}_N$ only through its type-configuration, neglecting all the information about the internal connection structure. It is a natural wish to have also a more detailed analysis, for example by distinguishing each component as a sub*graph *instead of the type-configuration. From such a refined LDP, one could derive Theorem 1 via the contraction principle.*

For the microscopic components it is indeed not too difficult to derive a refined version of the LDP of Theorem 1, since for each type-configuration $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$, the statistics of the $\approx N\lambda(k)$ components with type-configuration k follow an explicit multinomial distribution. The form of the term $\tau(k)$ gives the hint that only spanning trees survive. The macroscopic components are much more involved and it is not clear which type of structure give the decisive contribution in the limit.

Here is a standard corollary from the LDP in Theorem 1 about separate LDPs for Mi_N and Ma_N.

Corollary 4 (Separate LDPs for Mi_N and Ma_N)**.** *Under the assumptions of Theorem 1, the empirical measures $(\mathrm{Mi}_N)_{N\in\mathbb{N}}$ and $(\mathrm{Ma}_N)_{N\in\mathbb{N}}$ each satisfy an LDP on $\mathcal{L}$, respectively on $\mathcal{A}$, with rate functions*

$$\mathcal{I}_{\mathrm{Mi}}(\lambda) = \inf_{\alpha\in\mathcal{A}} I(\lambda, \alpha) \qquad \textit{and} \qquad \mathcal{I}_{\mathrm{Ma}}(\alpha) = \inf_{\lambda\in\mathcal{L}} I(\lambda, \alpha).$$

The LDP assertion directly follows from the contraction principle (see [DZ10]), since both projections $(\lambda, \alpha) \mapsto \lambda$ and $(\lambda, \alpha) \mapsto \alpha$ are continuous. The identification of the two contracted rate functions is formulated in Theorem 9 and discussed in Section 2.1.

Remark 5. (LDP for the mesoscopic part) *Analogously to the corresponding result in [AKP19], we could formulate and prove also a corollary about the* mesoscopic *part of the configuration $(\mathcal{C}_j)_j$ of the components of $\mathcal{G}_N$, i.e., about those components whose cardinalities satisfy $R < |\mathcal{C}_j| < \varepsilon N$ in the limit $N \to \infty$, followed by $R \to \infty$ and $\varepsilon \downarrow 0$. It is clear that we cannot consider the empirical measure of all these components anymore, but only the empirical measure of the total number of vertices of a given type in any of the mesoscopic components. Our conjecture is that (similarly to [AKP19, Corollary 1.4]), this measure on $\mathcal{S}$ satisfies an LDP as $N \to \infty$ for fixed $R \in \mathbb{N}$ and $\varepsilon > 0$ with a rate that converges towards I_{Me} defined in* (1.13) *as $R \to \infty$ and $\varepsilon \downarrow 0$. We abstained from writing out the details.*

1.3 Interpretation of the LDP

Our main result, the LDP of Theorem 1, is highly compressed and contains a number of interesting results as special cases, so let us comment on the impact and draw some conclusions from it. We

will restrict here to the large-deviations issues; the limiting issues and the consequences for the giant-cluster phase transition are deferred to Section 2.

We are examining the probability of the event $\{\mathrm{Mi}_N = \lambda, \mathrm{Ma}_N = \alpha\}$, asymptotically for large N, for any $\lambda \in \mathcal{L}$ and $\alpha = \sum_n \delta_{y_n} \in \mathcal{A}$. Indeed, we want to heuristically argue that one has

$$\mathbb{P}_N(\mathrm{Mi}_N = \lambda, \mathrm{Ma}_N = \alpha) \approx \mathrm{e}^{-NI(\lambda,\alpha)}. \tag{1.16}$$

Recall from Section 1.1 that this is the event that $\mathcal{G}_N$ has $\sim N\lambda(k)$ components with type-configuration k, for any $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$, and a macroscopic component with type-configuration $\sim Ny_n$, for any n. We can clearly restrict to the case that $c_\lambda + c_\alpha \leq \mu$, since otherwise the number of vertices of some type in all the microscopic or macroscopic components together would be larger than the number of existing vertices of that type. However, it might be that the difference $\nu = \mu - c_\lambda - c_\alpha$ is a positive measure; this means that there are $\sim N\nu(R)$ of the vertices with type in R in mesoscopic components for any $R \subset \mathcal{S}$, e.g., in components with N-dependent cardinalities like $\log N$ or $N^{1/3}$, or any mixture.

Recall that the types of all the vertices of $\mathcal{G}_N$ are approximately distributed as $\frac{1}{N}\sum_{i=1}^N \delta_{x_i} \approx \mu$. The probability of $\{\mathrm{Mi}_N = \lambda, \mathrm{Ma}_N = \alpha\}$ consists of a number of terms that are more or less independent, i.e., lead to a sum of exponential rates. These terms are the following:

- a combinatorial term that expresses the number of decompositions of $[N]$ into the collection of subsets as above (respecting all the types),
- the probability that each of these subsets are connected (this depends on the type-configurations k of the microscopic components and on the type-configuration Ny_n of the n-th macroscopic component, respectively),
- the probability that any two of all these vertex sets are not connected.

The above decomposition is the starting point of our combinatorial analysis in Lemma 31, but the large-deviations rate terms that we finally obtained are organized and interpreted in a slightly different fashion as follows.

Let us first consider the microscopic part. The first two terms in $I_{\mathrm{Mi}}(\lambda)$, $\mathbb{H}(\lambda|\mathbb{Q}_\mu) - 1$, describe the number of labellings of $N|c_\lambda|$ vertices into microscopic subsets, according to λ and respecting the type configurations. A priori, it is only notationally convenient to write this as an entropy, but this interpretation allowed us to make the step from discrete to continuous setting. The next three terms, $-\langle\lambda, \log\tau\rangle + |c_\lambda| - |\lambda|$, describe the probability that all the considered subsets are connected, see the comment below (1.9). The last term, $\frac{1}{2}\langle c_\lambda, \kappa\mu\rangle$, collects the costs of isolating each microscopic component from the rest of the system.

For the macroscopic part, given a macroscopic measure $\alpha = \sum_n \delta_{y_n}$, the term $\sum_n \langle y_n, \log\frac{\mathrm{d}y_n}{\mathrm{d}\mu}\rangle$ comes from the number of labellings of $N|c_\alpha|$ vertices into macroscopic subsets, according to α and to the type configurations. The term $-\sum_n \langle y_n, \log(1-\mathrm{e}^{-\kappa y_n})\rangle$ summarizes the connection probabilities of all the macroscopic components, see Theorem 15. Finally, for each n, the term $\langle y_n, \kappa(\mu - y_n)\rangle$ is the cost of isolating the macroscopic component from the rest of the system.

In the mesoscopic part of the rate function, we see only the dependence on the measure ν of all the vertices of mesoscopic components, without any information about the components themselves. Again, the term $\langle \nu, \log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\rangle$ is a result of the relabelling of the $N|\nu|$ vertices according to the type configuration, and the integral of $-\log(\kappa\nu)$ with respect to ν describes in a summarizing way that

each of the vertex sets is connected. The term $\frac{1}{2}\langle \nu, \kappa\mu \rangle$ represents the cost of isolating such vertices from the rest of the graph.

1.4 Connectivity of inhomogeneous graphs

On our way to a proof of Theorem 1, we derive some interesting formulas for quantities that are of general interest in the theory of multi-type Erdős–Rényi graphs. Indeed, in Lemma 31 we give a closed formula for the joint distribution of the entire collection of the vertex sets of all the components of $\mathcal{G}_N$. One important ingredient there is the probability for a given subset of vertices $\subset [N]$ to be connected in $\mathcal{G}_N$. We give sharp estimates for this probability in Lemma 23 for a fixed type-configuration. Another crucial ingredient is a sharp estimate on the logarithm of this probability for a type-configuration of order N, which is given in Theorem 15. In particular, for the connection probability of the entire graph $\mathcal{G}_N$, we obtain the following as a consequence of Theorem 15.

Corollary 6 (Connectivity probability of $\mathcal{G}_N$)**.** *In the situation of Theorem 1 with $\mathcal{S}$ a finite set, then*

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big(\mathcal{G}(N, \mathbf{x}^{(N)}, \tfrac{1}{N}\kappa_N) \text{ is connected}\big) = \sum_{r\in\mathcal{S}} \mu_r \log\big(1 - \mathrm{e}^{-(\kappa\mu)_r}\big)\,. \tag{1.17}$$

Corollary 6 holds true even in the case of a general compact metric space $\mathcal{S}$ as a consequence of Theorem 1.

The question about the connection probability of a random graph has attracted quite some interest. Let us mention [DF14], which studies for the inhomogeneous random graph the regime where the edge probability is of order $\frac{\log N}{N}$ and proves a phase transition: the probability of the graph to be connected either converges to 1 or to 0, depending on the parameters μ and κ (in our notation). In our case, where the edge probability is of order $\frac{1}{N}$, we are in the case in which the probability of the graph being connected is always going to 0. Corollary 6 above identifies the exponential rate of its decay, which we think is of independent interest.

2 Limiting consequences

In this section we present and discuss the second part of our main results, some consequences of the LDP of Theorem 1 that imply detailed and comprehensive limiting assertions about the inhomogeneous random graph. These results rely on involved variational analysis, using recursive formulas and elements of combinatorial power series analysis as methods to explicitly construct minimizers.

Like many large-deviations principles, also Theorem 1 implies a law of large numbers. This is particularly interesting here, since it implies and illustrates the well-known phase transition about the emergence of a giant cluster that was established in [BJR07], and which we record and discuss in Section 2.1. There we also reveal our identification and interpretation of that phase transition in terms of the minimizer(s) of the rate function of our LDP in Theorem 1. In Section 2.2, we compare to the description of this phase transition that was given in [BJR07] in terms of a strongly related multi-type branching process. Furthermore, in Section 2.3 we comment on the irreducibility of κ and explain what the LDP looks like if this assumption is dropped. One of our main motivations for the present work is explained in Section 2.4 where we map the inhomogeneous Erdős–Rényi graph on a particular particle process with a random coagulating mechanism and discuss the consequences of our results for this process, in particular the phase transition of gelation type, i.e., the emergence of a gel, a macroscopic particle.

2.1 The phase transition

We give now comprehensive information about the well-known phase transition of the emergence of a giant component for the inhomogeneous Erdős-Rényi graph using the LDP of Theorem 1. Indeed, we derive a detailed picture of all the limiting microscopic and macroscopic clusters, according to their type-configurations. In this way, we go substantially beyond the work [BJR07], to which we compare in Section 2.2 below. Unlike [BJR07], our point of departure is not a multitype branching process, but the variational analysis of the rate function. This leads us in a natural way to deal with a transformed Poisson point process, which indeed shows deep connections with the multitype branching process.

We introduce first the minimizing microcluster distribution. Let us fix a kernel κ as in Theorem 1. Recall that, for any measure ν on $\mathcal{S}$, we denote by $\mathbb{Q}_\nu$ the distribution of a Poisson point process on $\mathcal{S}$ with intensity measure ν. For $c \in \mathcal{M}(\mathcal{S})$, we introduce the measure λ_c on $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ by

$$\lambda_c(\mathrm{d}k) = \mathrm{e}^{\theta_c(\mathcal{S})}\tau(k)\mathbb{Q}_{\theta_c}(\mathrm{d}k), \qquad \text{where } \theta_c(\mathrm{d}r) = \mathrm{e}^{-\kappa c(r)}\, c(\mathrm{d}r). \tag{2.1}$$

In words, λ_c is obtained by transforming the Poisson point process with intensity measure $\mathrm{e}^{-\kappa c(r)}\, c(\mathrm{d}r)$ with the function τ. It will turn out in the subcritical case (and is implicit in the following theorem) that this measure possesses the integrated type configuration c, that is, the choice of the measure θ_c implies the crucial property that $c_{\lambda_c} = c$, where we recall the notation $c_\lambda(\mathrm{d}r) = \int \lambda(\mathrm{d}k)\, k(\mathrm{d}r)$.

We now introduce an important quantity (which was shown to be crucial in [BJR07]) that we use for separating the sub- and supercritical regimes. For any measure ν on $\mathcal{S}$, we denote by $L^2(\nu)$ the usual L^2-space of functions $\mathcal{S} \to \mathbb{R}$ with respect to the measure ν. We introduce the operator

$$T_{\kappa,\nu}\colon L^2(\nu) \to L^2(\nu), \qquad T_{\kappa,\nu}f(x) = \int_{\mathcal{S}} \kappa(x,y) f(y)\, \nu(\mathrm{d}y), \tag{2.2}$$

and its operator norm

$$\Sigma(\kappa,\nu) = \|T_{\kappa,\nu}\|_{L^2(\nu)} = \sup_{f\in L^2(\nu)\colon \|f\|_{L^2(\nu)}=1} \|T_{\kappa,\nu}f\|_{L^2(\nu)}. \tag{2.3}$$

Informally (an argument will follow in Section 3.1), in the special case that the support of ν is finite (i.e., the case of a finite set $\mathcal{S}$), then $T_{\kappa,\nu}$ can be identified with the matrix $(\kappa(r,s)\nu(s))_{r,s\in\mathcal{S}}$, and $\Sigma(\kappa,\nu)$ is its spectral radius.

Recall the rate function I and the reference measure μ from Theorem 1, now we describe its minimizer.

Theorem 7 (Minimizers of the rate function)**.** *Suppose that κ and μ are as in Theorem 1, then the following hold.*

(i) If $\Sigma(\kappa,\mu) \le 1$, then the unique minimizer of I is equal to $(\lambda_\mu, 0)$.

(ii) If $\Sigma(\kappa,\mu) > 1$, then the unique minimizer of I is equal to $(\lambda_{c^*}, \delta_{\mu-c^*})$, where the subprobability measure c^* is the unique solution to the characteristic equation

$$\mathrm{e}^{-\kappa c(r)}\, c(\mathrm{d}r) = \mathrm{e}^{-\kappa\mu(r)}\, \mu(\mathrm{d}r) \qquad \text{on } \mathcal{S}, \tag{2.4}$$

satisfying $c^* \le \mu$ and $\Sigma(\kappa, c^*) < 1$.

In particular,

$$\big(\mathrm{Mi}_N, \mathrm{Ma}_N\big) \overset{N\to\infty}{\Longrightarrow} \begin{cases} (\lambda_\mu, 0) & \text{if } \Sigma(\kappa,\mu) \le 1, \\ (\lambda_{c^*}, \delta_{\mu-c^*}) & \text{if } \Sigma(\kappa,\mu) > 1. \end{cases} \tag{2.5}$$

The proof of Theorem 7 is in Section 7. Most of it is original research, but take from [BJR07] the discussion of the solutions of the fixed point equation (2.4), see Lemma 16 where we summarize it.

The law of large numbers in (2.5) is a standard consequence of an LDP with a unique minimizer for the rate function. This is a very precise and detailed formulation of the famous giant-cluster phase transition in the graph $\mathcal{G}_N$. Indeed, the following happens with probability tending to one exponentially fast:

(i) In the subcritical phase $\Sigma(\kappa,\mu) < 1$, all vertices (meaning all up to $o(N)$) are in microscopic components, more precisely in the unique optimal configuration encoded by λ_μ. That is, for any $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$, the number of components with vertex set given by k is asymptotically $N\mathrm{e}^{\theta_\mu(\mathcal{S})}\tau(k)\,\mathbb{Q}_{\theta_\mu}(\mathrm{d}k)$, with $\theta_\mu(\mathrm{d}r) = \mathrm{e}^{-\kappa\mu(r)}\,\mu(\mathrm{d}r)$. We have $c_{\lambda_\mu} = \mu$, no giant component appears, and the number of vertices in mesoscopic components is $o(N)$.

(ii) In the case $\Sigma(\kappa,\mu) > 1$, a unique giant cluster appears with $\sim N(1-c^*(\mathcal{S}))$ vertices and type-configuration asymptotically equal to $N(\mu - c^*)$ with c^* characterized by (2.4), since $\theta_\mu = \theta_{c^*}$. The microscopic components are distributed according to the optimal distribution λ_{c^*}, and the number of vertices in mesoscopic components is $o(N)$. Note that this microscopic distribution is not saturated, that is $\Sigma(\kappa,c^*) < 1$, as in the one-type setting [AKP19]. That is, we encounter a phase transition of explosion type, rather than of saturation type, see Remark 8 and [AKP19, Section 1.6].

Remark 8. (Phase transition: saturation versus explosion) *Here is an explanation of the phase transition in terms of a dynamical process. Consider a process of Erdős–Rényi graphs in increasing connection probability, i.e., by adding more and more bonds between the vertices, such that components grow. A suitable growth parameter is $\Sigma(\kappa,c)$, where c stands for the rescaled type-configuration of all the vertices; however, we consider $\Sigma(\kappa,c)$ as a growing function of κ, the bond density. As we explained in [AKP19, Section 1.6], the well-known giant-cluster phase transition (see also the discussion below Theorem 1 below) is an explosion phase transition in the sense that, when crossing the threshold one, a positive fraction of finite-size clusters merges rapidly into one giant cluster and, at any time, every cluster keeps participating in merge events. In particular, the total microscopic mass starts decreasing at the phase transition. In contrast, in condensation phase transitions like the famous Bose–Einstein condensation first all microscopic components reach their maximal size (the saturated state), before a macroscopic component, the condensate, appears, and then the microsocopic ones do not change anymore, but all of the additional mass goes exclusively into the condensate.*

Now we give the description of the two rate functions for the contracted LDPs of $(\mathrm{Mi}_N)_{N\in\mathbb{N}}$ and $(\mathrm{Ma}_N)_{N\in\mathbb{N}}$, respectively, from Corollary 4.

Theorem 9 (Minimizers of the contracted rate function)**.** *Suppose that κ and μ are as in Theorem 1. Then the following hold for $\lambda \in \mathcal{L}$ and for $\alpha \in \mathcal{A}$, respectively.*

$$\mathcal{I}_{\mathrm{Mi}}(\lambda) = \inf_{\alpha\in\mathcal{A}} I(\lambda,\alpha) = \begin{cases} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Ma}}(\delta_{\mu-c_\lambda}) & \text{if } c_\lambda \le \mu, \\ \infty & \text{otherwise,} \end{cases} \tag{2.6}$$

and

$$\mathcal{I}_{\mathrm{Ma}}(\alpha) = \inf_{\lambda\in\mathcal{L}} I(\lambda,\alpha) = \begin{cases} I_{\mathrm{Ma}}(\alpha) + J(\mu - c_\alpha) & \text{if } c_\alpha \le \mu, \\ \infty & \text{otherwise,} \end{cases} \tag{2.7}$$

where for $c \in \mathcal{M}(\mathcal{S})$

$$J(c) = \begin{cases} I_{\mathrm{Mi}}(\lambda_c) = \langle c, \log \frac{\mathrm{d}c}{\mathrm{d}\mu}\rangle + \frac{1}{2}\langle c, \kappa(\mu - c)\rangle & \text{if } \Sigma(\kappa, c) \le 1, \\ I_{\mathrm{Mi}}(\lambda_{b^*}) + I_{\mathrm{Me}}(c - b^*) & \text{if } \Sigma(\kappa, c) > 1, \end{cases} \tag{2.8}$$

where $b^* = b^*(c) \in \mathcal{M}(\mathcal{S})$ *is the minimal non-trivial (i.e., not equal to* c*) solution to*

$$\kappa(c - b^*)(r)\, b^*(\mathrm{d}r) = (c - b^*)(\mathrm{d}r), \quad b^* \le c, \tag{2.9}$$

and it holds that $\Sigma(\kappa, b^*) = 1$*.*

The proof of Theorem 9 is in Section 6 for $\mathcal{I}_{\mathrm{Mi}}$ and in Section 7 for $\mathcal{I}_{\mathrm{Ma}}$. The above theorem suggests us a conditional law of large numbers. Informally, if we fix $\alpha \in \mathcal{A}$ and a sequence of $\alpha_N \in \mathcal{A}$ such that $\alpha_N \to \alpha$. Then, under the probability $\mathbb{P}_N(\cdot|\mathrm{Ma}_N = \alpha_N)$ we have

$$\mathrm{Mi}_N \overset{N\to\infty}{\Longrightarrow} \begin{cases} \lambda_{\mu - c_\alpha} & \text{if } \Sigma(\kappa, \mu - c_\alpha) \le 1, \\ \lambda_{b^*} & \text{if } \Sigma(\kappa, \mu - c_\alpha) > 1. \end{cases} \tag{2.10}$$

Notice that when $\Sigma(\kappa, \mu - c_\alpha) > 1$, then b^* is not equal to $\mu - c_\alpha$, therefore $b^* + c_\alpha \neq \mu$ and the missing mass is interpreted as being mesoscopic.

Remark 10. (Conditional limit with saturation phase transition) *In formula* (2.8) *we encounter a phase transition of saturation type in a conditional limit, in contrast to a transition of explosion type of the unconditional one, see Remark 8. We refer back to Section 1.3 for the interpretation. Recall that* $\mathcal{I}_{\mathrm{Ma}}(\alpha)$ *is the negative exponential rate of the probability of the event* $\{\mathrm{Ma}_N = \alpha\}$*. When* $\Sigma(\kappa, \mu - c_\alpha) \le 1$*,* (2.8) *shows that* $\mathcal{I}_{\mathrm{Ma}}(\alpha) = I_{\mathrm{Ma}}(\alpha) + I_{\mathrm{Mi}}(\lambda_{\mu - c_\alpha})$*, implying that* $\mathbb{P}_N(\mathrm{Ma}_N \approx \alpha) = \mathbb{P}_N(\mathrm{Mi}_N \approx \lambda_{\mu-c_\alpha}, \mathrm{Ma}_N \approx \alpha)\mathrm{e}^{o(N)}$*. The interpretation of this is that, conditionally on the event* $\{\mathrm{Ma}_N = \alpha\}$*, the non-macroscopic mass optimally organizes according to the microscopic measure* $\lambda_{\mu-c_\alpha}$*. In contrast, in the case* $\Sigma(\kappa, \mu - c_\alpha) > 1$*, from* (2.8) *we see that* $\mathcal{I}_{\mathrm{Ma}}(\alpha) = I_{\mathrm{Ma}}(\alpha) + I_{\mathrm{Mi}}(\lambda_{b^*}) + I_{\mathrm{Me}}(\mu - c_\alpha - b^*)$*. This means that, conditionally on the event* $\{\mathrm{Ma}_N = \alpha\}$*, the non-macroscopic mass cannot be organized fully in microscopic clusters, but it is organized microscopically according to* λ_{b^*} *and the remaining vertices, with type-configuration* $N(\mu - c_\alpha - b^*)$*, are put in mesoscopic components. The particular rescaled type-configuration* b^* *is saturated in the sense that* $\Sigma(\kappa, b^*) = 1$*. This means, given a fixed macroscopic type-configuration, if more bonds are thrown into the graph, then first all microscopic clusters grow until they reach the saturated state* λ_{b^*}*, and then this is frozen and only mesoscopic clusters grow. The latter effect is present in literature under the name of frozen percolation, see for example [CRY21, MN14, RT*$^+$*09, Yeo18] and the difference between the two phase transitions is reflected in substantial differences of the hydrodynamic limit, as we summarize in Section 2.4.*

2.2 Comparison to [BJR07]: branching-process interpretation

Our description of the limiting quantities that we presented in Section 2.1 is based on and derived from our analysis of the minimizer of I. Therefore we found it most suitable to present them in terms

of transformed Poisson point processes. However, in the analysis of finite-size components of random graphs, it is common and was often successful to employ well-adapted branching processes for the description. The main idea is that a component can be efficiently (sampled and) analyzed by exploring it via such a branching algorithm. This idea was also a cornerstone in the seminal paper [BJR07], and it produced a description of the limiting macroscopic component in terms of the extinction probability of a crucial branching process. In this section, we recall this description and compare it to our Poisson point process description, also including the microscopic components.

The main tool that is utilized in [BJR07] is a *multitype branching process* with type space $\mathcal{S}$, in which each particle of type $r \in \mathcal{S}$ has offspring with distribution that is a Poisson process with intensity measure $\kappa(r,s)\,\mu(\mathrm{d}s)$. We define $\rho(r) \in [0,1]$ as the probability of non-extinction of the branching process, if it starts with precisely one particle that has type $r \in \mathcal{S}$. We summarize the most important facts from [BJR07] that have relevance for our comparison as follows (see [BJR07, Th. 3.1, Th. 3.12, Th. 6.1, Th. 9.10]).

Theorem 11 (Existence of a giant component, [BJR07])**.** *Suppose the situation of Theorem 1 is given. Abbreviate $\mathcal{G}_N = \mathcal{G}(N, \mathbf{x}, \frac{1}{N}\kappa_N)$, then the following hold.*

(i) $\rho\colon \mathcal{S} \to [0,\infty)$ *is the maximal solution of*

$$\rho = 1 - \mathrm{e}^{-T_{\kappa,\mu}\rho}. \tag{2.11}$$

(ii) If $\Sigma(\kappa,\mu) \le 1$, then the largest component of $\mathcal{G}_N$ has size $O(\log N)$ as $N \to \infty$ with high probability.

(iii) If $\Sigma(\kappa,\mu) > 1$, then the largest component $\mathcal{C}_1$ of $\mathcal{G}_N$ has size $\asymp N$. More precisely, its normalized empirical measure $\frac{1}{N}\eta_{\mathbf{x}}(\mathcal{C}_1)$ (recall (1.1)*) converges weakly towards the measure $\rho(r)\,\mu(\mathrm{d}r)$, and ρ is positive μ-almost everywhere in $\mathcal{S}$.*

Part (iii) identifies the limiting type-configuration of the giant component as N times the measure with density ρ with respect to μ, and part (i) characterizes ρ via the functional identity (2.11). It is easily seen to be equivalent to the characteristic equation (2.4) that we use via the substitution $\rho(r)\,\mu(\mathrm{d}r) = \mu(\mathrm{d}r) - c^*(\mathrm{d}r)$ or $c^*(\mathrm{d}r) = (1-\rho(r))\,\mu(\mathrm{d}r)$. In our analysis of the minimizer of I, (2.4) arose via the Euler–Lagrange equations, while (2.11) emerged in [BJR07, Lemma 5.4] via a standard formula for mixed moments of the offspring of the branching process (which itself uses standard Poisson point process theory).

The statement in part (ii) about the order of the largest component is out of reach of our large-deviations ansatz, which implies that all but $o(N)$ vertices are in components of finite size.

About the distribution of the microscopic clusters of $\mathcal{G}_N$, however, there is no explicit result contained in [BJR07]. However, we can give a description in terms of the above branching process as well. We derive this description now from our form of the minimizer λ_μ defined in (2.1). Let $\Xi(\mathrm{d}r)$ be the total progeny of type r of the branching process; then Ξ is a random measure on $\mathcal{S}$. Then, if $\mathtt{P}_r$ denotes the measure if the process starts from one individual of type r at time 0, we have

$$\mu(\mathrm{d}r)\mathtt{P}_r(\Xi \in \mathrm{d}k) = \lambda_\mu(\mathrm{d}k)\,k(\mathrm{d}r), \qquad k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \setminus \{0\}, r \in \mathcal{S}. \tag{2.12}$$

In Remark 21 we explain this relation in the setting where $\mathcal{S}$ is a finite set. In words, the empirical statistics of the microscopic components in $\mathcal{G}_N$ in the subcritical case approximate the distribution of the total offspring of the characteristic branching process.

2.3 The reducible case

Let us briefly comment on the case where the kernel κ is reducible with respect to μ, i.e., $\mathcal{S}$ is composed of least two irreducible classes (maximal irreducible subsets). Then the graph $\mathcal{G}_N$ decomposes into disconnected subgraphs with types in only one of these classes. Accordingly, the collection of all the connected components can be decomposed into collections for each subgraph. In principle one can apply the LDP of Theorem 1 to each of the micro/macro empirical measures of the subgraphs. However, one might have the wish to have a joint LDP for the entire collection. Here one might expect that the same LDP holds true, and the decomposition into the subgraphs reappears in the rate function in a natural way.

It turns out that this expectation is not disappointed, as it concerns the microscopic part, but is disappointed for the macroscopic part. Actually, the formulation of the LDP slightly changes. The main point is that two macroscopic components can very cheaply be connected to form a significantly larger macroscopic component, just by throwing in one connecting edge, which cannot be seen on the exponential scale. Hence, the macroscopic part is very unstable on the exponential scale under adding edges. This argument fails for the microscopic part, since here we are talking now about $\asymp N$ independent copies of a component of a finite size; if one wants to connect them such that the microscopic statistics change, then one needs to change $\asymp N$ edges, whose probability is clearly seen on an exponential scale.

As a consequence of this effect, we will see that the rate function is finite only if the macroscopic measure $\alpha = \sum_n \delta_{y_n}$ is such that for each n the rescaled type-configuration y_n is supported in one of the irreducible classes. To be precise, we say that a measure $y \in \mathcal{M}(\mathcal{S})$ is *connectable* (with respect to κ and μ) if its support is contained in an irreducible class. Furthermore, a measure $\alpha \in \mathcal{A}$ is called *connectable* if each of its atoms is connectable.

For the microscopic configurations λ connectability is implicitly ensured by the fact that the measure λ has to be absolutely continuous with respect to $\tau\,\mathbb{Q}_\mu$ in order to have a finite value of the rate function; notice that $\tau(k) = 0$ if $\mathrm{supp}(k)$ is not concentrated on a irreducible class of $\mathcal{S}$. We also have to restrain to a sequence of random graphs that are defined with respect to the same kernel κ, rather than an approximating sequence κ_N, since each κ_N might be irreducible.

Theorem 12 (LDP in the reducible case)**.** *Suppose the setting as in Theorem 1, with the only exception that the kernel $\kappa = \kappa_N$ on $\mathcal{S} \times \mathcal{S}$ is now assumed to be reducible, and that $\mathcal{S}$ is equal to the support of μ. Then $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ satisfies an LDP with rate function $\widetilde{I}$ defined by*

$$\widetilde{I}(\lambda, \alpha) = \begin{cases} I(\lambda, \alpha), & \text{if } \alpha \text{ is connectable,} \\ +\infty & \text{otherwise,} \end{cases}$$

where I is as in Theorem 1.

The proof follows in a straightforward way from our results, see Remark 36 for the finite type case. The intuitive reason is that, for α that is not connectable, on the event $\{\mathrm{Ma}_N \approx \alpha\}$, there is a macroscopic component who has two non-trivial parts (i.e., each with $\asymp N$ types) in two different irreducible classes, even though there cannot be any edge between these sets. Hence this event has the probability zero.

We made the choice to state the theorem under the additional assumption that $\mathcal{S}$ corresponds to the support of μ. If this was not the case, one could still approach a not connectable measure $y \in \mathcal{M}(\mathcal{S})$ with a finite exponential cost if, for each finite N, the support of y is contained in an irreducible class

of κ with respect to $\mu^{(N)}$. The proof would anyway be an extension of our finite type case results, but it is out of our scope to cover this particular framework.

2.4 Motivation: an inhomogeneous coagulation process

With the present work, we actually continue our study of random particle models with coagulation in the light of large-deviation arguments initiated in [AKP19]. Indeed, we make here the first step towards a *spatial* model.

The model that we are interested in is the following. Fix N atoms $1, \dots, N$ at the locations $x_1, \dots, x_N$ in a compact metric space $\mathcal{S}$. We consider a Markov process in continuous time on the set of partitions of $[N] = \{1, \dots, N\}$. Starting with the monodispersed configuration $M(0) = (\{i\})_{i\in[N]}$, at any time any two subsets A, B in the current partition are replaced by their union $A \cup B$ after an exponentially distributed random time with parameter

$$\frac{1}{N} \sum_{i\in A, j\in B} \kappa(x_i, x_j), \tag{2.13}$$

where a symmetric $\kappa\colon \mathcal{S} \times \mathcal{S} \to [0, \infty)$ is given. All these random times are supposed to be independent. If $M(t)$ denotes the partition at time t, then $M(s)$ is a refinement of $M(t)$ for any $s < t$. Hence, the number of elements of $M(t)$ is a non-increasing (random) process starting at N. The special case of a singleton $\mathcal{S}$ and $\kappa \equiv 1$ (the homogeneous case) is the case of the *Marcus–Lushnikov model* that we studied in [AKP19]. There we also explained how the Erdős–Rényi model can be mapped onto the Marcus–Lushnikov model, and this works also in the inhomogeneous setting. Indeed, to any unordered pair $\{i, j\} \subset [N]$ with $i \neq j$ we associate an exponential random time $e(i, j)$ with parameter $\frac{1}{N}\kappa(x_i, x_j)$. These random times are independent and we put a bond between i and j as soon as $e(i, j)$ elapses. At a fixed time $t \in (0, \infty)$, this graph has the distribution of the inhomogeneous random graph $\mathcal{G}_{t,N} = \mathcal{G}([N], (x_1, \dots, x_N), \frac{1}{N}\kappa_{t,N})$ with type space $\mathcal{S}$ and

$$\kappa_{t,N}(r, s) = N\big(1 - \mathrm{e}^{-\frac{1}{N}t\kappa(r,s)}\big), \qquad r, s \in \mathcal{S}.$$

Notice that the random partition $M(t)$ of the above coagulation model is equal in distribution to the collection $(\mathcal{C}_j)_j$ of the vertex sets of the components of $\mathcal{G}_{t,N}$. The two main reasons for this fact are the memorylessness of the exponential distribution and the property that the minimum of independent exponential times is also exponential with a parameter that is the sum of all the parameters. The only difference between the two models is that the graph model registers all the bonds that arrive within each of the components (and do not change anything in the connectedness), while the coagulation model just registers that a given set is connected.

We are interested in an LDP for the micro and the macro empirical measure of the partition sets of $M(t)$ in the limit $N \to \infty$, assuming the initial locations of particles are such that $\frac{1}{N}\sum_{i=1}^{N} \delta_{x_i} \to \mu$ for some measure μ on $\mathcal{S}$. Since $\kappa_{t,N} \to t\kappa$, Theorem 1 applies also to the above inhomogeneous coagulation process under the appropriate assumptions at a fixed time t. Furthermore, from Theorem 7, we obtain that the process $(M(t))_{t\in[0,\infty)}$ has a phase transition at the time

$$t_{\mathrm{c}} = \frac{1}{\Sigma(\kappa, \mu)},$$

and we have a limiting distribution of the empirical micro and macro measures. This phase transition is of explosion type, as described in Remark 8, and in the coagulation literature is usually called

gelation. Further consequences for the limiting distribution of $M(t)$ as $N \to \infty$ follow in a natural way, but we refrain from writing them down.

Interestingly, we can deduce that the minimizing process of microscopic cluster sizes satisfies the multitype version of the *Flory equation*, which is a modification of the well-known *Smoluchowski equation*. This is formulated as follows. We think of an inhomogeneous deterministic coagulation process $(\lambda_t)_{t\in[0,\infty)}$, conceived as a process in $\mathcal{L}$. Each particle $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ consists of $k(\mathcal{S})$ atoms, $k(\{r\})$ of which have the type r for any $r \in \mathcal{S}$. Coagulation is nothing but addition of measures in this formulation, i.e., two particles k and $\widetilde{k}$ coagulate to a particle $k+\widetilde{k}$. The kernel of this process is given as

$$K(k,\widetilde{k}) = \langle \widetilde{k}, \kappa k\rangle = \int_{\mathcal{S}^2} \kappa(r,s)\, k(\mathrm{d}r)\widetilde{k}(\mathrm{d}s), \qquad k,\widetilde{k} \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}).$$

Then the weak formulation of the Flory equation is, for any test function $f \in \mathcal{C}_{\mathrm{c}}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}))$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} f(k)\,\lambda_t(\mathrm{d}k) =& \frac{1}{2}\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})^2} f(k+\widetilde{k})K(k,\widetilde{k})\,\lambda_t(\mathrm{d}k)\lambda_t(\mathrm{d}\widetilde{k})\\ &- \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})^2} f(k)K(k,\widetilde{k})\,\lambda_t(\mathrm{d}k)\lambda_t(\mathrm{d}\widetilde{k})\\ &- \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})^2} f(k)K(k,\widetilde{k})\,\lambda_t(\mathrm{d}k)(\lambda_0-\lambda_t)(\mathrm{d}\widetilde{k}),\end{aligned} \tag{2.14}$$

where $\lambda_0(\mathrm{d}k) = \int_{\mathcal{S}} \mu(\mathrm{d}r)\,\delta_{\delta_r}(\mathrm{d}k)$ is the initial condition, which expresses that μ is the atom type distribution. In words, the time-evolution of $(\lambda_t)_{t\in[0,\infty)}$ is described by saying that any coagulation of two particles k and $\widetilde{k}$ (i.e., replacement of k and $\widetilde{k}$ by $k+\widetilde{k}$) happens with rate $K(k,\widetilde{k})$. In our model the last term in the right-hand side can be rewritten as $-\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} f(k)\langle k, \kappa(\mu - c_{\lambda_t})\rangle\,\lambda_t(\mathrm{d}k)$. It captures the interaction between the microscopic particles and the gel (the macroscopic mass) once it forms.

See [NZ11, Section 3] for a mathematical discussion of the (well-known) homogeneous version of the Flory equation and [Nor00, Section 2] for an introduction of the inhomogeneous version of the equation. We now identify a solution $(\lambda_t)_{t\in[0,\infty)}$ to (2.14).

Lemma 13 (Solution to the Flory equation)**.** *Assume that $\mathcal{S}$ is a finite state space and κ an irreducible nonnegative symmetric matrix on $\mathcal{S}$. Let $\lambda_0(k) = \sum_{r\in\mathcal{S}} \mu_r\,\delta_{\delta_r}(k)$ and for $t \in (0,\infty)$, define λ_t to be the first component of the minimizer appearing in Theorem 7 with κ replaced by $t\kappa$.*

Then $t \mapsto \lambda_t$ is a solution to the Flory equation (2.14) *on* $[0,\infty)$.

The proof of Lemma 13, as well as an explicit expression for $(\lambda_t)_{t\geq 0}$, is given in Section 7.4. We are confident that Lemma 13 is also true in the general setting of Theorem 7.

The Flory equation is closely related to the *Smoluchowski equation*, which we write in its multitype version:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} f(k)\,\lambda_t(\mathrm{d}k) =& \frac{1}{2}\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})^2} f(k+\widetilde{k})K(k,\widetilde{k})\,\lambda_t(\mathrm{d}k)\lambda_t(\mathrm{d}\widetilde{k})\\ &- \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})^2} f(k)K(k,\widetilde{k})\,\lambda_t(\mathrm{d}k)\lambda_t(\mathrm{d}\widetilde{k}).\end{aligned} \tag{2.15}$$

The Smoluchowski equation only considers the microscopic clusters, that is, it excludes any interaction with a possible gel, which we see in the third line of the Flory equation (2.14). The solutions of

equations (2.14) and (2.15) coincide until the gelation time $t_c = 1/\Sigma(\kappa, \mu)$, after which differences appear. At the level of the underlying stochastic microscopic models this difference is seen in terms of the type of the phase transitions, as cited in Remark 8. The microscopic models of *frozen percolation*-type, as in [CRY21, MN14, RT$^+$09, Yeo18], correspond to (2.15), while models like ours correspond to (2.14).

3 Connection probabilities and the characteristic equation

In this and the subsequent section we will exclusively deal with the finite type case, i.e., we assume that $\mathcal{S}$ is a finite set. We assume that $\kappa = (\kappa(r,s))_{r,s\in\mathcal{S}} \in [0,\infty)^{\mathcal{S}\times\mathcal{S}}$ is a nonnegative and symmetric matrix. For any $N \in \mathbb{N}$, let κ_N be another such matrix, and assume that κ_N converges pointwise to κ as $N \to \infty$. We write vectors $y \in \mathbb{R}^{\mathcal{S}}$ as $(y_r)_r$ and $\langle a, f\rangle = \sum_r a_r f_r$ for the integral of a function f with respect to a measure a on $\mathcal{S}$.

Let $k \in \mathbb{N}_0^{\mathcal{S}}$ and let $\mathbf{x} = (x_1, \dots, x_{|k|}) \in \mathcal{S}^{|k|}$ be a type vector which is compatible with k, meaning that $\sum_{i=1}^{|k|} \delta_{x_i} = k$. The aim of this section is to understand the asymptotic behavior of the connection probability of the graph $\mathcal{G}(|k|, \mathbf{x}, \frac{1}{N}\kappa_N)$ given by

$$p_N(k) = \mathbb{P}\big(\mathcal{G}(|k|, \mathbf{x}, \tfrac{1}{N}\kappa_N) \text{ is connected}\big). \tag{3.1}$$

It will be crucial to distinguish between the following two cases. In the first case, we keep $k \in \mathbb{N}_0^{\mathcal{S}}$ fixed, whereas $N \to \infty$. We will refer to $p_N(k)$ as the connection probability of a microscopic cluster. In the second case we consider a sequence $k^{(N)} \in \mathbb{N}_0^{\mathcal{S}}$ and assume that for any $s \in \mathcal{S}$ the limit $\lim_{N\to\infty} \frac{k_s^{(N)}}{N} = y_s$ exists and the vector $y = (y_s)_{s\in\mathcal{S}}$ is non-trivial. In that case we refer to $p_N(k^{(N)})$ as the connection probability of a macroscopic cluster.

It will be easy to see that the connection probability of a microscopic cluster is asymptotically close to the cost of connecting the cluster via a spanning tree. Let us recall the combinatorial quantity τ that collects the weight of all spanning trees. For $k \in \mathbb{N}_0^{\mathcal{S}}$ we defined

$$\tau(k) = \sum_{T\in\mathcal{T}(k)} \prod_{\{i,j\}\in E(T)} \kappa(x_i, x_j), \tag{3.2}$$

where $\mathbf{x} \in \mathcal{S}^{|k|}$ is compatible with k and $\mathcal{T}(k)$ is the set of spanning trees on $[|k|]$. By τ_N we will denote the same quantity, but defined with respect to the kernel κ_N. The following lemma is an immediate consequence of Lemma 23 that we will prove in Section 3.2.

Lemma 14. *Fix $k \in \mathbb{N}_0^{\mathcal{S}}$. Then, as $N \to \infty$,*

$$p_N(k) = N^{-(|k|-1)}\tau_N(k)(1 + o(1)). \tag{3.3}$$

Understanding the asymptotics of the connection probability of a macroscopic cluster is much harder. Most of this section is dedicated to the verification of the following result, which is the multi-type version of a result from [Ste70], see [AKP19, Lemma 2.4].

Theorem 15 (Asymptotics of the connection probability of macroscopic clusters)**.** *Fix $y \in [0,1]^{\mathcal{S}} \setminus \{0\}$. Let $\{k^{(N)}\}_{N\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0^{\mathcal{S}}$ such that $\lim_{N\to\infty} \frac{k_r^{(N)}}{N} = y_r$ for all $r \in \mathcal{S}$ and assume that $\{k_r^{(N)}\}_N$ is bounded in N for all $r \notin \operatorname{supp}(y)$. Further, assume that $\tau(k^{(N)}) > 0$ for all but finitely many $N \in \mathbb{N}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \log p_N(k^{(N)}) = \sum_{r\in\mathcal{S}} y_r \log\left(1 - \mathrm{e}^{-\kappa y_r}\right) \in [-\infty, 0]. \tag{3.4}$$

We use the convention $0 \log 0 = 0$ *and* $\log 0 = -\infty$.

The assumptions that $\{k_r^{(N)}\}_N$ is bounded for all $r \notin \mathrm{supp}(y)$ and $\tau(k^{(N)}) > 0$ for almost all $N \in \mathbb{N}$ are only needed for the lower bound in (3.4). If the right-hand side of (3.4) is not equal to $-\infty$, then the assumption on $\tau(k^{(N)})$ is crucial for the lower bound, since it implies that $p_N(k^{(N)}) > 0$ for all but finitely many N. Note that if $y > 0$ on $\mathcal{S}$ and κ is irreducible with respect to y, then both assumptions for the lower bound are satisfied.

Roughly, the idea of the proof is to construct a larger graph that contains with high probability a macroscopic component $k^{(N)}$. In order to choose the right parameters for the larger graph we have to understand a certain characteristic equation that describes how a graph will split into several microscopic clusters and a macroscopic one.In Section 3.1 we investigate the characteristic equation from the point of view of the microscopic clusters, more precisely, from the point of view of weighted trees and their generating function. In Section 3.2 we provide the asymptotics of the connection probabilities for the two cases.

3.1 The characteristic equation and tree combinatorics

In this section we discuss a power series representation of the solution of fixed point equation (2.4) in the finite type setting. This will be crucial both in the proof of Theorem 15 and in the analysis of the minimizers of the rate function in Section 6.

We rewrite the equation in the following way. Fix $\nu = (\nu_s)_{s\in\mathcal{S}} \in [0,\infty)^{\mathcal{S}}$. We say that $\nu^* = (\nu_s^*)_{s\in\mathcal{S}}$ is a solution to the characteristic equation with respect to ν if

$$\nu_s^* \mathrm{e}^{-(\kappa\nu^*)_s} = \nu_s \mathrm{e}^{-(\kappa\nu)_s}, \quad s \in \mathcal{S}. \tag{3.5}$$

Recall that this is equivalent to the characteristic equation (2.11) that was studied in [BJR07] via the substitution $\nu = \mu$ and $\rho = 1 - \frac{\mathrm{d}\nu^*}{\mathrm{d}\nu}$. However, we will need equation (3.5) on different occasions and for several choices of ν. Note that any solution ν^* to (3.5) is necessarily non-negative and for any $s \in \mathcal{S}$ we have that $\nu_s^* > 0$ if and only if $\nu_s > 0$. Also note that ν itself is always a solution. Whether there exists a non-trivial solution ν^* (i.e., $\nu^* \neq \nu$) or not will turn out to depend on the quantity $\Sigma(\kappa,\nu)$, introduced in (2.2) and (2.3):

$$\begin{aligned}\Sigma(\kappa,\nu) &= \|T_{\kappa,\nu}\|_\nu = \sup\Big\{\|T_{\kappa,\nu}f\|_\nu \colon f \in [0,\infty)^{\mathcal{S}}, \|f\|_\nu = 1\Big\} \\ &= \sup\Big\{\Big(\sum_{r\in\mathcal{S}} \nu_r \Big(\sum_{s\in\mathcal{S}} \kappa(r,s) f(s) \nu_s\Big)^2\Big)^{1/2} \colon f \in [0,\infty)^{\mathcal{S}}, \sum_{r\in\mathcal{S}} \nu_r f(r)^2 \le 1\Big\},\end{aligned} \tag{3.6}$$

where we write $\|\cdot\|_\nu$ for the norm on $L^2(\mathbb{R}^{\mathcal{S}},\nu)$ and also for the corresponding operator norm. We note that $\Sigma(\kappa,\nu)$ is equal to the spectral radius of the matrix $T_{\kappa,\nu} = (\kappa(r,s)\nu_s)_{r,s\in\mathcal{S}}$, as is seen from an elementary analysis of (3.6), also using Frobenius eigenvalue theory. Indeed, the variational equations for the maximizer f in (3.6) (with $\sum_r \nu_r f(r)^2 = 1$ instead of ≤ 1) say that f is a positive eigenvector of the matrix $T_{\kappa,\nu}^2$. Since also the (up to positive multiples, unique) positive eigenvector of $T_{\kappa,\nu}$ is a positive eigenvector of $T_{\kappa,\nu}^2$, we have that f is the Frobenius eigenvector of $T_{\kappa,\nu}$. The corresponding Frobenius eigenvalue, i.e., the spectral radius, is equal to $\Sigma(\kappa,\nu)$.

Another elementary application of Frobenius eigenvalue theory yields that the map $\nu \mapsto \Sigma(\kappa,\nu)$ is non-decreasing with respect to componentwise ordering.

Now we cite the results from [BJR07] regarding the solutions of the characteristic equation.

Lemma 16. *Let $\nu \in [0,\infty)^{\mathcal{S}}$.*

(i) If $\Sigma(\kappa,\nu) \leq 1$, then the only solution ν^* to the characteristic equation (3.5) *satisfying $\nu^* \leq \nu$ is given by $\nu^* = \nu$.*

(ii) If $\Sigma(\kappa,\nu) > 1$, then there exists a solution ν^* to (3.5) *that satisfies $\nu^* \leq \nu$ and $\nu^* \neq \nu$. If additionally κ is irreducible, then ν^* is the only solution to* (3.5) *that satisfies $\nu^* \leq \nu$ and $\nu^* \neq \nu$. Further, $\Sigma(\kappa,\nu^*) < 1$.*

Proof. See Theorem 6.2 and Theorem 6.7 in [BJR07] and substitute $\rho = 1 - \frac{\mathrm{d}\nu^*}{\mathrm{d}\nu}$. □

Our aim is to verify the following.

Proposition 17. *Let $\nu \in [0,\infty)^{\mathcal{S}}$. Then for any $r \in \mathcal{S}$*

$$\sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \tau(k)k_r \prod_{s\in\mathcal{S}} \frac{(\nu_s \mathrm{e}^{-(\kappa\nu)_s})^{k_s}}{k_s!} = \nu_r^* \tag{3.7}$$

where ν^* is the smallest solution to (3.5)*.*

The result of Proposition 17 will be used in Section 3.2 to derive the asymptotics of the probability that a macroscopic set of vertices is connected. Further, it will be used in Section 6 to optimize the microscopic part of the rate function. The left-hand side of equation (3.7), when divided by ν_r, is equal to the extinction probability of the branching process that we mentioned in Section 2.2. This observation, together with additional results from [BJR07], is already enough to prove Proposition 17. However, we will provide a different proof that mainly uses the structure of the power series and additionally gives us a refined control of the convergence of the series, thanks to the estimates given in Lemma 20. They will be used later to bound probabilistic terms that we derive from components of mesoscopic sizes, but also in Section 6 where we need uniform convergence for a certain family of power series.

We now prepare for the proof of Proposition 17. We define a function $\Gamma = (\Gamma_r)_{r\in\mathcal{S}} \colon [0,\infty)^{\mathcal{S}} \to [0,\infty)^{\mathcal{S}}$ by putting, for $\theta = (\theta_s)_{s\in\mathcal{S}} \in [0,\infty)^{\mathcal{S}}$,

$$\Gamma_r(\theta) := \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \tau(k)k_r \prod_{s\in\mathcal{S}} \frac{\theta_s^{k_s}}{k_s!} \in [0,\infty] \quad \text{for } r \in \mathcal{S}. \tag{3.8}$$

The idea of the proof is to show that $\theta \mapsto \Gamma(\theta)$ is the inverse of the function $\nu \mapsto \theta(\nu) := \nu \mathrm{e}^{-\kappa\nu}$ on the domain $\{\nu[0,\infty)^{\mathcal{S}} \colon \Sigma(\kappa,\nu) \leq 1\}$. It turns out that this is an easy example of a technique known as Lagrange inversion, where directed trees are used to extract the variables of a power series, see [JKT19, JKT21] for more general results on this topic. The relation of our combinatorial quantity τ to directed trees is given in the following lemma. The second statement will be useful to derive a criterion for analyticity of the power series defined in (3.8).

Lemma 18. *Let $k \in \mathbb{N}_0^{\mathcal{S}}$ and $r \in \mathcal{S}$. Then the following holds.*

1 Let $\overrightarrow{\mathcal{T}}_r(k)$ denote the set of directed trees with vertex set $\{1,\dots,|k|\}$, root of type r and edges directed away from the root. Then

$$\tau(k)k_r = \sum_{T\in\overrightarrow{\mathcal{T}}_r(k)} \prod_{(i,j)\in E(T)} \kappa(x_i,x_j). \tag{3.9}$$

2 *Write* $S_0 := \mathrm{supp}(k) \subset \mathcal{S}$ *and* $\overrightarrow{T}_r(S_0)$ *for the set of all directed trees with vertex set* S_0*, root vertex given by* r *and edges directed away from the root. Then*

$$\tau(k)k_r = \left(\prod_{s\in S_0}(\kappa k)_s^{k_s-1}\right) \times \Delta_r(k) \tag{3.10}$$

where

$$\Delta_r(k) = \sum_{A\in\overrightarrow{T}_r(S_0)} \prod_{(s,s')\in E(A)} \kappa(s,s')k_s. \tag{3.11}$$

Proof. (1) Given some vertex $i \in [|k|]$ of type r, we can turn any undirected tree from $\mathcal{T}(k)$ uniquely into a directed tree that has vertex i as its root by giving to each edge the direction away from the root. For each undirected tree T in $\mathcal{T}(k)$ there are k_r ways to choose the root, hence the weight of T appears k_r times on the right-hand side of (3.9).

(2) We use the formula derived in [BM14, Theorem 2] with $x_{s,s',i} := \kappa(s,s') = \kappa(s',s)$. Note that in [BM14] the edges of the trees are directed towards the root. Adapted to our notation their formula reads as

$$\sum_{T\in\overrightarrow{\mathcal{T}}_r(k)} \prod_{s,s'\in S_0} \kappa(s,s')^{\#\{(i,j)\colon x_i=s, x_j=s'\}} = \left(\prod_{s\in S_0}(\kappa k)_s^{k_s-1}\right) \times \Delta_r(k). \tag{3.12}$$

By (3.9) we have that

$$\tau(k)k_r = \sum_{T\in\overrightarrow{\mathcal{T}}_r(k)} \prod_{(i,j)\in E(T)} \kappa(x_i,x_j) = \sum_{T\in\overrightarrow{\mathcal{T}}_r(k)} \prod_{s,s'\in S_0} \kappa(s,s')^{\#\{(i,j)\colon x_i=s, x_j=s'\}},$$

so together with (3.12) we have proven equation (3.10). Note that in the case $r \notin S_0$, we have that $\overrightarrow{T}_r(S_0) = \emptyset$, so both sides of (3.12) are 0. □

Lemma 19. *If* Γ *is analytic in* θ *(i.e.,* Γ_r *is analytic in* θ *for all* $r \in \mathcal{S}$*), then for all* $r \in \mathcal{S}$

$$\Gamma_r(\theta) = \theta_r \exp\left((\kappa\Gamma(\theta))_r\right). \tag{3.13}$$

Proof. Using formula (3.9) we can rewrite the power series using directed trees. Given a tree T in $\overrightarrow{\mathcal{T}}_r(k)$ and some $h \in \mathbb{N}_0^{\mathcal{S}}$ we say that a vertex $i \in [|k|]$ is of parent-type (r,h) if i is of type r and points to exactly h_s vertices of type s for any $s \in \mathcal{S}$. In that case, its weight is defined by $w(i) = \prod_{s\in\mathcal{S}} \kappa(r,s)^{h_s} =: g_{r,h}$. The weight of the tree T is defined by $w(T) := \prod_{i=1}^{|k|} w(i)$. Defining $g_r(x) = \sum_{h\in\mathbb{N}_0^{\mathcal{S}}} g_{r,h} \prod_{s\in\mathcal{S}} \frac{x_s^{h_s}}{h_s!}$, for $x \in \mathbb{R}^{\mathcal{S}}$, we can apply the multinomial theorem to see that $g_r(x) = \exp\{(\kappa x)_r\}$. Althogether, we have that

$$\Gamma_r(\theta) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \Big(\sum_{T\in\overrightarrow{\mathcal{T}}_r(k)} w(T)\Big) \prod_{s\in\mathcal{S}} \frac{\theta_s^{k_s}}{k_s!},$$

which is an exponential generating function, evaluated in $\theta \in \mathbb{R}^{\mathcal{S}}$ and shall be understood as a formal power series. We now cite a fact from [Ges87] (see the proof of Theorem 1):

$$\Gamma_r(\theta) = \theta_r \sum_{h\in\mathbb{N}_0^{\mathcal{S}}} g_{r,h} \prod_{s\in\mathcal{S}} \frac{(\Gamma_s(\theta))^{h_s}}{h_s!} = \theta_r g_r(\Gamma(\theta)), \qquad r \in \mathcal{S}, \tag{3.14}$$

where all equations denote equality of formal power series, i.e., we have equality of the coefficients of $\theta_s^{k_s}$. In particular, $\Gamma_r(\theta)$ might be infinite. However, by our assumption that Γ is analytic in θ, equation (3.14) holds in the usual sense, i.e., all series in (3.14) converge absolutely. Recalling that $g_r(x) = \exp\{(\kappa x)_r\}$, we have verified that (3.13) holds.

The idea behind proving (3.14) is to decompose the set of trees with root of type r into sets of trees with root of parent-type (r, h) for each $h \in \mathbb{N}_0^{\mathcal{S}}$. For fixed $h \in \mathbb{N}_0^{\mathcal{S}}$, one then decomposes a tree with root of parent-type (r, h) into a single vertex of type r and exactly h_s many trees with root of type s for each $s \in \mathcal{S}$. By studying how this decomposition affects the exponential generating function, one obtains the formula. □

Next, we derive a criterion for analyticity of the map $\theta \mapsto \Gamma(\theta)$, i.e., about the domain of convergence of the corresponding power series. We need to introduce the quantity

$$\chi(\kappa, \theta) = \inf\Big\{\Big\langle v, \log\frac{v}{(\kappa v)\theta}\Big\rangle \colon v \in \mathcal{M}_1(\mathcal{S}), v \ll \theta\Big\}, \tag{3.15}$$

where $\mathcal{M}_1(\mathcal{S})$ denotes the set of probability measures on $\mathcal{S}$. One can easily see that $\theta \mapsto \chi(\kappa, \theta)$ is lower semi-continuous.

Lemma 20. *1 For any $\theta \in [0, \infty)^{\mathcal{S}}$ and $r \in \mathcal{S}$,*

$$\sum_{k \in \mathbb{N}_0^{\mathcal{S}} \colon |k| = n} \tau(k) k_r \prod_{s \in \mathcal{S}} \frac{\theta_s^{k_s}}{k_s!} = \mathrm{e}^{o(n)} \mathrm{e}^{-n[\chi(\kappa,\theta)-1]}, \qquad n \to \infty. \tag{3.16}$$

2 Fix $\theta \in [0, \infty)^{\mathcal{S}}$. If $\chi(\kappa, \theta) > 1$, then for any $r \in \mathcal{S}$ the series Γ_r defined in (3.8) *is analytic in θ. If on the other hand $\chi(\kappa, \theta) < 1$, then $\Gamma_r(\theta)$ diverges.*

3 For arbitrary $\nu \in (0, \infty)^{\mathcal{S}}$ and $\theta_s(\kappa, \nu) = \nu_s \mathrm{e}^{-(\kappa\nu)_s}$, $s \in \mathcal{S}$, we have that

$$\chi(\kappa, \theta(\kappa, \nu)) = \Sigma(\kappa, \nu) - \log \Sigma(\kappa, \nu) \geq 1 \tag{3.17}$$

with equality if and only if $\Sigma(\kappa, \nu) = 1$.

Proof. (1) We will use the identity (3.10) from Lemma 18. Let $n \in \mathbb{N}$. We write $\mathcal{M}_1^{(n)}(\mathcal{S})$ for the set of all probability measures v on $\mathcal{S}$ satisfying $nv \in \mathbb{N}_0^{\mathcal{S}}$. For any $k \in \mathbb{N}_0^{\mathcal{S}}$ with $|k| = n$ we substitute $k = nv$ to get

$$\begin{aligned}\sum_{k \in \mathbb{N}_0^{\mathcal{S}} \colon |k| = n} \tau(k) k_r \prod_{s \in \mathcal{S}} \frac{\theta_s^{k_s}}{k_s!} &= \sum_{v \in \mathcal{M}_1^{(n)}(\mathcal{S})} \Big(\prod_{s \in \mathrm{supp}(v)} (n(\kappa v)_s)^{nv_s - 1} \Big) \prod_{s \in \mathcal{S}} \frac{\theta_s^{nv_s}}{(nv_s)!} \Delta_r(nv) \\ &= \sum_{v \in \mathcal{M}_1^{(n)}(\mathcal{S})} \mathrm{e}^{o(n)} n^{-|\mathrm{supp}(v)|} \mathrm{e}^{n\left[1 - \sum_s v_s \log \frac{v_s}{(\kappa v)_s \theta_s}\right]} \Delta_r(nv),\end{aligned}$$

where we used Stirlings formula $N! = N^N \mathrm{e}^{-N} \mathrm{e}^{o(N)}$ as $N \to \infty$ and assumed throughout that $k \ll \theta$ and $v \ll \theta$ respectively. Note that $\big|\mathcal{M}_1^{(n)}(\mathcal{S})\big| = \mathrm{e}^{o(n)}$ and that the terms $n^{-|\mathcal{S}|}$ and $\Delta_r(nv)$ are only polynomial in n, thus also of order $\mathrm{e}^{o(n)}$. Collecting everything, we arrive at (3.16).

(2) As a consequence of (1) we get the following. If $\chi(\kappa, \theta) > 1$, then the series $\Gamma_r(\theta)$ converges and since the mapping $\theta \mapsto \chi(\kappa, \theta)$ is lower-semicontinuous, we get that Γ_r is analytic in θ. If $\chi(\kappa, \theta) < 1$, then $\Gamma_r(\theta)$ diverges.

(3) Put $\phi(x) = x\log x$ for $x \geq 0$. Inserting the definition $\theta_s(\nu) = \nu_s \mathrm{e}^{-(\kappa\nu)_s}$, $s \in \mathcal{S}$, we can estimate for any $v \in \mathcal{M}_1(\mathcal{S})$

$$\Big\langle v, \log\frac{v}{(\kappa v)\theta(\nu)}\Big\rangle = \Big\langle v, \log\frac{v}{(\kappa v)\nu}\Big\rangle + \langle v, \kappa\nu\rangle = \langle \nu, \kappa v\rangle \sum_{r\in\mathcal{S}} \frac{(\kappa v)_r \nu_r}{\langle \nu, \kappa v\rangle} \phi\Big(\frac{v_r}{(\kappa v)_r \nu_r}\Big) + \langle v, \kappa\nu\rangle$$
$$\geq \langle \nu, \kappa v\rangle \phi\Big(\frac{\sum_r v_r}{\langle \nu, \kappa v\rangle}\Big) + \langle v, \kappa\nu\rangle = -\log\langle \nu, \kappa v\rangle + \langle \nu, \kappa v\rangle$$

where the estimate is due to Jensen's inequality applied to the convex function ϕ. Further, we used the fact that v is a probability measure and the symmetry of κ. Since ϕ is even strictly convex, the application of Jensen's inequality gives an equality if and only if v is such that the map $\mathcal{S} \ni r \mapsto v_r/((\kappa v)_r \nu_r)$ is constant, i.e., if there is some $a \in \mathbb{R}$ such that $(\kappa v)_r = a\frac{v_r}{\nu_r}$ for any $r \in \mathcal{S}$. Clearly, $a > 0$. It follows that $\langle \nu, \kappa v\rangle = a$ and that $w = (v_r/\nu_r)_{r\in\mathcal{S}}$ is an eigenvector of the matrix $T_{\kappa,\nu} = (\kappa(r,s)\nu_s)_{r,s\in\mathcal{S}}$ with eigenvalue a. Our assumption $\nu > 0$ and the irreducibility of κ imply that $T_{\kappa,\nu}$ is also irreducible. Hence the Perron–Frobenius theorem gives that, if a would be smaller than the spectal radius $\Sigma(\kappa,\nu)$ of $T_{\kappa,\nu}$, then w could not be non-negative, implying that $v = (w_r\nu_r)_r$ would not be in $\mathcal{M}_1(\mathcal{S})$. Hence, a is necessarily equal to $\Sigma(\kappa,\nu)$ and equation (3.17) holds. Since $x - 1 \geq \log x$ holds for any $x \geq 0$ with equality if and only if $x = 1$, the rest of the statement in (3) holds.

□

Now, we can give the proof of Proposition 17.

Proof of Proposition 17. We will proceed in three steps and weaken the assumptions on ν gradually.

(1) Assume that $\nu \in (0,\infty)^{\mathcal{S}}$ and that κ is irreducible with respect to ν. Consider the case $\Sigma(\kappa,\nu) \neq 1$, then by Lemma 20 we have that $\chi(\kappa,\theta(\nu)) > 1$ and thus for any $r \in \mathcal{S}$ the power series Γ_r defined as in (3.8) is analytic in $\theta(\nu)$. Applying Lemma 19 and using the definition of $\theta(\nu)$ we get that

$$\nu_r \exp(-(\kappa\nu)_r) = \theta_r(\nu) = \Gamma_r(\theta(\nu)) \exp(-(\kappa\Gamma(\theta(\nu)))_r) \tag{3.18}$$

for any $r \in \mathcal{S}$. In other words, $\Gamma(\theta(\nu))$ is a solution of (3.5). Now, Lemma 16 gives that $\nu = \Gamma(\theta(\nu))$ if $\Sigma(\kappa,\nu) \leq 1$ and the claim follows. If $\Sigma(\kappa,\nu) > 1$, then by Lemma 16 there exists (a strictly positive) $\nu^* \leq \nu$, $\nu^* \neq \nu$, solving equation (3.5) and satisfying $\Sigma(\kappa,\nu^*) < 1$. So, $\Gamma(\theta(\nu)) = \Gamma(\theta(\nu^*)) = \nu^*$ follows by applying the previous case.

Now, consider the case $\Sigma(\kappa,\nu) = 1$, which is equivalent to $\chi(\theta(\nu)) = 1$. Let $\nu^{(n)} \nearrow \nu$, in particular $\Sigma(\kappa,\nu^{(n)}) < 1$ for all n. Then by Fatou's lemma and the first case we get that for any $r \in \mathcal{S}$

$$\Gamma_r(\theta(\nu)) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \tau(k) k_r \prod_s \frac{\theta_s(\nu)^{k_s}}{k_s!} \leq \liminf_{n\to\infty} \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \tau(k) k_r \prod_s \frac{\theta_s(\nu^{(n)})^{k_s}}{k_s!} = \liminf_{n\to\infty} \nu_r^{(n)} = \nu_r. \tag{3.19}$$

Put $\nu^* := \Gamma(\theta(\nu))$ and assume towards a contradiction that $\nu^* \neq \nu$. Then there exists $s \in \mathcal{S}$ such that $\nu_s^* < \nu_s$ and by irreducibility of κ there exists at least one s' such that $\kappa(s',s) > 0$, and thus $\kappa(s',s)\nu_s^* < \kappa(s',s)\nu_s$. Hence the Perron–Frobenius theorem implies that $\Sigma(\kappa,\nu^*) < 1$. Therefore the power series is analytic in $\theta(\nu)$, so by Lemma 19 and the definition of $\theta(\nu)$ we get that $\nu^* \exp(-\kappa\nu^*) = \nu \exp(-\kappa\nu)$. Applying Lemma 16 yields $\nu^* = \nu$ in contradiction to our assumption.

(2) Let $\nu \in (0,\infty)^{\mathcal{S}}$ and an arbitrary κ. Then we can decompose $\mathcal{S}$ into disjoint sets S_j, $j \in J$, such that $\kappa^{(j)} = \kappa|_{S_j \times S_j}$ is irreducible with respect to $\nu^{(j)} = \nu|_{S_j}$ for any $j \in J$. For any $j \in J$ we can apply (1) to get that

$$\Gamma_r^{(j)}(\theta(\nu^{(j)})) := \sum_{k \in \mathbb{N}_0^{S_j}} \tau(k) k_r \prod_{s \in S_j} \frac{(\nu^{(j)} \mathrm{e}^{-(\kappa^{(j)}\nu^{(j)})_s})^{k_s})}{k_s!} = \nu_r^{(*,j)}, \quad r \in S_j, \tag{3.20}$$

where $\nu^{(j)}$ solves equation (3.5) on S_j. Observe that if $k \in \mathbb{N}_0^{\mathcal{S}}$ is such that $\mathrm{supp}(k) \not\subset S_j$ holds for any $j \in J$, then $\tau(k) = 0$. Consequently for fixed r and j such that $r \in S_j$ we get that

$$\Gamma_r(\theta(\nu)) = \Gamma_r^{(j)}(\theta(\nu^{(j)})) \tag{3.21}$$

where we also used that $(\kappa\nu)_s = (\kappa^{(j)}\nu^{(j)})_s$ for $s \in S_j$ holds by construction. Define $\nu^* = (\nu_s^*)_{s \in \mathcal{S}}$ by putting $\nu_s^* := \nu_s^{(*,j)}$ if $s \in S_j$. Then it is easy to verify that ν^* is the smallest solution to (3.5).

(3) Let $\nu \in [0,\infty)^{\mathcal{S}}$. Observe that for $r \notin \mathrm{supp}(\nu)$ we have $\Gamma_r(\theta(\nu)) = 0$. The rest follows by restricting to $\mathrm{supp}(\nu)$ and applying (2). □

Remark 21. Connection with branching processes. *For the reader who is familiar with the results and techniques used in [BJR07], a natural question that arises is about the connection between the power series that we study in Proposition 17 and the multi-type branching process that is used in [BJR07] to explore the clusters of the random graph, as we mention in Section 2.2. Indeed, both objects carry the same information, as is shown in the following lemma.*

Lemma 22. *Fix $\nu \in (0,\infty)^{\mathcal{S}}$. Let $\mathcal{X}$ be a multi-type branching process, where the individuals are equipped with types from $\mathcal{S}$ and an individual of type $r \in \mathcal{S}$ gives birth to a number of individuals of type $s \in \mathcal{S}$ that is Poisson distributed with parameter $\kappa(r,s)\nu_s$, independently for each $s \in \mathcal{S}$. Let Ξ be the vector in $\mathbb{N}_0^{\mathcal{S}}$ that counts the total progeny of the process $\mathcal{X}$ according to their types. For $r \in \mathcal{S}$ let P_r be the probability measure under which $\mathcal{X}$ starts with one single individual of type r. Then*

$$\mathrm{P}_r(\Xi = k) = \tau(k)\frac{k_r}{\nu_r} \prod_{s \in \mathcal{S}} \frac{(\nu_s \mathrm{e}^{-(\kappa\nu)_s})^{k_s}}{k_s!}, \qquad k \in \mathbb{N}_0^{\mathcal{S}}. \tag{3.22}$$

We omit the proof of the lemma, which comes from combinatorial manipulations and properties of the Poisson distribution of the offspring.

As a consequence, the result of Proposition 17 can be reformulated in terms of the branching process $\mathcal{X}$. If ρ_r^* denotes the probability that the process $\mathcal{X}$ goes extinct under P_r, then

$$\rho_r^* = \sum_{k \in \mathbb{N}_0^{\mathcal{S}}} \mathrm{P}_r(\Xi = k) = \frac{1}{\nu_r} \sum_{k \in \mathbb{N}_0^{\mathcal{S}}} \tau(k) k_r \prod_{s \in \mathcal{S}} \frac{(\nu_s \mathrm{e}^{-(\kappa\nu)_s})^{k_s}}{k_s!}, \qquad r \in \mathcal{S}. \tag{3.23}$$

Furthermore, by substituting $\rho^* := \frac{\nu^*}{\nu}$, the statement of Proposition 17 is equivalent to the fact that the survival probability $\rho = 1 - \rho^*$ is the maximal non-negative solution to

$$\rho = 1 - \mathrm{e}^{-\kappa(\rho\nu)}$$

as stated in Theorem 11.

Recall that, in the single-type case $|\mathcal{S}| = 1$, the right-hand side of (3.22) *is identical to the Borel distribution on $\mathbb{N}_0$. Hence, we call it also in the general case where $\mathcal{S}$ is a finite set the* multi-type Borel distribution *and denote it by $\mathrm{Bo}_{\kappa,\nu}$. It is a probability measure on $\mathbb{N}_0^{\mathcal{S}}$ if and only if $\rho^* = 1$.*

3.2 Asymptotics for the connection probability

Using the results from Section 3.1 we can now study the connection probabilities for microscopic and macroscopic clusters, respectively. The first result for the microscopic clusters is elementary and uses well-known arguments. The result formulated in Theorem 15 concerning the connection probabilities of macroscopic clusters is, to the best of our knowledge, a new one and might be of independent interest in the theory of random graphs.

Recall the definition of $\tau(k)$ from (3.2) and that $\kappa \in [0,\infty)^{\mathcal{S}\times\mathcal{S}}$ is the limiting matrix of the sequence κ_N as $N \to \infty$. We define $\tau_N(k)$ as in (3.2) with respect to κ_N instead of κ. We will assume that $\kappa_N(r,s) \le N$ for any $r, s \in \mathcal{S}$, which holds in the finite type setting if $N \in \mathbb{N}$ is large enough.

Lemma 23 (Bounds for the connection probability of microscopic clusters)**.** *For any $N \in \mathbb{N}$ and $k \in \mathbb{N}_0^{\mathcal{S}}$,*

$$\left(1 - \frac{\|\kappa_N\|_\infty}{N}\right)^{|k|^2/2} \le \frac{p_N(k)}{N^{1-|k|}\tau_N(k)} \le 1. \tag{3.24}$$

Proof. We start with the upper bound. For $T \in \mathcal{T}(k)$ we denote by Ω_T the event that the edge set $E(T)$ of T is contained in the edge set of $\mathcal{G}(|k|, \mathbf{x}, \frac{1}{N}\kappa_N)$. Since $\mathcal{G}(|k|, \mathbf{x}, \frac{1}{N}\kappa_N)$ has to contain at least one spanning tree in order to be connected, we have

$$p_N(k) = \mathbb{P}\Big(\bigcup_{T\in\mathcal{T}(k)} \Omega_T\Big) \le \sum_{T\in\mathcal{T}(k)} \mathbb{P}(\Omega_T) = \sum_{T\in\mathcal{T}(k)} \prod_{\{i,j\}\in E(T)} \frac{\kappa_N(x_i,x_j)}{N} \le \left(\frac{1}{N}\right)^{|k|-1} \tau_N(k),$$

where we used the fact that each $T \in \mathcal{T}(k)$ has exactly $|k|-1$ edges.

Now we continue with the lower bound. For $T \in \mathcal{T}(k)$ we denote by $\widetilde{\Omega}_T$ the event that the edge set of $\mathcal{G}(|k|, \mathbf{x}, \frac{1}{N}\kappa_N)$ is equal to $E(T)$. Note that the events $\widetilde{\Omega}_T$, $T \in \mathcal{T}(k)$, are disjoint and therefore

$$\begin{aligned} p_N(k) &\ge \sum_{T\in\mathcal{T}(k)} \mathbb{P}(\widetilde{\Omega}_T) = \sum_{T\in\mathcal{T}(k)} \Big(\prod_{\{i,j\}\in E(T)} \frac{\kappa_N(x_i,x_j)}{N}\Big) \times \prod_{\{i,j\}\notin E(T)} \Big(1 - \frac{\kappa_N(x_i,x_j)}{N}\Big) \\ &\ge \left(1 - \frac{\|\kappa_N\|_\infty}{N}\right)^{\binom{|k|}{2}-(|k|-1)} \left(\frac{1}{N}\right)^{|k|-1} \tau_N(k) \\ &\ge \left(1 - \frac{\|\kappa_N\|_\infty}{N}\right)^{|k|^2/2} \left(\frac{1}{N}\right)^{|k|-1} \tau_N(k). \end{aligned}$$

□

Now we turn to the proof of Theorem 15, which comes as a consequence of the following Lemmas 25 – 28. The intuitive idea of the proof is to embed $\mathcal{G}(|k^{(N)}|, \mathbf{x}, \frac{1}{N}\kappa_N)$ in a larger random graph with vertex set given by some $m^{(N)} \in \mathbb{N}_0^{\mathcal{S}}$ such that the component $k^{(N)}$ appears with high probability as the typical giant component in this graph. We make heavy use of the expansion in Lemma 24, which is a straight forward multi-type generalization of equation (4) from [Gil59].

Lemma 24. *Fix $m \in \mathbb{N}_0^{\mathcal{S}}$ and $r \in \mathcal{S}$ such that $m_r \ge 1$. Then the following formula holds*

$$1 = \sum_{h\in\mathbb{N}_0^{\mathcal{S}}\,:\,\mathbf{e}_r\le h\le m} \Big[\prod_{s\in\mathcal{S}} \binom{m_s - \delta_{r,s}}{h_s - \delta_{r,s}}\Big] p_N(h) \prod_{s,\tilde{s}\in\mathcal{S}} \Big(1 - \frac{\kappa_N(s,\tilde{s})}{N}\Big)^{h_s(m_{\tilde{s}}-h_{\tilde{s}})}. \tag{3.25}$$

Proof. Consider the graph $\mathcal{G}(|m|, x, \frac{1}{N}\kappa_N)$ where x is a fixed vector compatible with m. Fix a vertex $i \in \{1, \dots, |m|\}$ of type r. For fixed $h \in \mathbb{N}_0^{\mathcal{S}}$ with $\mathbf{e}_r \leq h \leq m$ denote by $\Omega(h)$ the event that the connected component containing i is given by some set $C \subset \{1, \dots, |m|\}$ that contains exactly h_s vertices of type s for each $s \in \mathcal{S}$. We claim that the summand on the right-hand side of (3.25) is the probability of $\Omega(h)$. Indeed, there are $\prod_{s\in\mathcal{S}} \binom{m_s-\delta_{r,s}}{h_s-\delta_{r,s}}$ possibilities to choose a set $C \setminus \{i\}$ from $\{1, \dots, |m|\} \setminus \{i\}$. The probability that C is the connected component containing vertex i is given as the product of $p_N(h)$, i.e., the probability that it is connected, and the probability that no edge exists between C and its complement $\{1, \dots, |m|\} \setminus C$, which is easily seen to be equal to $\prod_{s,\tilde{s}\in\mathcal{S}}(1 - \frac{\kappa_N(s,\tilde{s})}{N})^{h_s(m_{\tilde{s}}-h_{\tilde{s}})}$. Equation (3.25) follows by the observation that the events $\Omega(h)$, $\mathbf{e}_r \leq h \leq m$, form a decomposition of the underlying probability space. □

The idea is now to pick $m = m^{(N)}$ in such a way that the summand for $h = k^{(N)}$ is maximal, such that, on the exponential scale, the right-hand side of (3.25) can be replaced by just this summand. It will turn out that the correct choice is $m^{(N)} \sim N\nu$ with $\nu_r = y_r/(1 - \mathrm{e}^{-(\kappa y)_r})$ for $r \in \mathcal{S}$. Intuitively, this choice of ν comes from inverting equation (2.11): indeed one can see that $y = \rho\nu$ where ρ solves

$$\rho = 1 - \mathrm{e}^{-T_{\kappa,\nu}\rho}.$$

Notice that the assumption $y \ll \kappa y$ is crucial for ν to be well-defined. While in the upper bound of Lemma 25 this is not important (in that case we see from (3.27) that the upper bound of $p_N(k^{(N)})$ converges to 0), we need it in order to get a non-trivial lower bound in Lemma 27.

Lemma 25 (Upper bound in (3.4))**.** *Let $k \in \mathbb{N}_0^{\mathcal{S}}$ with $k_r \geq 1$ for some $r \in \mathcal{S}$. Let $m \in \mathbb{N}_0^{\mathcal{S}}$ such that $m \leq k$. Then*

$$p_N(k) \leq \Big[\prod_{s\in\mathcal{S}} \binom{m_s - \delta_{r,s}}{k_s - \delta_{r,s}}\Big]^{-1} \prod_{s,\tilde{s}\in\mathcal{S}} \Big(1 - \frac{\kappa(s,\tilde{s})}{N}\Big)^{-k_s(m_{\tilde{s}}-k_{\tilde{s}})}. \tag{3.26}$$

Fix $y \in [0,1]^{\mathcal{S}} \setminus \{0\}$. Let $\{k^{(N)}\}_{N\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0^{\mathcal{S}}$ such that $\lim_{N\to\infty} \frac{k_r^{(N)}}{N} = y_r$ for all $r \in \mathcal{S}$. Then

$$\limsup_{N\to\infty} \frac{1}{N} \log p_N(k^{(N)}) \leq \sum_{r\in\mathcal{S}} y_r \log\big(1 - \mathrm{e}^{-(\kappa y)_r}\big), \tag{3.27}$$

where the right-hand side takes the value $-\infty$ when $y \not\ll \kappa y$.

Proof. The inequality (3.26) is a direct consequence of (3.25), where from the right-hand side we pick only the summand for $h = k$, which is present since $m \geq k$.

Let us focus now on (3.27). Let $m^{(N)} \in \mathbb{N}_0^{\mathcal{S}}$ be such that $m_r^{(N)} := \lfloor k_r^{(N)}(1 - \mathrm{e}^{-\frac{1}{N}(\kappa k^{(N)})_r})^{-1}\rfloor$ for all $r \in \mathcal{S}$, therefore $m^{(N)} \geq k^{(N)}$. Fix $r \in \mathcal{S}$ such that $k_r^{(N)} \geq 1$ and consider (3.26).

For brevity we will write k and m instead of $k^{(N)}$ and $m^{(N)}$. With the help of Stirling's formula $n! = (n/\mathrm{e})^n \mathrm{e}^{o(n)}$ and the exponential limit theorem $\lim_{n\to\infty}(1 + \frac{c}{n})^n = \mathrm{e}^c$ and some elementary calculation, we get that, as $N \to \infty$,

$$\begin{aligned}
&\Big[\prod_{s\in\mathcal{S}} \binom{m_s - \delta_{r,s}}{k_s - \delta_{r,s}}\Big] \prod_{s,\tilde{s}\in\mathcal{S}} \Big(1 - \frac{\kappa_N(s,\tilde{s})}{N}\Big)^{k_s(m_{\tilde{s}}-k_{\tilde{s}})} \\
&\quad = \exp\Big(-\Big\langle k, \log\frac{k}{m}\Big\rangle - \Big\langle m-k, \log\frac{m-k}{m}\Big\rangle\Big) \times \exp\Big(-\langle m-k, \kappa k\rangle\Big)\mathrm{e}^{o(N)} \\
&\quad = \exp\Big(-\big\langle k, \log(1 - \mathrm{e}^{-\frac{1}{N}\kappa k})\big\rangle\Big)\mathrm{e}^{o(N)},
\end{aligned}$$

where in the last step we used that $k/m = 1 - \mathrm{e}^{-\frac{1}{N}\kappa k}$ (by the definition of m). Now, we insert the asymptotics in (3.26) and see that

$$p_N(k^{(N)}) \le \exp\Big(N \sum_{r\in\mathcal{S}} \frac{k_r^{(N)}}{N} \log\big(1 - \mathrm{e}^{-\frac{1}{N}(\kappa k^{(N)})_r}\big)\Big)\mathrm{e}^{o(N)}.$$

Consequently, the claim in (3.27) follows, including the convention for $y \not\ll \kappa y$. □

In order to prove the lower bound we need the following auxiliary lemma.

Lemma 26. *For $h', h \in \mathbb{N}_0^{\mathcal{S}}$ with $h' \le h$ we have, for any $N \in \mathbb{N}$,*

$$p_N(h') \le p_N(h) \prod_{r\in\mathcal{S}} \big(1 - \mathrm{e}^{-\frac{1}{N}(\kappa h')_r}\big)^{-(h_r - h'_r)}. \tag{3.28}$$

Proof. Let $I := [|h|]$ and let $x = (x_i)_{i\in I}$ be compatible with h, i.e., $\sum_{i\in I} \delta_{x_i} = h$. Consider the graph $G := \mathcal{G}(|h|, x, \frac{1}{N}\kappa_N)$ on the vertex set I. There exists $I' \subset I$ such that $x' = (x_i)_{i\in I'}$ is compatible with h'. The subgraph G' of G that is induced by I' can be identified with $\mathcal{G}(|h'|, x', \frac{1}{N}\kappa'_N)$, where κ' denotes the restriction of κ to $I' \times I'$. If G' is connected and for any $i \in I \setminus I'$ there is an edge $\{i, j\} \in E(G)$ with $j \in I'$, then also G is connected. Therefore, using that $1 - x \le \mathrm{e}^{-x}$ for any $x \in \mathbb{R}$, we have

$$p_N(h) \ge p_N(h') \prod_{r\in\mathcal{S}} \Big(1 - \prod_{s\in\mathcal{S}} \Big(1 - \frac{\kappa_N(r,s)}{N}\Big)^{h'_s}\Big)^{h_r - h'_r} \ge p_N(h') \prod_{r\in\mathcal{S}} \big(1 - \mathrm{e}^{-\frac{1}{N}(\kappa h')_r}\big)^{h_r - h'_r}.$$

□

In the following lemmas we give the lower bound for the connection probabilities in the macroscopic setting. It is sufficient to restrict to the case $y \ll \kappa y$, since otherwise the limit is already ensured to be $-\infty$ by Lemma 25.

Lemma 27 (Lower bound in (3.4) for irreducible κ)**.** *Fix $y \in (0,1]^{\mathcal{S}}$ satisfying $y \ll \kappa y$ and assume that κ is irreducible with respect to y. Let $\{k^{(N)}\}_{N\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0^{\mathcal{S}}$ such that $\lim_{N\to\infty} \frac{k_r^{(N)}}{N} = y_r$ for all $r \in \mathcal{S}$ and assume that $\tau(k^{(N)}) > 0$ holds for all $N \in \mathbb{N}$. Then*

$$\liminf_{N\to\infty} \frac{1}{N} \log p_N(k^{(N)}) \ge \sum_{r\in\mathcal{S}} y_r \log\big(1 - \mathrm{e}^{-\kappa y_r}\big). \tag{3.29}$$

Proof. For $\delta \ge 0$ satisfying $2\delta < \inf\{y_s \colon s \in \mathcal{S}\}$ define $y^{(\delta)} := y - \delta$ as well as

$$\nu_s^{(\delta)} := \frac{y_s^{(\delta)}}{1 - \mathrm{e}^{-(\kappa y^{(\delta)})_s}}, \quad \text{for } s \in \mathcal{S}. \tag{3.30}$$

For purely technical reasons that will become apparent later, we cannot work directly with $\nu := \nu^{(0)}$. Note that for all $\delta \ge 0$ our construction ensures that with respect to $\nu^{(\delta)}$ the characteristic equation (3.5) has a non-trivial solution, which is equivalent to $\Sigma(\kappa, \nu^{(\delta)}) > 1$ by Lemma 16. Let $m^{(N,\delta)} \in \mathbb{N}_0^{\mathcal{S}}$ be such that $m_r^{(N,\delta)} = \lfloor (k^{(N)} - N\delta)(1 - \mathrm{e}^{-\frac{1}{N}(\kappa(k - N\delta))_r})^{-1} \rfloor$. In particular, $\lim_{N\to\infty} \frac{m^{(N,\delta)}}{N} = \nu^{(\delta)}$. Clearly, we have that $\nu^{(\delta)} \to \nu$ as $\delta \to 0$. Note that $y \le \nu$. Moreover, there exist $\delta^* > 0$ and N_0 such that $m^{(N,\delta)} \ge k^{(N)}$ for all $\delta \le \delta^*$ and $N \ge N_0$, which we will assume from now on. For brevity we will write k and $m^{(\delta)}$ instead of $k^{(N)}$ and $m^{(N,\delta)}$.

Fix $r \in \mathcal{S}$ with $y_r > 0$ and note that by the assumption $y \ll \kappa y$ this implies $(\kappa y)_r > 0$ and thus $y_r < \nu_r$. From here on, h will always denote an element in the set $\{h \in \mathbb{N}_0^{\mathcal{S}} \colon \mathbf{e}_r \le h \le m^{(\delta)}\}$. For such h abbreviate

$$a_N^{(\delta)}(h) := \Big[\prod_{s\in\mathcal{S}} \binom{m_s^{(\delta)} - \delta_{r,s}}{h_s - \delta_{r,s}}\Big] p_N(h) \prod_{s,\tilde{s}\in\mathcal{S}} \Big(1 - \frac{\kappa_N(s,\tilde{s})}{N}\Big)^{h_s(m_{\tilde{s}}^{(\delta)} - h_{\tilde{s}})}. \tag{3.31}$$

Recall from Lemma 24 that the sum of $a_N^{(\delta)}(h)$ over the mentioned h's is equal to one. In the following, we will split this sum into the three parts where $|h| \le R$, and $R < |h| \le \varepsilon N$ and $|h| > \varepsilon N$ for some large $R \in \mathbb{N}$ and some small $\varepsilon > 0$. We will show that the first part converges towards something in $(0,1)$, the second part vanishes, and we will identify the exponential rate of the other one explicitly via some Laplace approximation. Explicitly, we will prove the following three claims:

Claim 1: $$\limsup_{\delta\downarrow 0} \limsup_{R\to\infty} \limsup_{N\to\infty} \sum_{h\colon |h|\le R} a_N^{(\delta)}(h) \le 1 - \frac{y_r}{\nu_r} \in [0,1). \tag{3.32}$$

Claim 2: If ε is small enough, for some sufficiently small $\delta_0 > 0$ and some sufficiently large $N_0 \in \mathbb{N}$,

$$\limsup_{R\to\infty} \sup_{N\ge N_0} \sup_{\delta\in(0,\delta_0]} \sum_{h\colon R<|h|\le\varepsilon N} a_N^{(\delta)}(h) = 0. \tag{3.33}$$

Claim 3: For any $\delta > 0$,

$$\sum_{h\colon |h|>\varepsilon N} a_N^{(\delta)}(h) \le \mathrm{e}^{o(N)} \mathrm{e}^{NC(\delta)} \mathrm{e}^{-N\widetilde{f}(y,\nu)} p_N(k) + \varepsilon_N, \tag{3.34}$$

for some sequence $(\varepsilon_N)_N = (\varepsilon_N(\delta))_N$ that converges to 0, where $\widetilde{f}(y,\nu) = \langle y, \log(1 - \mathrm{e}^{-\kappa y})\rangle$, and $C(\delta)$ is a constant only depending on δ and vanishing as $\delta \downarrow 0$.

Let us first explain how the assertion of the lemma follows from these three claims. We start by using Lemma 24 and Claim 3 to obtain, in the limit as $N \to \infty$,

$$1 - \sum_{h\colon |h|\le R} a_N^{(\delta)}(h) - \sum_{h\colon R<|h|<\varepsilon N} a_N^{(\delta)}(h) \le \mathrm{e}^{o(N)} \mathrm{e}^{NC(\delta)} \mathrm{e}^{-N\widetilde{f}(y,\nu)} p_N(k) + \varepsilon_N.$$

By (3.32) and (3.33), the left hand side is not smaller than y_r/ν_r, when taking the limits as $N \to \infty$, followed by $R \to \infty$ and $\delta \downarrow 0$, if ε is small enough. In particular, the left-hand side is bounded away from zero when taking these limits. Hence, the exponential rate of the right-hand side as $N \to \infty$ is nonnegative. This implies that $\liminf_{N\to\infty} \frac{1}{N} \log p_n(k) \ge \widetilde{f}(y,\nu)$, which is the assertion of the lemma. It remains to prove the three claims.

Proof of Claim 3: As we explained in the proof of Lemma 25, the exponential rate of the two products in the definition of $a_N^{(\delta)}(h)$ can easily be identified with the help of Stirling's formula and the exponential limit theorem $\lim_{n\to\infty}(1 + \frac{c}{n})^n = \mathrm{e}^c$ and elementary calculations. Indeed, for any sequence $h^{(N)}$ such that $\lim_{N\to\infty} \frac{h^{(N)}}{N} = x$ exists and $x \in [0,1]^{\mathcal{S}} \setminus \{0\}$, and for any $\delta > 0$,

$$a_N^{(\delta)}(h^{(N)}) = \mathrm{e}^{o(N)} \mathrm{e}^{-Nf(x,\nu^{(\delta)})} p_N(h^{(N)}), \qquad N \to \infty, \tag{3.35}$$

where we introduce $f(x,\tilde{x}) = \langle x, \log \frac{x}{\tilde{x}}\rangle + \langle \tilde{x} - x, \log \frac{\tilde{x}-x}{\tilde{x}}\rangle + \langle \tilde{x} - x, \kappa x\rangle$ for $x, \tilde{x} \in [0,1]^{\mathcal{S}} \setminus \{0\}$ and $x \le \tilde{x}$. We can write

$$\begin{aligned} f(x,\tilde{x}) &= \sum_{s\in\mathcal{S}} \tilde{x}_s \Big[\frac{x_s}{\tilde{x}_s} \log \frac{\frac{x_s}{\tilde{x}_s}}{1 - \mathrm{e}^{-(\kappa x)_s}} + \Big(1 - \frac{x_s}{\tilde{x}_s}\Big) \log \frac{1 - \frac{x_s}{\tilde{x}_s}}{\mathrm{e}^{-(\kappa x)_s}}\Big] + \sum_{s\in\mathcal{S}} x_s \log(1 - \mathrm{e}^{-(\kappa x)_s}) \\ &= \langle \tilde{x}, H(x;\tilde{x})\rangle + \langle x, \log(1 - \mathrm{e}^{-\kappa x})\rangle, \end{aligned} \tag{3.36}$$

where we write $H = (H_s)_{s\in\mathcal{S}}$ and for $s \in \mathcal{S}$ we write $H_s(x;\tilde{x})$ for the entropy of the Bernoulli distribution with parameter $x_s/\tilde{x}_s$ with respect to the one with parameter $1-\mathrm{e}^{-(\kappa x)_s}$.

The second term in the second line of (3.36) will be handled jointly with the exponential rate of $p_N(h^{(N)})$, so let us discuss here the minimum of the first. Since every component of H is an entropy between probability measures, we have that $H(x,\nu^{(\delta)}) \geq 0$ pointwise for all $x \leq \nu^{(\delta)}$ with equality if and only if $\frac{x}{\nu^{(\delta)}} = 1-\mathrm{e}^{-\kappa x}$. As κ is irreducible with respect to $\nu^{(\delta)}$ the latter condition is true if and only if $x = y^{(\delta)}$ or $x = 0$, see Lemma 16. Hence, we have seen that, for ε sufficiently small, $\min_{x\colon |x|\geq\varepsilon}\langle\nu^{(\delta)}, H(x;\nu^{(\delta)})\rangle = \langle\nu^{(\delta)}, H(y^{(\delta)};\nu^{(\delta)})\rangle$.

Let us now prove (3.34). The work we still have to do is to combine (3.35) with estimates for the p_N-term from Lemmas 25 and 26, and we have to distinguish the cases that h is left of k but close to k (then Lemma 26 applies), or bounded away from k (then Lemma 25 suffices). Let $h^{(N)}$ be such that $|h^{(N)}| > \varepsilon N$ and assume that $x := \lim_{N\to\infty}\frac{h^{(N)}}{N}$ exists. Note that $x \in [0,1]\setminus\{0\}$. We first examine the case where $h^{(N)} \in [k-2N\delta, k]$ for large N and hence $|x-y| \leq 2\delta$. With the help of (3.35) and using the estimate from Lemma 26 we get that

$$\begin{aligned} a_N^{(\delta)}(h^{(N)}) &\leq \mathrm{e}^{o(N)}\mathrm{e}^{N(f(y,\nu)-f(x,\nu^{(\delta)})-\langle y-x,\log(1-\mathrm{e}^{-\kappa x})\rangle)}\mathrm{e}^{-Nf(y,\nu)}p_N(k) \\ &\leq \mathrm{e}^{o(N)}\mathrm{e}^{NC(\delta)}\mathrm{e}^{-Nf(y,\nu)}p_N(k) \end{aligned} \tag{3.37}$$

where

$$C(\delta) = \sup_{x\colon |x-y|\leq 2\delta} |f(y,\nu) - f(x,\nu^{(\delta)}) - \langle y-x, \log(1-\mathrm{e}^{-\kappa x})\rangle| \tag{3.38}$$

and it can be easily verified that $C(\delta) \to 0$ as $\delta \to 0$. Now, let us examine the case $h^{(N)} \notin [k-2N\delta, k]$ for large N, i.e., $|x-y^{(\delta)}| \geq \delta$. We start with (3.35) and use the estimate (3.27) from Lemma 25 to get

$$a_N^{(\delta)}(h^{(N)}) \leq \mathrm{e}^{o(N)}\mathrm{e}^{-N\langle\nu^{(\delta)},H(x,\nu^{(\delta)})\rangle} \leq \mathrm{e}^{o(N)}\mathrm{e}^{-N\widetilde{C}(\delta)} \tag{3.39}$$

where the number

$$\widetilde{C}(\delta) := \inf_{x\colon |x|>\varepsilon, |x-y^{(\delta)}|\geq\delta} \langle\nu^{(\delta)}, H(x,\nu^{(\delta)})\rangle \tag{3.40}$$

is strictly positive since the function $x \mapsto H(x,\nu^{(\delta)})$ is continuous and its only zeros are at $x=0$ and $x = y^{(\delta)}$. Now, it is not hard to see that $\#\{h \in \mathbb{N}_0^{\mathcal{S}}\colon h \leq m^{(\delta)}\} \leq (\nu N)^{|\mathcal{S}|} = \mathrm{e}^{o(N)}$, hence a simple Laplace approximation argument implies that (3.34) holds with $\varepsilon_N := \mathrm{e}^{o(N)}\mathrm{e}^{-N\widetilde{C}(\delta)}$, which vanishes as $N\to\infty$ since $\widetilde{C}(\delta) > 0$.

Proof of Claim 1: Applying Lemma 23 to $p_N(h)$ we get,for any $N, R \in \mathbb{N}$ and $\delta > 0$,

$$\begin{aligned} &\sum_{h\colon |h|\leq R} a_N^{(\delta)}(h) \\ &\leq \frac{N}{m_r^{(\delta)}} \sum_{h\colon |h|\leq R} \tau_N(h) h_r \prod_{s\in\mathcal{S}} \frac{\frac{m_s^{(\delta)}}{N}\frac{m_s^{(\delta)}-1}{N}\cdots\frac{m_s^{(\delta)}-h_s+1}{N}\left(\prod_{\tilde{s}\in\mathcal{S}}(1-\frac{\kappa_N(s,\tilde{s})}{N})^{m_{\tilde{s}}^{(\delta)}-h_{\tilde{s}}}\right)^{h_s}}{h_s!} \\ &\leq \frac{N}{m_r^{(\delta)}} \sum_{h\colon |h|\leq R} \tau_N(h) h_r \prod_{s\in\mathcal{S}} \left[\frac{1}{h_s!}\left(\frac{m_s^{(\delta)}}{N}\mathrm{e}^{-\kappa_N(\frac{m^{(\delta)}-h}{N})_s}\right)^{h_s}\right], \end{aligned} \tag{3.41}$$

where we used that $1-x \leq \mathrm{e}^{-x}$ for any $x \in \mathbb{R}$ in the second line. Hence,

$$\limsup_{N\to\infty} \sum_{h\colon |h|\leq R} a_N^{(\delta)}(h) \leq \frac{1}{\nu_r^{(\delta)}} \sum_{h\colon |h|\leq R} \tau(h) h_r \prod_{s\in\mathcal{S}} \left[\frac{1}{h_s!}\left(\nu_s^{(\delta)}\mathrm{e}^{-(\kappa\nu^{(\delta)})_s}\right)^{h_s}\right] =: \Lambda_R^{(\delta)} \tag{3.42}$$

Observe that by Proposition 17 and by the definition of $\nu^{(\delta)}$ given in (3.30), we have that $\lim_{R\to\infty}\Lambda_R^{(\delta)} = \frac{\nu_r^{(\delta)}-y_r^{(\delta)}}{\nu_r^{(\delta)}} \in [0,1)$ and clearly $\lim_{\delta\to 0}\lim_{R\to\infty}\Lambda_R^{(\delta)} = \frac{\nu_r - y_r}{\nu_r} \in (0,1)$.

Proof of Claim 2: For the sum on h satisfying $R < |h| \le \varepsilon N$, we can use basically the same estimates as in (3.41), but we modify the second line by now estimating

$$\prod_{s,\tilde{s}\in\mathcal{S}} \big(1 - \tfrac{\kappa_N(s,\tilde{s})}{N}\big)^{(m_{\tilde{s}}^{(\delta)} - h_{\tilde{s}})h_s} \le \Big(\prod_{s\in\mathcal{S}} \mathrm{e}^{-\frac{1}{N}(\kappa_N m^{(\delta)})_s h_s}\Big)\mathrm{e}^{\frac{1}{N}\|\kappa_N\|_\infty |h|^2}, \tag{3.43}$$

since we do not have that $m^{(\delta)} \ge h$ on this sum. This gives

$$\sum_{h\colon R<|h|\le\varepsilon N} a_N^{(\delta)}(h) \le \frac{N}{m_r^{(\delta)}} \sum_{n=R+1}^{\varepsilon N} \mathrm{e}^{n\varepsilon\|\kappa_N\|_\infty} \sum_{h\colon |h|=n} \tau_N(h) h_r \prod_{s\in\mathcal{S}} \Big[\frac{1}{h_s!}\Big(\frac{m_s^{(\delta)}}{N}\mathrm{e}^{-(\frac{1}{N}\kappa_N m^{(\delta)})_s}\Big)^{h_s}\Big]. \tag{3.44}$$

This is the main part of a power series with coefficients τ, which we studied in Lemma 20. Note that the term in round brackets converges to $\nu_s \mathrm{e}^{-(\kappa\nu)_s}$ as $N\to\infty$, followed by $\delta \downarrow 0$. Furthermore, this vector $\theta(\kappa,\nu) = \nu\mathrm{e}^{-\kappa\nu}$ satisfies $\chi(\kappa,\theta(\kappa,\nu)) > 1$, since $\Sigma(\kappa,\nu) > 1$ by construction (see Lemma 20(3)). It will turn out that this is sufficient to estimate the sum on n against some convergent geometric series.

We pick a small threshold $\varepsilon' > 0$. We may pick δ_0 so small and then $N_0 \in \mathbb{N}$ so large that

$$\frac{m_s^{(\delta)}}{N}\mathrm{e}^{-(\frac{1}{N}\kappa_N m^{(\delta)})_s} \le \nu_s \mathrm{e}^{-(\kappa\nu)_s}\mathrm{e}^{\varepsilon'}, \qquad \delta\in(0,\delta_0], N\ge N_0, s\in\mathcal{S}.$$

Additionally, we can also assume that $N/m_r^{(\delta)} \le 1 + 1/\nu_r$ for these δ and N. By τ we denote the function defined as in (3.2) for the matrix κ, then, since $\lim_{N\to\infty}\kappa_N = \kappa$, we may also pick N_0 so large that additionally

$$\tau_N(h) \le \tau(h)\mathrm{e}^{\varepsilon'|h|}, \qquad h\in\mathbb{N}_0^{\mathcal{S}}, N\ge N_0.$$

We also may and will assume that $\|\kappa_N\|_\infty \le 1 + \|\kappa\|_\infty$ for $N \ge N_0$. Therefore, we can now estimate, for $\delta\in(0,\delta_0]$ and $N\ge N_0$, and any $R\in\mathbb{N}$,

$$\sum_{h\colon R<|h|\le\varepsilon N} a_N^{(\delta)}(h) \le \big(1+\tfrac{1}{\nu_r}\big) \sum_{n=R+1}^{\infty} \mathrm{e}^{n\varepsilon(1+\|\kappa\|_\infty)}\mathrm{e}^{2\varepsilon' n} \sum_{h\colon |h|=n} \tau(h) h_r \prod_{s\in\mathcal{S}} \Big[\frac{1}{h_s!}\big(\nu_s\mathrm{e}^{-(\kappa\nu)_s}\big)^{h_s}\Big].$$

Now we can apply (3.16) in Lemma 20 to estimate the sum on h against $\mathrm{e}^{o(n)}\mathrm{e}^{-n[\chi(\kappa,\theta(\kappa,\nu))-1]}$. We recall that $\chi(\kappa,\theta(\kappa,\nu)) - 1 > 0$ and pick now ε and ε' so small that the sum on n can be estimated against a convergent geometric series. Hence, the entire sum vanishes as $R\to\infty$, which ends the proof. □

In the following we want to verify the statement of Lemma 27 without assuming that κ is irreducible with respect to y but still under the assumption that $\tau(k^{(N)}) > 0$. What we have in mind are components $k^{(N)}$ where for a certain set of types $\tilde{S}\subset\mathcal{S}$ only $o(N)$ vertices of those types are available, i.e., $k_s^{(N)} \in o(N)$ for all $s\in\tilde{S}$. However, these types might be necessary in order to connect the vertices with types in $\mathcal{S}\setminus\tilde{S}$. For that case we prove in the following that on the exponential scale the asymptotics of the connection probability $p_N(k^{(N)})$ are the same as in the previous lemma.

Lemma 28 (Lower bound in (3.4) under generalized assumptions)**.** *Fix $y\in[0,1]^{\mathcal{S}}\setminus\{0\}$ satisfying $y \ll \kappa y$. Let $\{k^{(N)}\}_{N\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0^{\mathcal{S}}$ such that $\lim_{N\to\infty}\frac{k_r^{(N)}}{N} = y_r$ for all $r\in\mathcal{S}$ and*

assume that $\{k_s^{(N)}\}_N$ *is bounded for all* $s \notin \mathrm{supp}(y)$*. Assume further that* $\tau(k^{(N)}) > 0$ *holds for all but finitely many* $N \in \mathbb{N}$*. Then*

$$\liminf_{N\to\infty} \frac{1}{N} \log p_N(k^{(N)}) \geq \sum_{r\in\mathcal{S}} y_r \log\left(1 - \mathrm{e}^{-\kappa y_r}\right). \tag{3.45}$$

Proof. Throughout the proof we will assume without loss of generality that $\mathcal{S} = \bigcap_N \mathrm{supp}(k^{(N)})$. If $y > 0$ on $\mathcal{S}$ and κ is irreducible with respect to y, then we are in the setting of Lemma 27, which yields the claim. In the other case, it always holds that $\tilde{S} := \mathcal{S} \setminus \mathrm{supp}(y) \neq \emptyset$. We decompose $\mathrm{supp}(y)$ into disjoint sets S_j, $j \in J$, satisfying that the restriction of κ to $S_j \times S_j$ is irreducible with respect to y restricted to S_j.

We will start with the special case where $k_s^{(N)} = 1$ for all $s \in \tilde{S}$ and $N \in \mathbb{N}$. It is easily seen that the assumption $\tau(k^{(N)}) > 0$ for all but finitely many $N \in \mathbb{N}$ implies that there exists a tree T on the type set $\mathcal{S}$ such that $\prod_{\{r,s\}\in E(T)} \kappa(r,s) > 0$. Define

$$\tilde{E} := \{\{r,s\} \in E(T) \colon r \in \tilde{S} \text{ or } s \in \tilde{S}\}. \tag{3.46}$$

For every $r \in \mathcal{S}$ we now fix a vertex $i_r \in [|k^{(N)}|]$ that is of type r. Note that $\{i_r \colon r \in \tilde{S}\}$ already contains all the vertices with types in $\tilde{S}$. We will abbreviate $\mathcal{G}_N = \mathcal{G}(|k^{(N)}|, x, \frac{1}{N}\kappa_N)$. Note that the event $\{\{i_r, i_s\} \in E(\mathcal{G}_N) \colon \{r,s\} \in \tilde{E}\}$ is independent of the existence of edges $\{i, i'\}$, if the types of i and i' are in S_j for any $j \in J$. Therefore,

$$\begin{aligned} p_N(k^{(N)}) &\geq \Big[\prod_{j\in J} p_N(k^{(N)}\mathbb{1}_{S_j})\Big] \times \prod_{\{r,s\}\in\tilde{E}} \mathbb{P}(\{i_r, i_s\} \in E(\mathcal{G}_N)) \\ &\geq \Big[\prod_{j\in J} p_N(k^{(N)}\mathbb{1}_{S_j})\Big] \times \prod_{\{r,s\}\in\tilde{E}} \frac{1}{N}\kappa_N(r,s). \end{aligned}$$

Clearly the second product is $\geq \mathrm{e}^{o(N)}$, since κ_N is bounded away from zero, uniformly in all large N. Applying Lemma 27 to $k^{(N)}\mathbb{1}_{S_j}$ for every $j \in J$, we get that

$$\frac{1}{N}\log p_N(k^{(N)}) \geq o(1) + \sum_{j\in J}\sum_{r\in S_j} y_r \log(1 - \mathrm{e}^{-(\kappa(y\mathbb{1}_{S_j}))_r})) = o(1) + \sum_{r\in\mathcal{S}} y_r \log(1 - \mathrm{e}^{-(\kappa y)_r}), \tag{3.47}$$

where in the last equation we used that with fixed $j \in J$, for any $r \in S_j$ and $s \in \mathcal{S} \setminus S_j$ we have either $\kappa(r,s) = 0$ or $y_s = 0$. Thus, for $r \in S_j$ we get

$$(\kappa(y\mathbb{1}_{S_j}))_r = \sum_{s\in S_j} \kappa(r,s) y_s = \sum_{s\in\mathcal{S}} \kappa(r,s) y_s = (\kappa y)_r. \tag{3.48}$$

This implies that equation (3.45) holds in the special case.

Now, consider the general case, where $\{k_s^{(N)}\}_N$ is bounded in N for all $s \in \tilde{S}$. Then by Lemma 26

$$p_N(k^{(N)}) \geq p_N(k^{(N)}\mathbb{1}_{\mathcal{S}\setminus\tilde{S}} + \mathbb{1}_{\tilde{S}}) \times \prod_{r\in\tilde{S}} (1 - \mathrm{e}^{-(\kappa y)_r})^{k_r^{(N)}-1}.$$

Observe that the last product is $\mathrm{e}^{o(N)}$ by assumption, hence the claim is implied by the treatment of the first case.

□

4 The large-deviations principle for a finite type set

In this section we still assume that $\mathcal{S}$ is a finite set and derive the large-deviations principle (LDP) of Theorem 1 for this case, Theorem 29. This is not only an important special case that is worth being formulated and studied on its own, but it will also serve as an important input in the proof of Theorem 1 in Section 5. The formulation in the discrete case is notationally pretty different from the formulation in the general setting and many objects simplify because of the finiteness of $\mathcal{S}$. Therefore we are going to formulate the setting and the LDP from scratch in Section 4.1.

The proof will rely on the study of the connection probabilities that was done in Section 3.2. The organization of this section is as follows. In Section 4.2 we identify the distribution of Mi_N. We will decompose this distribution into a micro–, meso– and macroscopic part whose exponential rates we derive in Section 4.3. The proof of the LDP of Theorem 29 is finally finished in Section 4.4.

4.1 Formulation of the LDP

Fix a probability measure μ on $\mathcal{S}$, which we will denote as a vector $\mu = (\mu_s)_{s\in\mathcal{S}}$. For any $N \in \mathbb{N}$ let $\mathbf{x}^{(N)} = (x_1^{(N)}, \dots, x_N^{(N)}) \in \mathcal{S}^N$ be a type vector such that the normalized empirical measure, $\mu^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{x_i^{(N)}}$, converges to μ as $N \to \infty$.

We denote by $\{\mathcal{C}_j\}_j$ the collection of the vertex sets of all the connected components of $\mathcal{G}_N = \mathcal{G}(N, \mathbf{x}^{(N)}, \frac{1}{N}\kappa_N)$. We want to study empirical measures depending on the random collection $\{\mathcal{C}_j\}_j$. For this we introduce the type-registring mapping $\eta\colon \mathcal{P}([N]) \to \mathbb{N}_0^{\mathcal{S}}$ that gives the (type) composition of an arbitrary vertex set $A \subset [N]$, i.e., $\eta(A) = (\eta_s(A))_{s\in\mathcal{S}}$, and $\eta_s(A) = \#\{i \in A\colon x_i^{(N)} = s\}$ is the number of vertices in A with type s. Note that the mapping η depends on the entire type vector $\mathbf{x}^{(N)}$, not only on its normalized empirical measure, $\mu^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{x_i^{(N)}}$.

We identify $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ with $\mathbb{N}_0^{\mathcal{S}}$ and will work in $[0,1]^{\mathcal{S}}$ instead of $\mathcal{M}(\mathcal{S})$. Now we recall the definition of the main objects, the empirical measures of the connected components of $\mathcal{G}_N$, in microscopic, respectively macroscopic, registration. The *microscopic empirical measure* Mi_N is defined as a measure on $\mathbb{N}_0^{\mathcal{S}}$ via

$$\mathrm{Mi}_N(\mathrm{d}k) := \frac{1}{N}\sum_j \delta_{\eta(\mathcal{C}_j)}(\mathrm{d}k). \tag{4.1}$$

Since $\mathbb{N}_0^{\mathcal{S}}$ is a discrete space, we will abbreviate $\mathrm{Mi}_N(k) = \mathrm{Mi}_N(\{k\})$ for any $k \in \mathbb{N}_0^{\mathcal{S}}$. The *macroscopic empirical measure* is defined as a measure on $[0,1]^{\mathcal{S}} \setminus \{0\}$ via

$$\mathrm{Ma}_N(\mathrm{d}y) := \sum_j \delta_{\frac{1}{N}\eta(\mathcal{C}_j)}(\mathrm{d}y). \tag{4.2}$$

Therefore our state spaces for Mi_N and Ma_N are now

$$\mathcal{L} := \big\{\lambda = (\lambda_k)_{k\in\mathbb{N}_0^{\mathcal{S}}} \in [0,\infty)^{\mathbb{N}_0^{\mathcal{S}}} : |c(\lambda)| \le 1\big\} \tag{4.3}$$

and

$$\mathcal{A} := \big\{\alpha \in \mathcal{M}_{\mathbb{N}_0}([0,1]^{\mathcal{S}} \setminus \{0\})\colon |c(\alpha)| \le 1\big\}, \tag{4.4}$$

respectively, where

$$c_r(\lambda) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k k_r \qquad \text{and} \qquad c_r(\alpha) = \int_{[0,1]^{\mathcal{S}}\setminus\{0\}} \alpha(\mathrm{d}y)\, y_r, \quad r \in \mathcal{S} \tag{4.5}$$

and $|c| = \sum_s c_s$ for any $(c_s)_{s\in\mathcal{S}}$. One can easily verify that for any fixed N we have $c_r(\mathrm{Mi}_N) = \mu_r^{(N)}$ as well as $c_r(\mathrm{Ma}_N) = \mu_r^{(N)}$ for any $r \in \mathcal{S}$, so indeed $\mathrm{Mi}_N \in \mathcal{L}$ and $\mathrm{Ma}_N \in \mathcal{A}$. However, the idea is that due to the topologies that we choose, some of the (rescaled) vertex mass specified by $\mu^{(N)}$ will get lost when we take the limit for $N \to \infty$.

We equip $\mathcal{L}$ and $\mathcal{A}$ with the vague topologies that we introduced in Section 1.2. On $\mathcal{L}$, this is identical with the topology of pointwise convergence (i.e., $\lim_{N\to\infty} \lambda^{(N)} = \lambda$ if and only if $\lim_{N\to\infty} \lambda_k^{(N)} = \lambda_k$ for any $k \in \mathbb{N}_0^{\mathcal{S}}$). The vague topology on $\mathcal{A}$ is formulated by saying that $\lim_{N\to\infty} \alpha^{(N)} = \alpha$ if and only if $\lim_{N\to\infty} \int \alpha^{(N)}(\mathrm{d}y)f(y) = \int \alpha(\mathrm{d}y)f(y)$ for any continuous and compactly supported function $f\colon [0,1]^{\mathcal{S}} \setminus \{0\} \to \mathbb{R}$; note that for every such function f there is an $\varepsilon > 0$ such that $f = 0$ on $\{x \in [0,1]^{\mathcal{S}} \colon |x| \le \varepsilon\}$.

Recall the definition of τ given in (3.2). Here is the main result of Section 4:

Theorem 29 (LDP for $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ with finitely many types)**.** *Assume that the empirical measure $\mu^{(N)}$ of the type sequence $(x_1^{(N)}, \dots, x_N^{(N)})$ converges weakly towards a positive probability vector $\mu \in (0,1]^{\mathcal{S}}$ as $N \to \infty$ and that the kernel κ_N converges on $\mathcal{S} \times \mathcal{S}$ towards a μ-irreducible kernel $\kappa \in [0,\infty)^{\mathcal{S}\times\mathcal{S}}$.*

Then $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ satisfies a large deviations principle (LDP) with speed N and rate function $(\lambda,\alpha) \mapsto I(\lambda,\alpha)$ defined by

$$I(\lambda,\alpha) = \begin{cases} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Ma}}(\alpha) + I_{\mathrm{Me}}(\mu - c(\lambda) - c(\alpha)), & \text{if } c(\lambda) + c(\alpha) \le \mu, \\ +\infty & \text{otherwise,} \end{cases}$$

where

$$I_{\mathrm{Mi}}(\lambda) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k \log \frac{\lambda_k}{\tau(k)\prod_{s\in\mathcal{S}} \frac{\mu_s^{k_s}}{k_s!}} + \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k(|k|-1) + \frac{1}{2}\langle c(\lambda), \kappa\mu\rangle, \tag{4.6}$$

$$I_{\mathrm{Ma}}(\alpha) = \int_{\mathcal{M}(\mathcal{S})\setminus\{0\}} \alpha(\mathrm{d}y)\Big(\Big\langle y, \log\frac{y}{(1-\mathrm{e}^{-\kappa y})\mu}\Big\rangle + \frac{1}{2}\langle y, \kappa(\mu - y)\rangle\Big), \tag{4.7}$$

$$I_{\mathrm{Me}}(\nu) = \Big\langle \nu, \log\frac{\nu}{(\kappa\nu)\mu}\Big\rangle + \frac{1}{2}\langle \nu, \kappa\mu\rangle \tag{4.8}$$

and where we always use the convention that $\log 0 = -\infty$ and $0\log 0 = 0$.

Theorem 29 is indeed nothing but the special case of Theorem 1 for a finite set $\mathcal{S}$. Indeed, it is clear that the setting and the two rate functions I_{Ma} and I_{Me} are the discrete-space versions, but (4.6) looks a bit different from the formula for I_{Mi} in Theorem 1. But from substituting the notation of the entropy in (4.6) and noting that the distribution of a Poisson point process with intensity measure μ can here be identified as

$$\mathbb{Q}_\mu(k) = \prod_{r\in\mathcal{S}} \Big[\mathrm{e}^{-\mu_r}\frac{\mu_r^{k_r}}{k_r!}\Big], \qquad k \in \mathbb{N}_0^{\mathcal{S}},$$

one sees that (4.6) is indeed a discrete analog of (1.11). Furthermore, one can also write

$$I_{\mathrm{Mi}}(\lambda) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k \log\frac{\lambda_k}{\lambda_k(\mu)} + \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k(|k|-1) - \frac{1}{2}\langle c(\lambda), \kappa\mu\rangle \tag{4.9}$$

where $\lambda(\mu)$ is defined as the discrete analog of (2.1) with $c = \mu$, i.e.,

$$\lambda_k(\mu) = \tau(k)\prod_{s\in\mathcal{S}} \frac{(\mu_s \mathrm{e}^{-(\kappa\mu)_s})^{k_s}}{k_s!}, \quad k \in \mathbb{N}_0^{\mathcal{S}}. \tag{4.10}$$

This formula will be helpful in Section 6 when we will identify minimizers of I_{Mi}.

We will now give an extension of Theorem 29 for kernels κ that are not μ-irreducible. Recall the notion of connectability for $\alpha \in \mathcal{A}$ that was introduced in Section 2.3.

Theorem 30 (Finite-type LDP for $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ without irreducibility)**.** *For $(\lambda, \alpha) \in \mathcal{L} \times \mathcal{A}$ define*

$$\widetilde{I}(\lambda, \alpha) = \begin{cases} I(\lambda, \alpha), & \text{if } \alpha \text{ is connectable,} \\ +\infty & \text{otherwise.} \end{cases} \tag{4.11}$$

(i) Given all the assumptions from Theorem 29, except the assumption that κ is μ-irreducible, the pair $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ satisfies the lower large-deviations bound (1.14) *with speed N and rate function $\widetilde{I}$.*

(ii) Given all the assumptions from Theorem 29, except the assumption that κ is μ-irreducible and with the additional assumption that $\kappa_N = \kappa$ for all but finitely many $N \in \mathbb{N}$, the pair $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ satisfies an LDP with speed N and rate function $\widetilde{I}$.

The proof is given in Remark 36.

We omit to restate the finite-$\mathcal{S}$ analogs of Theorem 9, Theorem 7 and all the related corollaries, as they can be deduced as special cases. For the critical quantity $\Sigma(\kappa, \mu)$, we refer to (3.6), and we recall that it is in this setting equal to the spectral radius of the matrix $(\kappa(r, s)\mu_s)_{(r,s)\in\mathcal{S}^2}$.

4.2 The distribution of Mi_N

Let us identify the distribution of Mi_N for any N in explicit terms. Note that as long as $N \in \mathbb{N}$ is fixed, the measure Ma_N contains exactly the same information as Mi_N, hence, we are also deriving its distribution. We start by noting that $N\mathrm{Mi}_N$ takes values in

$$\mathcal{L}_N := \Big\{ \ell = (\ell_k)_{k\in\mathbb{N}_0^{\mathcal{S}}} : \ell_k \in \mathbb{N}_0 \text{ for all } k \text{ and } \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \ell_k k_r = N\mu_r^{(N)} \text{ for all } r \in \mathcal{S} \Big\}. \tag{4.12}$$

Lemma 31 (The distribution of Mi_N)**.** *Let $N \in \mathbb{N}$ and assume that $\kappa_N(r, s) \le N$ for all $r, s \in \mathcal{S}$. Then for any $\ell \in \mathcal{L}_N$ we have that*

$$\mathbb{P}\left(N\mathrm{Mi}_N(k) = \ell_k \ \forall k \in \mathbb{N}_0^{\mathcal{S}}\right) = \Big(\prod_{r\in\mathcal{S}} (N\mu_r^{(N)})!\Big) \times \prod_{k\in\mathbb{N}^{\mathcal{S}}} \zeta^{(N)}(\ell, k), \tag{4.13}$$

where

$$\zeta^{(N)}(\ell, k) = \frac{p_N(k)^{\ell_k}}{\ell_k! \prod_{r\in\mathcal{S}} (k_r!)^{\ell_k}} \Big(\prod_{r,s\in\mathcal{S}} \Big(1 - \frac{\kappa_N(r, s)}{N}\Big)^{\frac{1}{2}k_r[N\mu_s^{(N)} - k_s]} \Big)^{\ell_k}, \tag{4.14}$$

and $p_N(k)$ is defined in (3.1).

Proof. This is proved in an analogous way to [AKP19, Corollary 2.2]; we omit the details. □

The formula in (4.13) is easy to understand. Indeed, the combinatorial term on the right (with an additional factor $1/N!$) is equal to the inverse of the number of possible labelings of all the vertices; the event $\{N\mathrm{Mi}_N(k) = \ell_k \, \forall k\}$ means that ℓ_k is equal to the number of clusters in $[N]$ whose

vertex set has the type configuration k, for any multi-index k. The product of $p_N(k)^{\ell_k}$ over k is the probability that all these clusters are connected, the product over the powers of $1-\kappa_N(r,s)/N$ is equal to the probability that each two of them are not connected, and the product of the two remaining combinatorial terms (with an additional factor $N!$) is equal to the number of ways to decompose all the types into clusters having the prescribed vertex structure.

4.3 Exponential rates for micro-, meso- and macro parts

The proof of the LDP in Theorem 29 is carried out in the same way as in [AKP19, Section 3]. The main idea is to split the distribution that we obtained in Lemma 31 into three parts, which we will call micro-, meso- and macroscopic part. These parts roughly give the terms $\mathrm{e}^{-NI_{\mathrm{Mi}}(\lambda)}$, $\mathrm{e}^{-NI_{\mathrm{Me}}(\mu-c(\lambda)-c(\alpha))}$ and $\mathrm{e}^{-NI_{\mathrm{Ma}}(\alpha)}$, if a properly rescaled version of ℓ is close to (λ,α). In the next lemma we give the decomposition into the three parts. Afterwards, we derive the exponential asymptotics of them in Lemmas 33 – 35.

Lemma 32 (Decomposition into three contributions)**.** *Fix $\ell\in\mathcal{L}_N$. For $k\in\mathbb{N}_0^{\mathcal{S}}$ define*

$$z^{(N)}(\ell_k,k):=\frac{p_N(k)^{\ell_k}\prod_{s\in\mathcal{S}}(\frac{N\mu_s}{\mathrm{e}})^{k_s\ell_k}}{\ell_k!\prod_{s\in\mathcal{S}}(k_s!)^{\ell_k}}\prod_{r,s\in\mathcal{S}}\Big(1-\frac{\kappa_N(r,s)}{N}\Big)^{\frac12 k_s(N\mu_r-k_r)\ell_k}, \tag{4.15}$$

and for any two numbers $A,B\in[0,\infty)$ write

$$z^{(N)}_{A,B}(\ell):=\prod_{k\in\mathbb{N}_0^{\mathcal{S}}\,:\,A<|k|\le B} z^{(N)}(\ell_k,k).$$

Then, for any fixed $R\in\mathbb{N}$ and $\varepsilon>0$ we have that, as $N\to\infty$

$$\mathbb{P}_N(N\mathrm{Mi}_N=\ell)=\mathrm{e}^{o(N)}z^{(N)}_{0,R}(\ell)z^{(N)}_{R,\varepsilon N}(\ell)z^{(N)}_{\varepsilon N,N}(\ell).$$

Proof. On the right-hand side of (4.13) we apply Stirling formula $n=\mathrm{e}^{o(n)}(n/\mathrm{e})^n$ to the terms $(N\mu_r^{(N)})!$ and use that $N\mu_r^{(N)}=\sum_k\ell_k k_r$. Note that we also used that $\mu_r^{(N)}\to\mu_r$. □

Lemma 33 (Asymptotics of the micro part)**.** *Fix $R\in\mathbb{N}$ and let $\lambda\in\mathcal{L}$. Define*

$$I^{(R)}_{\mathrm{Mi}}(\lambda):=\sum_{k\in\mathbb{N}_0^{\mathcal{S}}\,:\,|k|\le R}\lambda_k\log\frac{\lambda_k\mathrm{e}^{|k|-1}}{\tau(k)\prod_s\frac{\mu_s^{k_s}}{k_s!}}+\frac12\langle c^{(R)}(\lambda),\kappa\mu\rangle, \tag{4.16}$$

where

$$c^{(R)}_r(\lambda)=\sum_{k\in\mathbb{N}_0^{\mathcal{S}}\,:\,|k|\le R}\lambda_k k_r,\quad r\in\mathcal{S}, \tag{4.17}$$

and where we use the convention that $\log 0=-\infty$ and $0\log 0=0$. In particular, the right-hand side of (4.16) *is equal to $+\infty$ if there is some $k\in\mathbb{N}_0^{\mathcal{S}}$ with $|k|\le R$ such that $\tau(k)=0$, but $\lambda_k>0$. Otherwise, the right-hand side is finite.*

Then for all $\ell^{(N)}=(\ell_k^{(N)})_{k\in\mathbb{N}_0^{\mathcal{S}}}$ satisfying $\ell_k^{(N)}\in\mathbb{N}_0$ and $\lambda_k=\lim_{N\to\infty}\frac1N\ell_k^{(N)}$ for all k with $|k|\le R$ we have

$$\lim_{N\to\infty}\frac1N\log z^{(N)}_{0,R}(\ell^{(N)})=-I^{(R)}_{\mathrm{Mi}}(\lambda). \tag{4.18}$$

Proof. We use Stirling's formula for the terms $\ell_k^{(N)}!$ as well as the fact that for any $k \in \mathbb{N}_0^{\mathcal{S}}$ with $|k| \leq R$ we have that $p_N(k) = \mathrm{e}^{o(1)} N^{1-|k|} \tau_N(k)$ as $N \to \infty$, which is a consequence of the bounds obtained in Lemma 23. Therefore, as $N \to \infty$, we have

$$\begin{aligned}\prod_{k\colon |k|\leq R} \frac{p_N(k)^{\ell_k^{(N)}} \prod_{s\in\mathcal{S}} (\frac{N\mu_s}{\mathrm{e}})^{k_s \ell_k^{(N)}}}{\ell_k^{(N)}! \prod_{s\in\mathcal{S}} (k_s!)^{\ell_k^{(N)}}} &= \mathrm{e}^{o(N)} \mathrm{e}^{-\frac{1}{2}\sum_{k\colon |k|\leq R} \log \ell_k} \prod_{|k|\leq R} \Big(\frac{\tau_N(k) \prod_s \frac{\mu_s^{k_s}}{k_s!}}{(\ell_k^{(N)}/N)\mathrm{e}^{|k|-1}} \Big)^{\ell_k^{(N)}} \\ &= \mathrm{e}^{o(N)} \exp\Big(N \sum_{|k|\leq R} \lambda_k \log \frac{\tau(k) \prod_s \frac{\mu_s^{k_s}}{k_s!}}{\lambda_k \mathrm{e}^{|k|-1}} \Big),\end{aligned} \tag{4.19}$$

where, we have used that $0 \leq \sum_{k\colon |k|\leq R} \log \ell_k \leq |\{k\colon |k| \leq R\}| \log \big(\frac{\sum_{k\colon |k|\leq R} \ell_k}{|\{k\colon |k|\leq R\}|} \big) = o(N)$. Further, we use that $\lim_{N\to\infty} (1 + \frac{x}{N})^N = \mathrm{e}^x$ to get that, as $N \to \infty$,

$$\begin{aligned}&\prod_{|k|\leq R} \prod_{r,s\in\mathcal{S}} \big(1 - \tfrac{\kappa_N(r,s)}{N}\big)^{\frac{1}{2}k_s(N\mu_r - k_r)\ell_k^{(N)}} = \\ &\qquad \exp\Big(-\tfrac{N}{2} \Big\langle \sum_{|k|\leq R} \tfrac{\ell_k^{(N)}}{N} k, \kappa(\mu - \tfrac{k}{N}) \Big\rangle \Big) = \mathrm{e}^{o(N)} \mathrm{e}^{-\frac{N}{2}\langle c^{(R)}(\lambda), \kappa\mu\rangle},\end{aligned} \tag{4.20}$$

where in the last step we used that

$$0 \leq \tfrac{1}{2} \sum_{|k|\leq R} \tfrac{\ell_k^{(N)}}{N} \sum_{r,s} \tfrac{k_r}{N} \kappa_N(r,s) k_s \leq \tfrac{1}{2} \|\kappa_N\|_\infty \tfrac{R}{N} |c^{(R)}(\tfrac{\ell^{(N)}}{N})|,$$

where the right-hand side converges to 0 as $N \to \infty$. Combining the asymptotics from (4.19) and (4.20) gives the claim. □

Lemma 34 (Asymptotics of the macro part)**.** *Fix $\alpha \in \mathcal{A}$, and note that α can be written as $\alpha = \sum_{j\in J} \delta_{y^{(j)}}$ where $y^{(j)} \in [0,1]^{\mathcal{S}} \setminus \{0\}$ for all $j \in J$ and J is a countable set. Fix any $\varepsilon > 0$ with $\varepsilon \notin \{|y^{(j)}|\colon j \in J\}$. Define $J_\varepsilon(\alpha)\colon = \{j \in J\colon |y^{(j)}| > \varepsilon\}$, which is a finite set, and*

$$I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha) := \int \alpha(\mathrm{d}y) \mathbb{1}_{\{|y|>\varepsilon\}} \Big\langle y, \log \frac{y}{(1 - \mathrm{e}^{-\kappa y})\mu} \Big\rangle + \frac{1}{2} \int \alpha(\mathrm{d}y) \mathbb{1}_{\{|y|>\varepsilon\}} \langle y, \kappa(\mu - y)\rangle, \tag{4.21}$$

where we use the convention that $\log 0 = -\infty$ and $0 \log 0 = 0$. In particular, the right-hand side of (4.21) is equal to $+\infty$, if there is some $i \in J_\varepsilon(\alpha)$ such that the condition $y^{(i)} \ll \kappa y^{(i)}$ fails. Then we have the following.

1 *For any sequence $\ell^{(N)} \in \mathcal{L}_N$ denote $\alpha^{(N)} = \sum_k \ell_k^{(N)} \delta_{\frac{k}{N}}$ and assume that $\alpha^{(N)}$ restricted to $\{y\colon |y| > \varepsilon\}$ converges to α restricted to $\{y\colon |y| > \varepsilon\}$, as $N \to \infty$. Then it holds that*

$$\limsup_{N\to\infty} \frac{1}{N} \log z_{\varepsilon N, N}^{(N)}(\ell^{(N)}) \leq -I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha). \tag{4.22}$$

2 *For all $j \in J_\varepsilon(\alpha)$, let $\{k^{(j,N)}\}_{N\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0^{\mathcal{S}}$ such that $\tau(k^{(j,N)}) > 0$ for all $N \in \mathbb{N}$, $\lim_{N\to\infty} \frac{k^{(j,N)}}{N} = y^{(j)}$ and $\{k_s^{(j,N)}\}_{N\in\mathbb{N}}$ is bounded for all $s \notin \mathrm{supp}(y^{(j)})$. Let $\ell^{(N)}$ be an element of $\mathcal{L}_N$ such that $\ell_k^{(N)} = \#\{j \in J\colon k^{(j,N)} = k\}$ for $|k| > \varepsilon N$. Denote $\alpha^{(N)} = \sum_k \ell_k^{(N)} \delta_{\frac{k}{N}}$, then $\alpha^{(N)}$ restricted to $\{y\colon |y| > \varepsilon\}$ converges to α restricted to $\{y\colon |y| > \varepsilon\}$, and*

$$\lim_{N\to\infty} \frac{1}{N} \log z_{\varepsilon N, N}^{(N)}(\ell^{(N)}) = -I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha). \tag{4.23}$$

Proof. We start with the first statement. Let us first turn to the first term on the right of (4.15). We apply Stirling's bound $n! \geq n^n \mathrm{e}^{-n}$ to each of the terms $k_s!$ and the simple bound $\ell_k^{(N)}! \geq 1$ for all $k \in \mathbb{N}_0^{\mathcal{S}}$. Using the upper bound (3.27) for $p_N(k)$ that was derived in Lemma 25 we obtain, as $N \to \infty$,

$$\prod_{k\colon |k|>\varepsilon N} \frac{p_N(k)^{\ell_k^{(N)}} \prod_{s\in\mathcal{S}} (\frac{N\mu_s}{\mathrm{e}})^{k_s \ell_k^{(N)}}}{\ell_k^{(N)}! \prod_{s\in\mathcal{S}} (k_s!)^{\ell_k^{(N)}}} \leq \mathrm{e}^{o(N)} \exp\Big(N \int_{\{|y|>\varepsilon\}} \Big\langle y, \log \frac{(1-\mathrm{e}^{-\kappa y})\mu}{y} \Big\rangle \, \alpha^{(N)}(\mathrm{d}y) \Big), \tag{4.24}$$

where we used that $\sum_{|k|>\varepsilon N} \ell_k^{(N)} \leq \frac{1}{\varepsilon}$ The second term on the right of (4.15) is estimated using $1-x \leq \mathrm{e}^{-x}$ for $x = \kappa_N(r,s)/N$ as follows for $N \to \infty$,

$$\prod_{|k|>\varepsilon N} \prod_{r,s\in\mathcal{S}} \big(1 - \tfrac{\kappa_N(r,s)}{N}\big)^{\frac{1}{2} k_s (N\mu_r - k_r) \ell_k^{(N)}} \leq \mathrm{e}^{o(N)} \exp\Big(-\frac{N}{2} \int_{\{|y|>\varepsilon\}} \langle y, \kappa(\mu - y)\rangle \, \alpha^{(N)}(\mathrm{d}y) \Big). \tag{4.25}$$

Note that the product of the right-hand sides of (4.24) and (4.25) is equal to $\mathrm{e}^{o(N)} \mathrm{e}^{-N I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha^{(N)})}$. Using the convergence assumption on $\alpha^{(N)}$ and the fact that $\varepsilon < \inf_{j\in J_\varepsilon} |y^{(j)}|$ one can verify that $I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha^{(N)}) \to I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha)$ as $N \to \infty$. This gives the result.

To show the second statement, let us first notice that it is clear from the definition that $\alpha^{(N)}$ restricted to $\{y\colon |y| > \varepsilon\}$ converges to α restricted to $\{y\colon |y| > \varepsilon\}$. Therefore we can apply the first assertion and we have the upper bound (4.22). In order to get also the lower bound, we lower estimate $z_{\varepsilon N, N}^{(N)}(\ell^{(N)})$ against the sum on $k^{(N,j)}$ over the finite set $j \in J_\varepsilon(\alpha)$ and note that these are the only summands k with $|k| > \varepsilon N$ such that $\ell_k^{(N)} > 0$. For each such j we apply the asymptotic (3.4) from Theorem 15 and obtain the corresponding lower bound, also noting that $\prod_{|k|>\varepsilon N} (\ell_k^{(N)}!)^{-1} \geq \mathrm{e}^{-\frac{1}{\varepsilon}\log N} = \mathrm{e}^{o(N)}$ and, by estimating $k_s \leq N$, we get that $(k_s)^{-\frac{\ell_k}{2}} \geq \mathrm{e}^{-\frac{1}{2\varepsilon}\log N}$ for any $s \in \mathcal{S}$. To derive the lower bound equivalent of (4.25) we use that $(1-\frac{c}{N})^N = \mathrm{e}^{-c}(1+o(1))$. This proves the claim. □

In Lemma 34(2) we restrict to sequences $\ell^{(N)}$ such that each connected component k with $\ell_k^{(N)} > 0$ is connectable in the sense that $\tau(k) > 0$. We will see in the proof of Theorem 29 that for each $\alpha \in \mathcal{A}$ such that $I_{\mathrm{Ma}}(\alpha) < \infty$ and for each $\varepsilon > 0$, we will always be able to find such a sequence. Our second restriction $\varepsilon \notin \{|y^{(j)}|\colon j \in J\}$ is clearly only technical and gives no problem at all when we later take the limit as $\varepsilon \downarrow 0$; it frees us from unwanted terms.

Lemma 35 (Asymptotics of the meso part)**.** *Fix $R \in \mathbb{N}$ and $\varepsilon > 0$. For any sequence $\ell^{(N)} \in \mathcal{L}_N$ let $\nu_s^{(N)} := \frac{1}{N} \sum_{R<|k|\leq\varepsilon N} \ell_k^{(N)} k_s$ for $s \in \mathcal{S}$. Assume that $\nu_s := \lim_{N\to\infty} \nu_s^{(N)}$ exists (it depends on R and ε). Recall the definition of I_{Me} from* (1.13)*. Then*

$$\limsup_{N\to\infty} \frac{1}{N} \log z_{R,\varepsilon N}^{(N)}(\ell^{(N)}) \leq -I_{\mathrm{Me}}(\nu) + C(R,\varepsilon,\kappa), \tag{4.26}$$

where $C(R,\varepsilon,\kappa)$ is a constant neither depending on $\ell^{(N)}$ nor ν, which vanishes for $R \to \infty$ and $\varepsilon \to 0$.

Proof. We first derive an estimate for the connection probability. For fixed $k \in \mathbb{N}_0^{\mathcal{S}}$ denote $S_k := \mathrm{supp}(k)$. For $p_N(k)$ and $r \in S_k$ we use the upper bound in (3.24) from Lemma 23 and also the formula for $\tau(k)k_r$ given in (3.10) to obtain

$$p_N(k) \leq N^{1-|k|}(1+o(1))\tau(k) = N^{1-|k|}(1+o(1)) \frac{1}{k_r} \Big(\prod_{s\in S_k} (\kappa k)_s^{k_s - 1} \Big) \Delta_r(k), \tag{4.27}$$

where $\Delta_r(k)$ is defined in (3.11). Note that

$$\Delta_r(k) = \sum_{A\in\overrightarrow{\mathcal{T}}_r(S_k)} \prod_{\{s,s'\}\in E(A)} \kappa(s,s')k_{s'} \le \|\kappa\|_\infty^{|S_k|-1}|\mathcal{T}_r(S_k)| \prod_{s\in S_k\setminus\{r\}} k_s \le C \prod_{s\in S_k\setminus\{r\}} k_s, \quad (4.28)$$

where $C = (\|\kappa\|_\infty|\mathcal{S}|)^{|\mathcal{S}|-1}$. Next, we use both estimates together with the Stirling bound $n! \ge n^n\mathrm{e}^{-n}$, which we apply to $\ell_k^{(N)}!$ as well as the terms $k_s!$ for all $s \in \mathcal{S}$. This gives that

$$\begin{aligned}\prod_{R<|k|\le\varepsilon N} \frac{p_N(k)^{\ell_k^{(N)}} \prod_{s\in\mathcal{S}}(\frac{N\mu_s}{\mathrm{e}})^{k_s\ell_k^{(N)}}}{\ell_k^{(N)}!\prod_{s\in\mathcal{S}}(k_s!)^{\ell_k^{(N)}}} &\le \prod_{R<|k|\le\varepsilon N}\Big(\frac{N\Delta_r(k)\prod_{s\in S_k}(\kappa k)_s^{k_s-1}}{k_r\ell_k^{(N)}\mathrm{e}^{-1}\prod_{s\in S_k}\big(\frac{k_s}{\mu_s}\big)^{k_s}}\Big)^{\ell_k^{(N)}} \\ &\le \mathrm{e}^{N\langle\nu^{(N)},\log\mu\rangle} \times \prod_{R<|k|\le\varepsilon N}\prod_{s\in S_k}\Big(\frac{(\kappa k)_s}{k_s}\Big)^{(k_s-1)\ell_k^{(N)}} \times \prod_{R<|k|\le\varepsilon N}\Big(\frac{N\mathrm{e}C}{\ell_k^{(N)}(k_r)^2}\Big)^{\ell_k^{(N)}}.\end{aligned} \quad (4.29)$$

The first term on the right-hand side of (4.29) is clearly equal to $\mathrm{e}^{o(N)}\mathrm{e}^{N\langle\nu,\log\mu\rangle}$. We now take a look at the second one. For $s \in \mathcal{S}$ we put

$$C_s(\ell^{(N)}) := \sum_{k:\, R<|k|\le\varepsilon N,\, k_s>0} \ell_k^{(N)}(k_s-1) = \nu_s^{(N)} - \sum_{k:\, R<|k|\le\varepsilon N,\, k_s>0} \ell_k^{(N)}. \quad (4.30)$$

Then $C_s(\ell^{(N)}) \in [\nu_s^{(N)} - \frac{1}{R}, \nu_s^{(N)}]$ since $\sum_{k:\, R<|k|\le\varepsilon N}\ell_k^{(N)} \le \frac{1}{R}\sum_{k:\, R<|k|}|k|\ell_k^{(N)} \le N/R$. Next, we apply Jensen's inequality and the fact that $x \mapsto \log x$ is concave to get that

$$\begin{aligned}\prod_{R<|k|\le\varepsilon N}\prod_{s\in S_k}\Big(\frac{(\kappa k)_s}{k_s}\Big)^{(k_s-1)\ell_k^{(N)}} &= \exp\Big(\sum_{s\in\mathcal{S}} C_s(\ell^{(N)}) \sum_{R<|k|\le\varepsilon N,\, k_s>0} \frac{\ell_k^{(N)}(k_s-1)}{C_s(\ell^{(N)})}\log\frac{(\kappa k)_s}{k_s}\Big) \\ &\le \exp\Big(\sum_{s\in\mathcal{S}} C_s(\ell^{(N)})\log\Big(\sum_{R<|k|\le\varepsilon N,\, k_s>0}\frac{\ell_k^{(N)}(k_s-1)}{C_s(\ell^{(N)})}\,\frac{(\kappa k)_s}{k_s}\Big)\Big) \\ &\le \exp\Big(N\sum_{s\in\mathcal{S}}\tfrac{1}{N}C_s(\ell^{(N)})\log\frac{(\kappa\nu^{(N)})_s}{\frac{1}{N}C_s(\ell^{(N)})}\Big) = \mathrm{e}^{o(N)}\exp\Big(N\Big\langle\nu,\log\frac{\kappa\nu}{\nu}\Big\rangle\Big)\mathrm{e}^{N\delta_R}\end{aligned}$$

for some $\delta_R \in \mathbb{R}$ that converges to 0 as $R \to \infty$ and does not depend on $\ell^{(N)}$ nor on ν.

It remains to argue that the large-N exponential scale of the last term on the right-hand side of (4.29) vanishes when taking $R \to \infty$ afterwards. Recall that the choice of r may depend on k; we will denote it as r_k. We first use that

$$\prod_{R<|k|\le\varepsilon N}\Big(\frac{N\mathrm{e}C}{\ell_k^{(N)}(k_{r_k})^2}\Big)^{\ell_k^{(N)}} = \exp\Big(N\sum_{R<|k|\le\varepsilon N}\frac{\ell_k^{(N)}}{N}\log\frac{N\mathrm{e}C}{\ell_k^{(N)}(k_{r_k})^{2+|\mathcal{S}|}}\Big)\exp\Big(N|\mathcal{S}|\sum_{R<|k|\le\varepsilon N}\frac{\ell_k^{(N)}}{N}\log|k|\Big).$$

Abbreviating $D := \sum_{R<|k|<\varepsilon N}\ell_k^{(N)}/N$ and using Jensen's inequality we get that

$$\exp\Big(N\sum_{R<|k|\le\varepsilon N}\frac{\ell_k^{(N)}}{N}\log\frac{N\mathrm{e}C}{\ell_k^{(N)}(k_{r_k})^{2+|\mathcal{S}|}}\Big) \le \exp\Big(ND\log\Big(\tfrac{\mathrm{e}C}{D}\sum_{R<|k|\le\varepsilon N}k_{r_k}^{-2-|\mathcal{S}|}\Big)\Big). \quad (4.31)$$

It is easy to see that choosing r_k such that $k_{r_k} = \max_{s\in\mathcal{S}} k_s$ ensures the convergence of $\sum_k k_{r_k}^{-2-|\mathcal{S}|}$ and the fact that $\sum_{|k|\ge R} k_{r_k}^{-2-|\mathcal{S}|}$ is polynomial in R. Further, as we remarked below (4.30), $D \le 1/R$

and therefore the right-hand side of (4.31) is bounded by $\exp(N\delta'_R)$ for some δ'_R that vanishes as $R \to \infty$. Next, we use that $R/\log R \le |k|/\log|k|$ holds on our summation area of k if R is large enough and therefore

$$\exp\left(N|\mathcal{S}| \sum_{R<|k|\le\varepsilon N} \frac{\ell_k^{(N)}}{N} \log|k|\right) \le \exp\left(N|\mathcal{S}| \frac{\log R}{R} \sum_{R<|k|\le\varepsilon N} \frac{\ell_k^{(N)}}{N}|k|\right) = \exp\left(N|\mathcal{S}|\frac{\log R}{R}\right). \tag{4.32}$$

Noting that $\frac{1}{R}\log R \to 0$ as $R \to \infty$, we have shown that the last term on the right-hand side of (4.29) can be bounded by $\mathrm{e}^{N\delta''_R}$ for some δ''_R that vanishes as $R \to \infty$. So far, we have handled the first term in the definition of $z^{(N)}_{R,\varepsilon N}(\ell^{(N)})$, and we saw that its exponential rate is not larger than the first term in $-I_{\mathrm{Me}}(\nu)$.

Let us now handle the second and last part of $z^{(N)}_{R,\varepsilon N}(\ell^{(N)})$. We use $1-x \le \mathrm{e}^{-x}$ for $x = \kappa_N(r,s)/N$ as well as $\frac{1}{2}\sum_{R<k\le\varepsilon N} \ell_k^{(N)}\langle k, \kappa k\rangle \le \frac{1}{2}\|\kappa\|_\infty \varepsilon \sum_{R<|k|\le\varepsilon N} \ell_k^{(N)}|k| \le \frac{N}{2}\|\kappa\|_\infty\varepsilon$ to get that

$$\prod_{R<|k|\le\varepsilon N} \prod_{r,s\in\mathcal{S}} \left(1 - \tfrac{\kappa_N(r,s)}{N}\right)^{\frac{1}{2}k_s(N\mu_r - k_r)\ell_k^{(N)}} \le \mathrm{e}^{o(N)}\mathrm{e}^{-\frac{N}{2}\langle\nu,\kappa\mu\rangle}\mathrm{e}^{\frac{N}{2}\|\kappa\|_\infty\varepsilon}.$$

Noting that the exponential rate of the last factor on the right-hand side vanishes as $\varepsilon \to 0$, we see that we have completed the proof. □

4.4 Proof of Theorem 29

Here we finish now the proof of the LDP of Theorem 29. Let $d_{\mathcal{L}}$ and $d_{\mathcal{A}}$ be, respectively, metrics that induce the vague topologies on the state spaces $\mathcal{L}$ and $\mathcal{A}$ (see (4.3) and (4.4)). Notice that, thanks to the contraints $|c(\lambda)| \le 1$ and $|c(\alpha)| \le 1$, the spaces $\mathcal{L}$ and $\mathcal{A}$ endowed with the vague topologies are compact by the Bolzano–Weierstrass theorem and Fatou's lemma. Moreover the rate function I is lower semicontinuous on $\mathcal{L} \times \mathcal{A}$, which makes it a good rate function. Hence, a weak LDP implies the claim of Theorem 29 and it suffices to prove that

$$\lim_{\delta,\rho\to 0} \lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big(\mathrm{Mi}_N \in B_\delta(\lambda), \mathrm{Ma}_N \in B_\rho(\alpha)\big) = -I(\lambda,\alpha), \qquad \lambda \in \mathcal{L}, \alpha \in \mathcal{A}. \tag{4.33}$$

where $B_\delta(\lambda)$ and $B_\rho(\alpha)$ are closed balls centered at λ and α with radii δ and ρ, respectively.

Fix $\lambda \in \mathcal{L}$ and $\alpha \in \mathcal{A}$.

Step 1: Cardinality of $\mathcal{L}_N$. Recall the definition of $\mathcal{L}_N$ that was given in (4.12). Each $\ell \in \mathcal{L}_N$ has a unique representation as a product measure $\ell = \bigotimes_{r\in\mathcal{S}} \ell^{(r)}$, where $\ell^{(r)} = (\ell_j^{(r)})_{j\in\mathbb{N}_0}$ and $\sum_j \ell_j^{(r)} j = N\mu_r^{(N)}$ for all $r \in \mathcal{S}$. By the same argument as in [AKP19, Lemma 3.2] there are at most $e^{o(N\mu_r^{(N)})}$ ways to choose the marginal $\ell^{(r)}$. Consequently we have

$$|\mathcal{L}_N| = \mathrm{e}^{o(N)}.$$

Fix $R \in \mathbb{N}$ and $\varepsilon > 0$. We denote by $d_{\mathcal{L}}^R$ and $d_{\mathcal{A}}^\varepsilon$ the distance of the projections of measures on $\{k\colon |k| \le R\}$ and $\{y\colon |y| > \varepsilon\}$, respectively. Then $d_{\mathcal{L}} \ge d_{\mathcal{L}}^R$ and $d_{\mathcal{A}} \ge d_{\mathcal{A}}^R$. For $\delta > 0$ and $\rho > 0$ denote by $\mathcal{L}_N^{(R,\varepsilon)}(\delta,\rho)$ the set of all $\ell \in \mathcal{L}_N$ with $d_{\mathcal{L}}^R(\frac{1}{N}\ell, \lambda) < \delta$ and $d_{\mathcal{A}}^\varepsilon(\ell_{\lfloor N\cdot\rfloor}, \alpha) < \rho$. Note that

$$\mathbb{P}\big(\mathrm{Mi}_N \in B_\delta(\lambda), \mathrm{Ma}_N \in B_\rho(\alpha)\big) \le \sum_{\ell\in\mathcal{L}_N^{(R,\varepsilon)}(\delta,\rho)} \mathbb{P}\left(N\mathrm{Mi}_N = \ell\right).$$

According to the preceding, also

$$|\mathcal{L}_N^{(R,\varepsilon)}(\delta,\rho)| \le |\mathcal{L}_N| = \mathrm{e}^{o(N)}.$$

Step 2: Case $c(\lambda)+c(\alpha)\not\le\mu$. Assume that there is some $r\in\mathcal{S}$ such that $c_r(\lambda)+c_r(\alpha)>\mu_r$. Then it is easy to see that $\mathcal{L}_N(\delta,\rho)=\emptyset$ for sufficiently large N and hence

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\mathrm{Mi}_N\in B_\delta(\lambda),\mathrm{Ma}_N\in B_\rho(\alpha)\big)=-\infty=-I(\lambda,\alpha),$$

which is proved with the same argument as in [AKP19, Lemma 3.7].

Step 3: Proof of the upper bound in (4.33) *for* $c(\lambda)+c(\alpha)\le\mu$. For any $R\in\mathbb{N}$ and any $\varepsilon\in(0,1]$ we have by Step 1 and Lemma 32 that

$$\begin{aligned}\mathbb{P}\big(\mathrm{Mi}_N\in B_\delta(\lambda),\mathrm{Ma}_N\in B_\rho(\alpha)\big)&\le \mathrm{e}^{o(N)}\sup_{\ell\in\mathcal{L}_N^{(R,\varepsilon)}(\delta,\rho)}\mathbb{P}\left(N\mathrm{Mi}_N=\ell\right)\\&\le \mathrm{e}^{o(N)}\sup_{\ell\in\mathcal{L}_N^{(R,\varepsilon)}(\delta,\rho)}z_{0,R}^{(N)}(\ell)z_{R,\varepsilon N}^{(N)}(\ell)z_{\varepsilon N,N}^{(N)}(\ell).\end{aligned}\tag{4.34}$$

For $\tilde\lambda\in\mathcal{L}$ and $\tilde\alpha\in\mathcal{A}$ we define

$$c_s^{(R)}(\tilde\lambda):=\sum_{k\colon|k|\le R}\tilde\lambda_k k_s\quad\text{and}\quad c_s^{(\varepsilon)}(\tilde\alpha):=\int_{\{|y|>\varepsilon\}}y_s\,\tilde\alpha(\mathrm{d}y),\qquad s\in\mathcal{S}.\tag{4.35}$$

We require that $\varepsilon\in(0,1]\setminus\{|y|\colon y\in\mathrm{supp}(\alpha)\}$ (recall that $\mathrm{supp}(\alpha)$ is countable). This is a prerequisite to apply Lemma 34 to $\tilde\alpha$. Now, applying Lemmas 33, 34(1) and 35 to the right-hand side of (4.34) we get

$$\begin{aligned}&\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\mathrm{Mi}_N\in B_\delta(\lambda),\mathrm{Ma}_N\in B_\rho(\alpha)\big)\\&\quad\le\sup_{\substack{\tilde\lambda\colon d_{\mathcal{L}}^R(\tilde\lambda,\lambda)<\delta\\ \tilde\alpha\colon d_{\mathcal{A}}^\varepsilon(\tilde\alpha,\alpha)<\rho}}\Big(-I_{\mathrm{Mi}}^{(R)}(\tilde\lambda)-I_{\mathrm{Ma}}^{(\varepsilon)}(\tilde\alpha)-I_{\mathrm{Me}}\big(\mu-c^{(R)}(\tilde\lambda)-c^{(\varepsilon)}(\tilde\alpha)\big)\Big)+C(R,\varepsilon,\kappa),\end{aligned}$$

where we used the cut-off versions of the rate functions defined in Lemmas 33, 34 and 35 and $C(R,\varepsilon,\kappa)$ is the constant introduced in Lemma 35. Note that the functions $I_{\mathrm{Mi}}^{(R)}$ and $c^{(R)}$ are continuous (in any point) and the functions $I_{\mathrm{Ma}}^{(\varepsilon)}$ and $c^{(\varepsilon)}$ are continuous in α due to our requirement that $\varepsilon\notin\{|y|\colon y\in\mathrm{supp}(\alpha)\}$.

Hence,

$$\begin{aligned}&\limsup_{\delta,\rho\to0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(\mathrm{Mi}_N\in B_\delta(\lambda),\mathrm{Ma}_N\in B_\rho(\alpha)\big)\\&\quad\le -I_{\mathrm{Mi}}^{(R)}(\lambda)-I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha)-I_{\mathrm{Me}}(\mu-c^{(R)}(\lambda)-c^{(\varepsilon)}(\alpha))+C(R,\varepsilon,\kappa)\end{aligned}$$

Observe that the right-hand sideconverges to $-I(\lambda,\alpha)$, if we let $R\to\infty$ and $\varepsilon\to0$, which proves the upper bound in (4.33). Notice that the requirement that $\varepsilon\notin\{|y|\colon y\in\mathrm{supp}(\alpha)\}$ is no a problem since $\mathrm{supp}(\alpha)$ is countable.

Step 4: Construction of a recovery sequence. In this step, we prepare for the proof of the lower bound in (4.33) (see Step 5) by constructing an almost optimal sequence of ℓ's. We handle here only the

case that κ is irreducible; see Remark 36 for hints how to handle the case of a reducible κ. We may assume that $c(\lambda)+c(\alpha)\le\mu$, since the rate function is equal to ∞ otherwise. For the same reason, we also may assume that the mesoscopic mass $\nu := \mu - c(\lambda) - c(\alpha)$ satisfies $\nu \ll \kappa\nu$, since otherwise $I_{\mathrm{Me}}(\nu)=\infty$.

For $R\in\mathbb{N}$ and $\varepsilon\in(0,1]$ we construct a suitable recovery sequence $\ell^{(N)}=\ell^{(N)}(R,\varepsilon)$ that will turn out in Step 5 as asymptotically optimal. To this end, we construct it in such a way that it will put all mesoscopic mass $\nu := \mu - c(\lambda) - c(\alpha)$ into several components that can all be described by the same configuration $k^{(\mathrm{Me},N)}$ and which are actually on the lower end of the macroscopic scale, such that (3.4) of Theorem 15 can be applied. Our construction also ensures that all components of macroscopic scale have a strictly positive probability of being connected. Denote by $\mathbb{1}$ the element of $\mathbb{N}_0^{\mathcal{S}}$ that is equal to 1 in each entry. Define $k^{(\mathrm{Me},N)} := \lfloor N\varepsilon\nu\rfloor + \mathbb{1}$. Using the representation $\alpha=\sum_i\delta_{y^{(i)}}$ put $k^{(i,N)} := \lfloor Ny^{(i)}\rfloor+\mathbb{1}$ for $i\in J_\varepsilon(\alpha) := \{i\colon |y^{(i)}|>\varepsilon\}$.

Let us check now that we can actually apply Theorem 15 to $p_N(k^{(i,N)})$ for all $i\in J_\varepsilon(\alpha)\cup\{\mathrm{Me}\}$. To check that $\tau(k^{(i,N)})>0$ for all $i\in J_\varepsilon(\alpha)\cup\{\mathrm{Me}\}$ note the following: Depending on the support of $y^{(i)}$ we might not have that κ is irreducible with respect to $y^{(i)}$. However, the vectors $k^{(i,N)}$ have support on the full type space $\mathcal{S}$ for all $N\in\mathbb{N}$ by construction. Consequently, $\tau(k^{(i,N)})>0$ is implied by the fact that κ is irreducible with respect to μ. Secondly, since $\nu\ll\kappa\nu$, we also have that $y^{(i)}\ll\kappa y^{(i)}$ for all $i\in J_\varepsilon(\alpha)$: if not, $I_{\mathrm{Ma}}(\alpha)=\infty$ by definition and the lower bound $-\infty$ in (4.33) trivially holds. By construction it holds for all $i\in J_\varepsilon(\alpha)$ that $k_s^{(i,N)}=1$ for $s\notin\operatorname{supp}(y^{(i)})$ and for all $N\in\mathbb{N}$, which ensures the boundedness condition. The same holds for $k^{(\mathrm{Me},N)}$.

Define now

$$\ell_k^{(N)} := \begin{cases} \lfloor\lambda_k N\rfloor & \text{if } 2\le|k|\le R\\ 0 & \text{if } R<|k|\le\varepsilon N \text{ and } k\ne k^{(\mathrm{Me},N)}\\ \lfloor\varepsilon^{-1}\rfloor & \text{for } k=k^{(\mathrm{Me},N)}\\ \#\{i\in J_\varepsilon(\alpha)\colon k=k^{(i,N)}\} & \text{if } \varepsilon N<|k|\\ N\mu_r^{(N)}-\sum_{m:2\le|m|}\ell_m^{(N)}m_r & \text{if } k=\mathbf{e}_r \text{ for any } r\in\mathcal{S}, \end{cases} \tag{4.36}$$

where the last line ensures that $\ell^{(N)}\in\mathcal{L}_N$. Observe that

$$\begin{aligned} &\lim_{N\to\infty}\frac{1}{N}\ell_{\mathbf{e}_r}^{(N)}-\lambda_{\mathbf{e}_r} = c_r(\lambda)-c_r^{(R)}(\lambda)+\nu_r-\lfloor\varepsilon^{-1}\rfloor(\varepsilon\nu_r)+c_r(\alpha)-c_r^{(\varepsilon)}(\alpha) && \text{for } r\in\mathcal{S}\\ &\lim_{N\to\infty}\frac{1}{N}\ell_k^{(N)}-\lambda_k=0 && \text{for } 2\le|k|\le R, \end{aligned}$$

which implies that $\lim_{R\to\infty,\varepsilon\to0}\lim_{N\to\infty}d_{\mathcal{L}}(\frac{1}{N}\ell^{(N)},\lambda)=0$. Similarly, $\lim_{\varepsilon\to0}\lim_{N\to\infty}d_{\mathcal{A}}(\alpha^{(N)},\alpha)=0$, where $\alpha^{(N)}=\sum_k\ell_k^{(N)}\delta_{\frac{k}{N}}$.

Step 5: Proof of the lower bound in (4.33). Now we finish the proof of the lower bound by showing that the recovery sequence $(\ell^{(N)})_{N\in\mathbb{N}}$ that we constructed in Step 4 is giving the right asymptotics.

Fix $\delta,\rho>0$. Then by choosing $R\in\mathbb{N}$ large enough and $\varepsilon>0$ small enough, we have that $\frac{1}{N}\ell^{(N)}\in B_\delta(\lambda)$ and $\alpha^{(N)}\in B_\rho(\alpha)$ for all but finitely many $N\in\mathbb{N}$. Hence, Lemma 32 implies that

$$\begin{aligned}\mathbb{P}\left(\mathrm{Mi}_N\in B_\delta(\lambda),\mathrm{Ma}_N\in B_\rho(\alpha)\right) &\ge \mathbb{P}(N\mathrm{Mi}_N=\ell^{(N)})\\ &= \mathrm{e}^{o(N)}z_{0,R}^{(N)}(\ell^{(N)})z_{R,\varepsilon N}^{(N)}(\ell^{(N)})z_{\varepsilon N,N}^{(N)}(\ell^{(N)}).\end{aligned}$$

Next, we want to use Lemmas 33 and 34(2) to get the exponential rates for the z-terms. Note that we do not have $\lim_{N\to\infty}\frac{1}{N}\ell^{(N)}=\lambda$ on $\{k\colon |k|\le R\}$, so we will instead apply Lemma 33 to the

sequence $\lfloor \lambda N \rfloor$, which gives

$$\lim_{N\to\infty} \frac{1}{N} \log z_{0,R}^{(N)}(\lfloor \lambda N \rfloor) = -I_{\mathrm{Mi}}^{(R)}(\lambda)$$

Using the definition (4.15), it is easy to verify that

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \log z_{0,R}^{(N)}(\ell^{(N)}) &= \lim_{N\to\infty} \frac{1}{N} \log z_{0,R}^{(N)}(\lfloor \lambda N \rfloor) + \lim_{N\to\infty} \frac{1}{N} \sum_{r\in\mathcal{S}} \log \frac{z^{(N)}(\ell_{\mathbf{e}_r}^{(N)}, \mathbf{e}_r)}{z^{(N)}(\lfloor \lambda_{\mathbf{e}_r} N \rfloor, \mathbf{e}_r)} \\ &= -I_{\mathrm{Mi}}^{(R)}(\lambda) + \widetilde{C}(R,\varepsilon).\end{aligned}$$

for some constant $\widetilde{C}(R,\varepsilon)$ vanishing as $R \to \infty$ and $\varepsilon \to 0$. For the mesoscopic part, we use the asymptotic formula (3.4) of Theorem 15 for $p_N(k^{(\mathrm{Me},N)})$. We proved in Step 4 that we can actually apply Theorem 15 to $p_N(k^{(\mathrm{Me},N)})$. This, together with Stirling's formula, gives us

$$\lim_{N\to\infty} \frac{1}{N} \log z_{R,\varepsilon N}^{(N)}(\ell^{(N)}) = \lfloor \varepsilon^{-1} \rfloor \Big\langle \varepsilon\nu, \log \frac{(1-\mathrm{e}^{-\kappa(\varepsilon\nu)})\mu}{\varepsilon\nu} \Big\rangle - \frac{1}{2} \lfloor \varepsilon^{-1} \rfloor \langle \varepsilon\nu, \kappa(\mu - \varepsilon\nu) \rangle.$$

The right-hand side clearly converges to $-I_{\mathrm{Me}}(\nu)$ as $\varepsilon \to 0$. The analysis of the macroscopic part in Step 4 shows that we can directly use Lemma 34(2) to see that

$$\lim_{N\to\infty} \frac{1}{N} \log z_{\varepsilon N,N}^{(N)}(\ell^{(N)}) = -I_{\mathrm{Ma}}^{(\varepsilon)}(\alpha).$$

Taking the limits $R \to \infty$ and $\varepsilon \to 0$ we obtain the lower bound in (4.33). This finishes the proof of Theorem 29.

Remark 36. (The reducible case: proof of Theorem 30) *Recall from Section 2.3 that in the case that κ is reducible, one can decompose the type space $\mathcal{S} = \bigcup_j \mathcal{S}_j$ in such a way that κ restricted to $\mathcal{S}_j \times \mathcal{S}_j$ is irreducible and $\kappa|_{\mathcal{S}_i \times \mathcal{S}_j} = 0$ for $i \neq j$. Assume that $\alpha \in \mathcal{A}$ is connectable as defined in Section 2.3: for each $y \in \mathrm{supp}(\alpha)$ there is some j such that $\mathrm{supp}(y) \subset \mathcal{S}_j$. Then we can argue that the lower bound in the proof of* (4.33) *holds. For any j we can construct a recovery sequence as before, where we approximate the macroscopic components by defining $k^{(i,N)} := \lfloor N y^{(i)} \rfloor + \mathbb{1}_{\mathcal{S}_j}$ for any $y^{(i)} \in \mathrm{supp}(\alpha)$ to ensure that $\tau(k^{(i,N)}) > 0$. In the same way we approximate the mesoscopic mass by defining $\nu^{(j)} := \nu \mathbb{1}_{\mathcal{S}_j}$ and $k^{(\mathrm{Me},j,N)} := \lfloor N\varepsilon\nu^{(j)} \rfloor + \mathbb{1}_{\mathcal{S}_j}$. The rest will work out as in the proof above. The only additional observation needed is that for any j we have that $(\kappa\nu^{(j)})_s = (\kappa\nu)_s$ if $s \in \mathcal{S}_j$.*

On the other hand, if $\alpha \in \mathcal{A}$ is not connectable, then we can argue that the upper bound of the LDP holds with $I(\lambda,\alpha) = \infty$ by additionally requiring that $\kappa_N = \kappa$ for all but finitely many $N \in \mathbb{N}$: Indeed, there exists $y \in \mathrm{supp}(\alpha)$ such that for any sequence $k^{(N)}$ with $\lim_{N\to\infty} \frac{k^{(N)}}{N} = y$ we have that $\tau(k^{(N)}) = 0$ and by using that $\kappa_N = \kappa$ we get that $p_N(k^{(N)}) \leq \tau(k^{(N)}) N^{-(|k|-1)} = 0$, which gives that $I(\lambda,\alpha) = \infty$.

5 Proof of Theorem 1

The aim of this section is to finally prove our main result, Theorem 1, for the general case, where the type space $\mathcal{S}$ is some compact metric space. The main idea is to approximate the general graph model introduced in Section 1.1 by a discretized model with a finite type space and discretized kernels. To

derive the LDP we will use the LDP of Theorem 29 for finite type spaces, together with the Dawson–Gärtner theorem, which we will slightly modify for our purposes.

In Section 5.1 we introduce the approximation scheme and derive lower and upper approximations for the distribution of the empirical measures of the general graph via a comparison with certain discretized graph models. In order to lift statements about the distribution of the discretized model to the general case we introduce in Section 5.2 a projective system. There, we also identify the projective limit spaces with the state spaces of our empirical measures, i.e., with $\mathcal{L} \times \mathcal{A}$, and verify that the projective limit topology is strong enough to imply an LDP result also with respect to our chosen topology. In Section 5.3 we conclude with the derivation of our main result by using the Dawson–Gärtner approach, adapted to our purposes: we have to use different distributions for the lower and the upper bound and will have to deal with some additional technical difficulty concerning the lower bound. Finally, Theorem 1 is implied by combining the results of Lemma 52, where we prove the upper bound, and Lemma 54, where we prove the lower bound.

The following objects will be fixed for the remainder of the section. Let $\mathcal{S}$ be a compact metric space. Fix a probability measure μ on $\mathcal{S}$. For any N fix a type vector $\mathbf{x} = \mathbf{x}^{(N)} = (x_1, \dots, x_N) \in \mathcal{S}^N$ such that its empirical measure $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ weakly converges to μ as $N \to \infty$. Furthermore, let a continuous kernel $\kappa\colon \mathcal{S} \times \mathcal{S} \to [0, \infty)$ be given that is irreducible w.r.t. μ, as well as a sequence of continuous kernels κ_N that converges to κ uniformly on $\mathcal{S} \times \mathcal{S}$ as $N \to \infty$. By $\mathcal{G}_N = \mathcal{G}([N], \mathbf{x}, \frac{1}{N}\kappa_N)$ we denote the inhomogeneous random graph introduced in Section 1.1 and by $(\mathcal{C}_i)_i$ we denote the vertex sets of the connected components of $\mathcal{G}_N$. We will denote the probability measure corresponding to the graph $\mathcal{G}_N$ by $\mathbb{P}_N$. Recall the definition of the microscopic and macroscopic empirical measures Mi_N and Ma_N given in (1.2). The goal is to derive the LDP for the pair $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ that is formulated in Theorem 1.

Before chosing a discretization scheme, we would like to collect some properties of our state space $\mathcal{L} \times \mathcal{A}$. Recall that

$$\mathcal{L} := \big\{\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}))\colon c_\lambda \le \mu \text{ or } c_\lambda \le \mu_N \text{ for some } N \in \mathbb{N},\ \lambda(\{0\}) = 0\big\}, \tag{5.1}$$

$$\mathcal{A} := \big\{\alpha \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{M}(\mathcal{S}) \setminus \{0\})\colon c_\alpha \le \mu \text{ or } c_\alpha \le \mu_N \text{ for some } N \in \mathbb{N}\big\}, \tag{5.2}$$

and that both are equipped with vague topologies, i.e., a sequence $(\lambda_n)_{n\in\mathbb{N}}$ in $\mathcal{L}$ converges to λ, if for any continuous, compactly supported test function $g\colon \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \to \mathbb{R}$ the integrals $\int g \,\mathrm{d}\lambda_n$ converge to $\int g \,\mathrm{d}\lambda$, as $n \to \infty$. In the same way, a sequence $(\alpha_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ converges to α, if for any continuous, compactly supported test function $g\colon \mathcal{M}(\mathcal{S}) \setminus \{0\} \to \mathbb{R}$ the integrals $\int g \,\mathrm{d}\alpha_n$ converge to $\int g \,\mathrm{d}\alpha$, as $n \to \infty$. Both on $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ and $\mathcal{M}(\mathcal{S}) \setminus \{0\}$ we consider the topologies of weak convergence. In the next lemma, we give a short characterization of compactness that is implied by this choice; a verification is left to the reader.

Lemma 37. *The following assertions hold:*

1. *A subset $\mathcal{N} \subset \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ is compact if and only if it is closed and $\sup\{\nu(\mathcal{S})\colon \nu \in \mathcal{N}\} < \infty$.*
2. *A subset $\mathcal{N} \subset \mathcal{M}_{\le 1}(\mathcal{S}) \setminus \{0\}$ is compact if and only if it is closed and $\inf\{\nu(\mathcal{S})\colon \nu \in \mathcal{N}\} > 0$.*

The following lemma implies that any vague accumulation points of Mi_N and Ma_N are indeed elements of $\mathcal{L}$ and $\mathcal{A}$ if they exist.

Lemma 38. *Both $\mathcal{L}$ and $\mathcal{A}$ are compact spaces.*

Proof. Denote $M := \{\mu\} \cup \{\mu_N \colon N \in \mathbb{N}\}$.

(1) Compactness of $\mathcal{L}$: For any $\lambda \in \mathcal{L}$ note that $|\lambda| \leq c_\lambda(\mathcal{S}) \leq \sup_{\tilde{\mu}\in M} \tilde{\mu}(\mathcal{S}) = 1$. The set $B_1 := \{\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})) \colon |\lambda| \leq 1\}$ is a bounded subset of the dual of $C_0(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \setminus \{0\})$ (the space of continuous functions g on $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ with $\lim_{k(\mathcal{S})\to\infty} g(k) = 0$), so by applying the Banach-Alaoglu Theorem we get that B_1 is compact w.r.t. the vague topology. Since $\mathcal{L} \subset B_1$ it remains to argue that $\mathcal{L}$ is closed. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}$ with vague limit λ. Then for each $n \in \mathbb{N}$ there exists $\tilde{\mu}_n \in M$ such that $c_{\lambda_n} \leq \tilde{\mu}_n$. Since M is compact w.r.t. the weak topology, we can find $\tilde{\mu} \in M$ and a subsequence (which we will also denote by $(\tilde{\mu}_n)_{n\in\mathbb{N}}$) such that $\tilde{\mu}_n \to \tilde{\mu}$ weakly, as $n \to \infty$. We now argue that this implies that $c_\lambda \leq \tilde{\mu}$. Let $f \colon \mathcal{S} \to [0,\infty)$ be a continuous and bounded function. For any $R \in \mathbb{N}$, let $\chi_R \colon [0,\infty) \to [0,\infty)$ be a smooth function satisfying $\mathbb{1}_{[0,R]} \leq \chi_R \leq \mathbb{1}_{[0,R+1]}$, such that $R \mapsto \chi_R$ is increasing pointwise. Define

$$\Phi_R^f \colon \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \to \mathbb{R}, \quad \Phi_R^f(k) = \Big(\int_{\mathcal{S}} f(s)\, k(\mathrm{d}s)\Big)\chi_R(k(\mathcal{S})) \tag{5.3}$$

and note that $\Phi_R^f(k) \nearrow \int f(s)\, k(\mathrm{d}s)$, as $R \to \infty$, pointwise for any k. By Lemma 37 we have that Φ_R^f is compactly supported and it can be easily seen that continuity of χ_R and f imply that Φ_R^f is continuous. Therefore

$$\begin{aligned}\int_{\mathcal{S}} f(s)\, c_\lambda(\mathrm{d}s) &= \sup_{R\in\mathbb{N}} \int \Phi_R^f(k)\, \lambda(\mathrm{d}k) = \sup_{R\in\mathbb{N}} \lim_{n\to\infty} \int \Phi_R^f(k)\, \lambda_n(\mathrm{d}k) \\ &\leq \sup_{R\in\mathbb{N}} \lim_{n\to\infty} \int_{\mathcal{S}} f(s)\, c_{\lambda_n}(\mathrm{d}s) \leq \lim_{n\to\infty} \int_{\mathcal{S}} f(s)\, \tilde{\mu}_n(\mathrm{d}s) = \int_{\mathcal{S}} f(s)\, \tilde{\mu}(\mathrm{d}s).\end{aligned}$$

Since this holds for any $f \geq 0$, we can conclude that $c_\lambda \leq \mu$ and thus $\lambda \in \mathcal{L}$.

(2) We only sketch the construction of a vague limit point. Fix a sequence $(\varepsilon_i)_{i\in\mathbb{N}}$, with $\varepsilon_i \searrow 0$ as $i \to \infty$, and define $N_{\varepsilon_i} := \{y \in \mathcal{M}_{\leq 1}(\mathcal{S}) \setminus \{0\} \colon |y| \geq \varepsilon_i\}$. Note that $\varepsilon_i \alpha_n(N_{\varepsilon_i}) \leq c_{\alpha_n}(\mathcal{S}) \leq 1$ implies that $\alpha_n(N_{\varepsilon_i}) \leq 1/\varepsilon_i$. Thus, for fixed $i \in \mathbb{N}$, the restricted measures $(\alpha_n|_{N_{\varepsilon_i}})_{n\in\mathbb{N}}$ are bounded and thus have a vaguely converging sub-sequence. By diagonalization we can construct for each $i \in \mathbb{N}$ a subsequence of $(\alpha_n)_{n\in\mathbb{N}}$ such that the restrictions to N_{ε_i} converge vaguely to some $\alpha^{(i)} \in \mathcal{M}(N_{\varepsilon_i})$ and in such a way that $\alpha^{(i+1)}|_{N_{\varepsilon_i}} = \alpha^{(i)}$ holds for all $i \in \mathbb{N}$. Thus the monotone limit $\lim_{i\to\infty} \alpha^{(i)} =: \alpha$ on $\mathcal{M}_{\leq 1}(\mathcal{S}) \setminus \{0\}$ is a countably additive extension of the measures $\alpha^{(i)}$, $i \in \mathbb{N}$. Since every compactly supported test function has its support contained in N_{ε_i} for some i by Lemma 37 one sees that in the vague topology $\alpha_n \to \alpha$.

To see that $\alpha \in \mathcal{A}$ (i.e., that $\mathcal{A}$ is closed) we can proceed as in the proof for $\mathcal{L}$. Note that for $\varepsilon > 0$ we can find a smooth function $\chi_\varepsilon \colon [0,\infty) \to [0,\infty)$ satisfying $\mathbb{1}_{[2\varepsilon,\infty)} \leq \chi_\varepsilon \leq \mathbb{1}_{[\varepsilon,\infty)}$, which allows us to define Φ_ε^f as above but by truncating with χ_ε. □

5.1 Discretization and Approximation

Let $(P_m)_{m\in\mathbb{N}}$ be a sequence of finite partitions of $\mathcal{S}$ into non-empty sets. For $m \in \mathbb{N}$ we denote $P_m = \{A_{m,i} \colon i = 1, \ldots, |P_m|\}$. We say that $(P_m)_{m\in\mathbb{N}}$ is nested if for $m \leq n$ and any $A_{m,i} \in P_m$ there is $J \subset \{1, \ldots, |P_n|\}$, such that $A_{m,i} = \bigcup_{j\in J} A_{n,j}$. For any subset A we write $\mathrm{diam}(A) := \sup\{d(x,y) \colon x, y \in A\}$. For any measure ν on $\mathcal{S}$ a set $A \subset \mathcal{S}$ is called a continuity set of ν if $\nu(\partial A) = 0$. The following lemma is a modification of Lemma 7.1 from [BJR07].

Lemma 39. *There exists a sequence of finite partitions $(P_m)_{m\in\mathbb{N}}$ of $\mathcal{S}$ with the following properties:*

1 *For any $m \in \mathbb{N}$ and any $i = 1, \dots, |P_m|$ we have that $A_{m,i}$ is measurable and a continuity set of μ and μ_N for all $N \in \mathbb{N}$.*

2 *$(P_m)_{m\in\mathbb{N}}$ is nested.*

3 *It holds that*

$$\lim_{m\to\infty} \max_{i=1,\dots,|P_m|} \operatorname{diam}(A_{m,i}) = 0. \tag{5.4}$$

Proof. The proof can be done as in the proof of Lemma 7.1 in [BJR07] with two small modifications. Let $A \subset \mathcal{S}$ be the set of points that are atoms of μ or μ_N for some $N \in \mathbb{N}$. Then A is still a countable set, so we may pick the balls B_{mi} in a way such that they are continuity sets of μ and μ_N for all $N \in \mathbb{N}$. Since in our case the set $\mathcal{S}$ is compact, we can cover it with finitely many of these balls, hence we get the stronger property formulated in (5.4). □

In the following we always assume that $(P_m)_{m\in\mathbb{N}}$ has all the properties given in Lemma 39. For $m \in \mathbb{N}$ and any $A_{m,i} \in P_m$ we pick exactly one point $x_{m,i}$ from the set A_m which we call the representative of $A_{m,i}$. We define

$$\mathcal{S}_m := \{x_{m,i} \colon i = 1, \dots, |P_m|\} \tag{5.5}$$

and define the projection

$$\pi_m \colon \mathcal{S} \to \mathcal{S}_m, \quad x \mapsto x_{m,i} \tag{5.6}$$

where i is such that $A_{m,i}$ is the unique set containing x.

Further we can lift the projection to the space $\mathcal{M}(\mathcal{S})$, the space of finite measures on $\mathcal{S}$. For any $m \in \mathbb{N}$ we introduce (by abuse of notation)

$$\pi_m \colon \mathcal{M}(\mathcal{S}) \to \mathcal{M}(\mathcal{S}_m), \quad \pi_m(\nu) := \nu \circ \pi_m^{-1} \quad \text{for } \nu \in \mathcal{M}(\mathcal{S})\text{..} \tag{5.7}$$

Going one level further, we also define

$$\pi_m \colon \mathcal{M}(\mathcal{M}(\mathcal{S})) \to \mathcal{M}(\mathcal{M}(\mathcal{S}_m)), \quad \pi_m(\lambda) := \lambda \circ \pi_m^{-1} \quad \text{for } \lambda \in \mathcal{M}(\mathcal{M}(\mathcal{S})). \tag{5.8}$$

We can now apply the projections to all the levels of our graph setting. On the first level of discretization, i.e., from $\mathcal{S}$ to $\mathcal{S}_m$, we approximate the type of a vertex by some type from the discrete set $\mathcal{S}_m$. On the second level, we approximate the type configuration of a vertex set which, depending on the context, may or may not be a cluster. On the third level, we approximate measures that count the multiplicities of type configurations and are therefore suited to register the number of clusters described by the different type configurations.

Fix $m \in \mathbb{N}$. We write $\pi_m(\mathbf{x}) = (\pi_m(x_1), \dots, \pi_m(x_N)) \in (\mathcal{S}_m)^N$ for the discretized type sequence and denote by $\mu_N^{(m)} = \frac{1}{N}\sum_{i=1}^N \delta_{\pi_m(x_i)}$ its empirical measure. It is easy to check that $\mu_N^{(m)} = \pi_m(\mu_N)$. Since our partition P_m has carefully been chosen such that the sets $A_{m,i} \in P_m$ are continuity sets of μ, it holds that $\mu_N^{(m)}$ converges weakly to $\mu^{(m)} := \pi_m(\mu)$, as $N \to \infty$. We also define $\eta_{\pi_m(\mathbf{x})} \colon \mathcal{P}([N]) \to \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)$ as the discrete analog of the type registering measure $\eta_{\mathbf{x}}$ from (1.1), i.e., for any $A \subset [N]$ we have that $\eta_{\pi_m(\mathbf{x})} = \sum_{i\in A} \delta_{\pi_m(x_i)} \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)$. Abbreviating $\eta_m = \eta_{\pi_m(\mathbf{x})}$ we can now define

$$\mathrm{Mi}_N^{(m)} = \frac{1}{N}\sum_i \delta_{\eta_m(\mathcal{C}_i)} \qquad \text{and} \qquad \mathrm{Ma}_N^{(m)} = \sum_i \delta_{\frac{1}{N}\eta_m(\mathcal{C}_i)}. \tag{5.9}$$

It is straight forward to show that $\mathrm{Mi}_N^{(m)} = \pi_m(\mathrm{Mi}_N)$ and $\mathrm{Ma}_N^{(m)} = \pi_m(\mathrm{Ma}_N)$. Both empirical measures $\mathrm{Mi}_N^{(m)}$ and $\mathrm{Ma}_N^{(m)}$ evaluate type information given by the vertices of the clusters only roughly by approximating the type of each vertex by an element from $\mathcal{S}_m$. Note that $c_{\mathrm{Mi}_N^{(m)}} = \pi_m(c_{\mathrm{Mi}_N}) \leq \mu_N^{(m)}$ and $c_{\mathrm{Ma}_N^{(m)}} = \pi_m(c_{\mathrm{Ma}_N}) \leq \mu_N^{(m)}$. Hence, the natural state spaces for the discretized empirical measures are given by

$$\mathcal{L}_m = \big\{\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m))\colon c_\lambda \leq \mu^{(m)} \text{ or } c_\lambda \leq \mu_N^{(m)} \text{ for some } N \in \mathbb{N},\ \lambda(\{0\}) = 0\big\}, \tag{5.10}$$

$$\mathcal{A}_m = \big\{\alpha \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{M}(\mathcal{S}_m) \setminus \{0\})\colon c_\alpha \leq \mu^{(m)} \text{ or } c_\alpha \leq \mu_N^{(m)} \text{ for some } N \in \mathbb{N}\big\}, \tag{5.11}$$

and we endow them with vague topologies.

At this point it is important to note that we cannot apply the discrete LDP from Theorem 29 directly to get an LDP for the measure $\mathbb{P}_N((\mathrm{Mi}_N^{(m)}, \mathrm{Ma}_N^{(m)}) \in \cdot)$, since the edges of the random graph $\mathcal{G}_N$ are drawn according to the non-discretized types of the vertices. Before applying Theorem 29 one has to approximate the underlying graph itself by a discrete version.

Let $\kappa_N^{(m)}\colon \mathcal{S}_m \times \mathcal{S}_m \to [0,\infty)$, $N \in \mathbb{N}$, be a sequence of kernels on $\mathcal{S}_m$. We will specify later, in which sense they should be an approximation for κ_N. Consider the inhomogeneous random graph $\mathcal{G}_N^{(m)} = \mathcal{G}([N], \pi_m(\mathbf{x}), \frac{1}{N}\kappa_N^{(m)})$ and denote by $(\mathcal{C}_i^{(m)})_i$ the collection of the vertex sets of its connected components.

Instead of choosing just one approximating kernel, we will consider a lower and an upper approximation for κ_N, which will allow us to find upper and lower bounds for the distribution $\mathbb{P}_N((\mathrm{Mi}_N^{(m)}, \mathrm{Ma}_N^{(m)}) \in \cdot)$. For fixed $m \in \mathbb{N}$ let $\kappa_N^{(m,-)}$ and $\kappa_N^{(m,+)}$, $N \in \mathbb{N}$, be two sequences of kernels on $\mathcal{S}_m$ satisfying

$$\kappa_N^{(m,-)}(\pi_m(x), \pi_m(\hat{x})) \leq \kappa_N(x,\hat{x}) \leq \kappa_N^{(m,+)}(\pi_m(x), \pi_m(\hat{x})), \quad \forall\, x, \hat{x} \in \mathcal{S},\ \forall\, N \in \mathbb{N}. \tag{5.12}$$

An obvious choice is given by

$$\kappa_N^{(m,-)}(r,s) := \inf\{\kappa_N(x,\hat{x})\colon x,\hat{x} \in \mathcal{S},\ \pi_m(x) = r,\ \pi_m(\hat{x}) = s\}, \quad \text{for } r,s \in \mathcal{S}_m \tag{5.13}$$

$$\kappa_N^{(m,+)}(r,s) := \sup\{\kappa_N(x,\hat{x})\colon x,\hat{x} \in \mathcal{S},\ \pi_m(x) = r,\ \pi_m(\hat{x}) = s\}, \quad \text{for } r,s \in \mathcal{S}_m. \tag{5.14}$$

Defining $\kappa^{(m,*)} = \lim_{N\to\infty} \kappa_N^{(m,*)}$ for both $* \in \{+,-\}$ it is obvious that

$$\kappa^{(m,-)}(\pi_m(x), \pi_m(\hat{x})) \leq \kappa(x,\hat{x}) \leq \kappa^{(m,+)}(\pi_m(x), \pi_m(\hat{x})), \quad \text{for all } x,\hat{x} \in \mathcal{S}, \tag{5.15}$$

and we see that

$$\lim_{m\to\infty} \kappa^{(m,*)}(\pi_m(x), \pi_m(\hat{x})) = \kappa(x,\hat{x}), \qquad \text{uniformly in } x,\hat{x} \in \mathcal{S}, * \in \{+,-\}. \tag{5.16}$$

For both $* \in \{+,-\}$ we will denote by $\mathbb{P}_N^{(m,*)}$ the probability measure corresponding to the graph $\mathcal{G}_N^{(m,*)} = \mathcal{G}([N], \pi_m(\mathbf{x}), \frac{1}{N}\kappa_N^{(m,*)})$.

The following comparison lemma shows how we can estimate the distribution $\mathbb{P}_N((\mathrm{Mi}_N^{(m)}, \mathrm{Ma}_N^{(m)}) \in \cdot)$ from below and above.

Lemma 40. *Fix $m \in \mathbb{N}$ and let $\kappa_N^{(m,-)}$ and $\kappa_N^{(m,+)}$, $N \in \mathbb{N}$, be two sequences of kernels on $\mathcal{S}_m$ satisfying* (5.12)*. Assume that $N \in \mathbb{N}$ is large enough such that $\frac{1}{N}\kappa_N^{(m,+)} \leq 1$. Then, for any $\ell \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m))$,*

$$(\Delta_N^{(m)})^{-1}\mathbb{P}_N^{(m,-)}(N\mathrm{Mi}_N = \ell) \leq \mathbb{P}_N(N\mathrm{Mi}_N^{(m)} = \ell) \leq \mathbb{P}_N^{(m,+)}(N\mathrm{Mi}_N = \ell)\Delta_N^{(m)}, \tag{5.17}$$

where

$$\Delta_N^{(m)} = \prod_{r,s\in\mathcal{S}_m} \Big(\frac{1 - \frac{1}{N}\kappa_N^{(m,-)}(r,s)}{1 - \frac{1}{N}\kappa_N^{(m,+)}(r,s)}\Big)^{\frac{1}{2}N^2\mu_N^{(m)}(r)\mu_N^{(m)}(s)}. \tag{5.18}$$

Proof. As in the discrete setting of Section 4.1 we can identify elements from $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)$ with elements from $\mathbb{N}_0^{\mathcal{S}_m}$ and identify ℓ with $(\ell_k)_{k\in\mathbb{N}_0^{\mathcal{S}_m}}$. For $*\in\{+,-\}$ and any $k\in\mathbb{N}_0^{\mathcal{S}_m}$ we write

$$p_N^{(m,*)}(k)=\mathbb{P}_N^{(m,*)}\big(\mathcal{G}(|k|,\mathbf{x}_k,\tfrac{1}{N}\kappa_N^{(m,*)})\text{ is connected }\big),\tag{5.19}$$

with $\mathbf{x}_k\in\mathcal{S}_m^{|k|}$ any $|k|$-dimensional vector compatible with k. The main idea is that we identify in the exact formula (4.13) in Lemma 31 those terms that are increasing in κ and those that are decreasing in κ. Indeed, the connection probabilities are increasing in κ since the event of being connected is increasing in the edge parameter. Indeed, for any $k\in\mathbb{N}_0^{\mathcal{S}_m}$ and $\mathbf{x}\in\mathcal{S}^{|k|}$ such that $\pi_m(\sum_{i=1}^{|k|}\delta_{x_i})=k$,

$$p_N^{(m,-)}(k)\le\mathbb{P}_N(\mathcal{G}(|k|,\mathbf{x},\tfrac{1}{N}\kappa_N)\text{ is connected })\le p_N^{(m,+)}(k).$$

Furthermore, the powers of $(1-\frac{1}{N}\kappa(\cdot))$ that describe the probabilities of not being connected are increasing for negative powers, and decreasing for positive powers. With those observations and combinatorial factors we can estimate

$$\begin{aligned}\mathbb{P}_N(N\mathrm{Mi}_N^{(m)}=\ell)\le&\Big(\prod_{r\in\mathcal{S}_m}(N\mu_N^{(m)}(r))!\Big)\Big(\prod_{k\in\mathbb{N}_0^{\mathcal{S}_m}}\frac{p_N^{(m,+)}(k)^{\ell_k}}{\ell_k!\prod_{r\in\mathcal{S}_m}(k_r!)^{\ell_k}}\Big)\\&\times\Big(\prod_{r,s\in\mathcal{S}_m}\big(1-\tfrac{1}{N}\kappa_N^{(m,+)}(r,s)\big)^{\frac{1}{2}\sum_k k_r(N\mu_N^{(m)}(s)-k_s)\ell_k}\Big)\Delta_N^{(m)}\end{aligned}\tag{5.20}$$

The lower bound and its proof are analogous. □

5.2 Projective system

Using our discretization scheme from Section 5.1 we now introduce a projective system that fits into the framework of [DZ10, Section 4.6]. Recall the notions introduced at the beginning of Section 5.1, in particular the family of finite partitions $(P_m)_{m\in\mathbb{N}}$ that has the properties formulated in Lemma 39 and the definition (5.5) of the discretized type spaces $\mathcal{S}_m$. For any pair $m,n\in\mathbb{N}$ with $m\le n$ we define

$$\pi_{m,n}\colon\mathcal{S}_n\to\mathcal{S}_m,\quad x_{n,j}\mapsto x_{m,i}\tag{5.21}$$

where $x_{m,i}$ is the representative of $A_{m,i}$ and $A_{m,i}$ is the unique set in P_m containing $x_{n,j}$. As before, we lift this definition to the measure spaces by defining (with abuse of notation)

$$\pi_{m,n}\colon\mathcal{M}(\mathcal{S}_n)\to\mathcal{M}(\mathcal{S}_m),\quad\pi_{m,n}(\nu)=\nu\circ\pi_{m,n}^{-1},\tag{5.22}$$

as well as

$$\pi_{m,n}\colon\mathcal{M}(\mathcal{M}(\mathcal{S}_n))\to\mathcal{M}(\mathcal{M}(\mathcal{S}_m)),\quad\pi_{m,n}(\lambda)=\lambda\circ\pi_{m,n}^{-1}.\tag{5.23}$$

It is easy to check that restricting the latter mapping to $\mathcal{L}_n$ and $\mathcal{A}_n$, respectively, gives us two well-defined mappings $\pi_{m,n}\colon\mathcal{L}_n\to\mathcal{L}_m$ and $\pi_{m,n}\colon\mathcal{A}_n\to\mathcal{A}_m$, respectively. Now, $(\mathcal{L}_m,\pi_{m,n})_{m\le n}$ (or $(\mathcal{A}_m,\pi_{m,n})_{m\le n}$) is called a projective system, if

- for any $m\in\mathbb{N}$ the space $\mathcal{L}_m$ is a Hausdorff topological space,
- for any $m,n\in\mathbb{N}$ with $m\le n$ the mapping $\pi_{m,n}\colon\mathcal{L}_n\to\mathcal{L}_m$ is continuous,
- for any $m,n,p\in\mathbb{N}$ with $m\le n\le p$ we have $\pi_{m,p}=\pi_{m,n}\circ\pi_{n,p}$.

The projective limit of the projective system $(\mathcal{L}_m, \pi_{m,n})_{m\leq n}$ is denoted by $\varprojlim \mathcal{L}_m$ and is defined as the subset of the product space $\prod_{m\in\mathbb{N}} \mathcal{L}_m$ that contains all elements $(\lambda_m)_{m\in\mathbb{N}}$ that satisfy $\pi_{m,n}(\lambda_n) = \lambda_m$ for any pair $m, n \in \mathbb{N}$ with $m \leq n$. The projective limit $\varprojlim \mathcal{L}_m$ is equipped with the topology that is induced by the product topology on $\prod_{m\in\mathbb{N}} \mathcal{L}_m$. We call this the projective limit topology. In particular, a sequence $\lambda^{(n)} \in \varprojlim \mathcal{L}_m$ converges to $\lambda \in \varprojlim \mathcal{L}_m$ as $n \to \infty$ if it holds for any $m \in \mathbb{N}$ that $\lambda_m^{(n)}$ converges to λ_m as $n \to \infty$.

Lemma 41. *$(\mathcal{L}_m, \pi_{m,n})_{m\leq n}$ and $(\mathcal{A}_m, \pi_{m,n})_{m\leq n}$ are projective systems.*

Proof. The fact that $\mathcal{L}_m$ and $\mathcal{A}_m$ are Hausdorff topological spaces is a consequence of the equivalence of their topologies with the discrete topologies that we described in Section 4.1. For $m \leq n$, the continuity of $\pi_{m,n}$ is easily verified for the lowest level, i.e., on $\mathcal{S}_n \to \mathcal{S}_m$, and then lifted to the higher levels. Indeed, for each open set $\mathcal{O} \in \mathcal{M}(\mathcal{S}_m)$ (which can be identified as an open set in $\mathbb{R}_{\geq 0}^{\mathcal{S}_m} \setminus \{0\}$), we see that $\pi_{m,n}^{-1}(\mathcal{O})$ is an open set in $\mathcal{M}(\mathcal{S}_m)$. The same is true at the higher level with open sets in $\mathcal{L}_m$ and $\mathcal{A}_m$, which are images of open sets in $\mathcal{L}_n$ and $\mathcal{A}_n$. □

Throughout this section we will denote $\mathcal{L}_\infty = \varprojlim \mathcal{L}_m$ and $\mathcal{A}_\infty = \varprojlim \mathcal{A}_m$. The aim of this section is to prove the following

Proposition 42. *The following assertions hold.*

1. *The set $\mathcal{L} \times \mathcal{A}$ can be identified with $\mathcal{L}_\infty \times \mathcal{A}_\infty$.*
2. *The projective limit topology on $\mathcal{L}_\infty \times \mathcal{A}_\infty$ is equivalent to the vague topology on $\mathcal{L} \times \mathcal{A}$.*

The proof will be a consequence of the following lemmas. In Lemma 44 we explain how to project from $\mathcal{L} \times \mathcal{A}$ to $\mathcal{L}_\infty \times \mathcal{A}_\infty$. As a direct consequence of Lemma 45 we get that this operation is continuous. Afterwards, we deal with the inverse operation. In Lemma 46 we show the existence of the inverse and in Lemma 48 we argue that it is continous.

Here is a basic property of the mappings π_m and $\pi_{m,n}$, which we need to prepare for Lemma 44.

Lemma 43. *Let $m \leq n$. Then the equality $\pi_m = \pi_{m,n} \circ \pi_n$ holds on all levels, i.e., for all mappings that were defined in* (5.6)–(5.8) *and* (5.21)–(5.23)*.*

Proof. On the lowest level, i.e., on $\mathcal{S} \to \mathcal{S}_m$, the equality is a direct consequence of the fact that $(P_m)_{m\in\mathbb{N}}$ is nested. On the higher levels, the equality follows by the definition of the image measure. □

Lemma 44. *The following holds.*

1. *Let $\lambda \in \mathcal{L}$. Then for any $m \in \mathbb{N}$ we have that $\pi_m(\lambda) \in \mathcal{L}_m$. Further, we have that the sequence $(\pi_m(\lambda))_{m\in\mathbb{N}}$ is an element of the projective limit $\mathcal{L}_\infty$.*
2. *Let $\alpha \in \mathcal{A}$. Then for any $m \in \mathbb{N}$ we have that $\pi_m(\alpha) \in \mathcal{A}_m$. Further, we have that the sequence $(\pi_m(\alpha))_{m\in\mathbb{N}}$ is an element of the projective limit $\mathcal{A}_\infty$.*

Consequently, the mapping $\Pi\colon \mathcal{L} \times \mathcal{A} \to \mathcal{L}_\infty \times \mathcal{A}_\infty$, $(\lambda, \alpha) \mapsto \big((\pi_m(\lambda))_{m\in\mathbb{N}}, (\pi_m(\alpha)_{m\in\mathbb{N}})\big)$ is well-defined.

Proof. (1) For any $\lambda \in \mathcal{L}$ and $m \in \mathbb{N}$ we have that $c_{\pi_m(\lambda)} = \pi_m(c_\lambda) \leq \pi_m(\tilde{\mu})$, where $\tilde{\mu} \in \{\mu\} \cup \{\mu_N \colon N \in \mathbb{N}\}$, hence $\pi_m(\lambda) \in \mathcal{L}_m$. As a direct consequence of Lemma 43 the sequence $(\pi_m(\lambda))_{m\in\mathbb{N}}$ satisfies the consistency condition and is therefore an element of $\mathcal{L}_\infty$.

(2) For $\alpha \in \mathcal{A}$ the proof is analogous. □

Lemma 45. *Fix any $m \in \mathbb{N}$. The mappings $\pi_m \colon \mathcal{L} \to \mathcal{L}_m$ and $\pi_m \colon \mathcal{A} \to \mathcal{A}_m$ are continuous w.r.t. the vague topologies. Consequently, the mapping Π defined as in Lemma 44 is continuous.*

Proof. On the lowest level, i.e., for $\pi_m \colon \mathcal{S} \to \mathcal{S}_m$, it is obvious that π_m is continuous on the set $\mathcal{S} \setminus \bigcup_{i=1}^{|P_m|} \partial A_{m,i}$. Consider the second level, i.e., $\pi_m \colon \mathcal{M}(\mathcal{S}) \to \mathcal{M}(\mathcal{S}_m)$. We claim that π_m is continuous on

$$\mathcal{N}_0 := \{\nu \in \mathcal{M}(\mathcal{S}) \colon \nu(\partial A_{m,i}) = 0 \text{ for all } i = 1, \dots, |P_m|\} \tag{5.24}$$

w.r.t. weak convergence: Given some $\nu \in \mathcal{N}_0$ and a sequence $(\nu_n)_{n\in\mathbb{N}}$ in $\mathcal{M}(\mathcal{S})$ such that $\nu_n \to \nu$ weakly as $n \to \infty$ and some continuous bounded function $f \colon \mathcal{S}_m \to \mathbb{R}$ we clearly have that

$$\int_{\mathcal{S}_m} f \,\mathrm{d}\pi_m(\nu_n) = \int_{\mathcal{S}} f \circ \pi_m \,\mathrm{d}\nu_n \to \int_{\mathcal{S}} f \circ \pi_m \,\mathrm{d}\nu = \int_{\mathcal{S}_m} f \,\mathrm{d}\pi_m(\nu),$$

as $n \to \infty$, since $f \circ \pi_m$ is continuous ν-almost everywhere. Hence, $\pi_m(\nu_n) \to \pi_m(\nu)$ weakly as $n \to \infty$.

Now, consider $\pi_m \colon \mathcal{L} \to \mathcal{L}_m$. For $\lambda \in \mathcal{L}$ we have that $c_\lambda \leq \tilde{\mu}$, where $\tilde{\mu} \in \{\mu\} \cup \{\mu_N \colon N \in \mathbb{N}\}$. This implies that for any $i = 1, \dots, |P_m|$ we have that

$$\lambda\big(\{k \colon k(\partial A_{m,i}) > 0\}\big) \leq \sum_{n=1}^{\infty} n\lambda\big(\{k \colon k(\partial A_{m,i}) = n\}\big) = c_\lambda(\partial A_{m,i}) \leq \tilde{\mu}(\partial A_{m,i}) = 0 \tag{5.25}$$

by Lemma (39). With other words, λ is concentrated on a subset of the set $\mathcal{N}_0$ given in (5.24). Now, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{L}$ that converges vaguely to λ as $n \to \infty$. Then for any function $g \colon \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m) \to \mathbb{R}$ that is continuous and compactly supported, we have that

$$\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)} g \,\mathrm{d}\pi_m(\lambda_n) = \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} g \circ \pi_m \,\mathrm{d}\lambda_n \to \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} g \circ \pi_m \,\mathrm{d}\lambda = \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)} g \,\mathrm{d}\pi_m(\lambda),$$

as $n \to \infty$, since $g \circ \pi_m$ is compactly supported and continuous λ-almost everywhere. Hence, $\pi_m(\lambda_n) \to \pi_m(\lambda)$ vaguely as $n \to \infty$.

The proof for $\pi_m \colon \mathcal{A} \to \mathcal{A}_m$ is analogous. □

In the next lemmas we will deal with the construction of the inverse of the projection mapping Π and verify its continuity. This requires, for any $m \in \mathbb{N}$, to identify measures $\lambda_m \in \mathcal{L}_m$ with measures $\bar{\lambda}_m \in \mathcal{L}$. To prepare for this identification we now define for any $m \in \mathbb{N}$

$$\mathcal{M}_m(\mathcal{S}) := \{\nu \in \mathcal{M}(\mathcal{S}) \colon \nu \text{ is concentrated on } \mathcal{S}_m\}, \tag{5.26}$$

where for any measure ν on some measure space $\mathcal{X}$ and any measurable $\mathcal{U} \subset \mathcal{X}$ we say that ν is concentrated on $\mathcal{U}$ if $\nu(\mathcal{X} \setminus \mathcal{U}) = 0$. It is clear that $\mathcal{M}_m(\mathcal{S})$ can be identified with $\mathcal{M}(\mathcal{S}_m)$. Observe that this is possible because $\mathcal{S}_m \subset \mathcal{S}$, as we have defined it via the representatives of each partition. For any $\nu_m \in \mathcal{M}(\mathcal{S}_m)$ we will denote the corresponding element by $\bar{\nu}_m \in \mathcal{M}_m(\mathcal{S})$ and we will write $\bar{\pi}_m \colon \mathcal{M}(\mathcal{S}) \to \mathcal{M}_m(\mathcal{S})$ for the mapping that we obtain by concatenating $\pi_m \colon \mathcal{M}(\mathcal{S}) \to \mathcal{M}(\mathcal{S}_m)$

as defined in (5.7) with the operation $\nu_m \mapsto \bar{\nu}_m$. Then we can identify the space $\mathcal{L}_m$ that was defined in (5.10) with

$$\bar{\mathcal{L}}_m := \big\{\lambda \in \mathcal{L} \colon \lambda \text{ is concentrated on } \mathcal{M}_m(\mathcal{S}) \cap \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})\big\}. \tag{5.27}$$

For any $\lambda_m \in \mathcal{L}_m$ we will denote the corresponding element by $\bar{\lambda}_m \in \bar{\mathcal{L}}_m$ and we will write $\bar{\pi}_m \colon \mathcal{L} \to \bar{\mathcal{L}}_m$ for the mapping that we obtain by concatenating $\pi_m \colon \mathcal{L} \to \mathcal{L}_m$ with the operation $\lambda_m \mapsto \bar{\lambda}_m$. In the same way, we identify the space $\mathcal{A}_m$ that was defined in (5.11) with

$$\bar{\mathcal{A}}_m := \big\{\alpha \in \mathcal{A} \colon \alpha \text{ is concentrated on } \mathcal{M}_m(\mathcal{S}) \cap \mathcal{M}(\mathcal{S}) \setminus \{0\}\big\}. \tag{5.28}$$

Now we construct the inverse of the projection mapping Π that was defined in Lemma 44.

Lemma 46. *The following assertions hold.*

1. *Let $(\lambda_m)_{m\in\mathbb{N}} \in \mathcal{L}_\infty$. Then there exists a unique $\lambda \in \mathcal{L}$ such that $\lambda_m = \pi_m(\lambda)$ holds for all $m \in \mathbb{N}$.*
2. *Let $(\alpha_m)_{m\in\mathbb{N}} \in \mathcal{A}_\infty$. Then there exists a unique $\alpha \in \mathcal{A}$ such that $\alpha_m = \pi_m(\alpha)$ holds for all $m \in \mathbb{N}$.*

Consequently, the mapping Π defined in Lemma 44 is bijective with inverse Π^{-1}.

Proof. Fix $(\lambda_m)_{m\in\mathbb{N}} \in \mathcal{L}_\infty$. The idea is to identify for any $m \in \mathbb{N}$ the measure $\lambda_m \in \mathcal{L}_m$ uniquely with the element $\bar{\lambda}_m \in \bar{\mathcal{L}}_m$ and to prove that the sequence $(\bar{\lambda}_m)_{m\in\mathbb{N}}$ has a vague limit point in $\mathcal{L}$, which we will denote by λ. It then remains to show that $\pi_m(\lambda) = \lambda_m$ holds for any $m \in \mathbb{N}$.

Next, we will argue for the existence of a vague limit point of $(\bar{\lambda}_m)_{m\in\mathbb{N}}$. As an element of $\mathcal{L}_\infty$ the sequence $(\lambda_m)_{m\in\mathbb{N}}$ satisfies the consistency condition $\lambda_m = \pi_{m,n}(\lambda_n)$ for any $m \le n$. Abbreviating $\mathcal{M}_{\mathbb{N}_0} := \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \setminus \{0\}$ and $\mathcal{M}_{\mathbb{N}_0,m} := \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m) \setminus \{0\}$, we get that

$$\bar{\lambda}_m(\mathcal{M}_{\mathbb{N}_0}) = \lambda_m(\mathcal{M}_{\mathbb{N}_0,m}) = \lambda_n(\pi_{m,n}^{-1}(\mathcal{M}_{\mathbb{N}_0,m})) = \lambda_n(\mathcal{M}_{\mathbb{N}_0,n}) = \bar{\lambda}_n(\mathcal{M}_{\mathbb{N}_0})$$

and consequently, $\bar{\lambda}_m$, $m \in \mathbb{N}$, is of constant total variation. Note that the measures $\bar{\lambda}_m$, $m \in \mathbb{N}$, are in the dual of $C_0(\mathcal{M}_{\mathbb{N}_0})$ and that, due to the Banach–Alaoglu theorem, they are compact w.r.t. the weak*-topology, which implies compactness w.r.t. the vague topology on $\mathcal{L}$. Hence, there exists a vague limit point $\lambda \in \mathcal{M}(\mathcal{M}_{\mathbb{N}_0})$ and a subsequence $(\bar{\lambda}_{m_i})_{i\in\mathbb{N}}$ in $\mathcal{L}$ converging vaguely to λ. Since $\mathcal{L}$ is compact by Lemma 38, we also have that $\lambda \in \mathcal{L}$.

Next, we fix $m \in \mathbb{N}$ and our goal is to show that $\lambda_m = \pi_m(\lambda)$. Observe that as a consequence of our identification between λ_n and $\bar{\lambda}_n$ we have that $\pi_n(\bar{\lambda}_n) = \lambda_n$ for any $n \in \mathbb{N}$. Together with the consistency and Lemma 43 we get for $n \ge m$ that

$$\lambda_m = \pi_{m,n}(\lambda_n) = \pi_{m,n}(\pi_n(\bar{\lambda}_n)) = \pi_m(\bar{\lambda}_n).$$

Choosing a subsequence $(\bar{\lambda}_{n_i})_{i\in\mathbb{N}}$ that converges vaguely to λ, we get that $\lambda_m = \lim_{i\to\infty} \pi_m(\bar{\lambda}_{n_i}) = \pi_m(\lambda)$ where we used the continuity of the mapping π_m that we showed in Lemma 45.

For a given $(\alpha_m)_{m\in\mathbb{N}} \in \mathcal{A}_\infty$ the proof is analogous. □

To prepare for the proof of the continuity of Π^{-1} we need the following lemma.

Lemma 47. *On bounded subsets of $\mathcal{M}(\mathcal{S})$, the mapping $\mathcal{M}(\mathcal{S}) \to \mathcal{M}(\mathcal{S})$, $\nu \mapsto \bar{\pi}_m(\nu)$, converges uniformly to the identity as $m \to \infty$.*

Proof. Recall that we equip $\mathcal{M}(\mathcal{S})$ with the weak topology, which is generated by all the test integrals against continuous bounded functions $\mathcal{S} \to \mathbb{R}$. The weak topology on $\mathcal{M}(\mathcal{S})$ admits a number of metrisations, especially since $\mathcal{S}$ is compact. We introduce the dual bounded-Lipschitz distance given by

$$d_{\mathrm{BL}}(\nu, \hat{\nu}) := \sup_{\phi\colon \|\phi\|_{\mathrm{BL}} \le 1} \left| \int_{\mathcal{S}} \phi \,\mathrm{d}\nu - \int_{\mathcal{S}} \phi \,\mathrm{d}\hat{\nu} \right|, \tag{5.29}$$

where $\|\phi\|_{\mathrm{BL}} = \|\phi\|_\infty + \mathrm{Lip}(\phi)$ and $\mathrm{Lip}(\phi)$ is the infimum of all Lipschitz constants of $\phi\colon \mathcal{S} \to \mathbb{R}$. Let $\phi\colon \mathcal{S} \to \mathbb{R}$ satisfy $\|\phi\|_{\mathrm{BL}} \le 1$ then

$$\left| \int_{\mathcal{S}} \phi \,\mathrm{d}\pi_m(\nu) - \int_{\mathcal{S}} \phi \,\mathrm{d}\nu \right| \le \sum_{i=1}^{|P_m|} \int_{A_{m,i}} |\phi(x_{m,i}) - \phi(x)| \,\nu(\mathrm{d}x) \le \max_{i=1,\dots,|P_m|} \mathrm{diam}(A_{m,i}) \nu(\mathcal{S}), \tag{5.30}$$

where $x_{m,i}$ is the representative of $A_{m,i}$ as defined right before definition (5.5). Now, if $\mathcal{N} \subset \mathcal{M}(\mathcal{S})$ is bounded, i.e., $\sup_{\nu \in \mathcal{N}} \nu(\mathcal{S}) < \infty$, then $d_{\mathrm{BL}}(\nu, \pi_m(\nu))$ vanishes as $m \to \infty$ uniformly on $\mathcal{N}$ by assumption (5.4). □

In the next lemma we verify that the mapping Π^{-1} constructed in Lemma 46 is continuous.

Lemma 48. *The following assertions hold.*

1. *Let $\lambda^{(n)}$, $n \in \mathbb{N}$, be a sequence in $\mathcal{L}$. Assume for all $m \in \mathbb{N}$ that $\pi_m(\lambda^{(n)})$ converges to $\pi_m(\lambda)$ vaguely in $\mathcal{L}_m$ as $n \to \infty$. Then $\lambda^{(n)}$ converges vaguely to λ.*
2. *Let $\alpha^{(n)}$, $n \in \mathbb{N}$, be a sequence in $\mathcal{A}$. Assume for all $m \in \mathbb{N}$ that $\pi_m(\alpha^{(n)})$ converges to $\pi_m(\alpha)$ vaguely in $\mathcal{A}_m$ as $n \to \infty$. Then $\alpha^{(n)}$ converges vaguely to α.*

Consequently, the mapping Π^{-1} constructed in Lemma 46 is continuous.

Proof. (1) Let $\lambda^{(n)}$, $n \in \mathbb{N}$, be a sequence in $\mathcal{L}$, such that for all $m \in \mathbb{N}$ it holds that $\pi_m(\lambda^{(n)})$ converges to $\pi_m(\lambda)$ vaguely in $\mathcal{L}_m$ as $n \to \infty$. Abbreviate $\mathcal{M}_{\mathbb{N}_0} := \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}) \setminus \{0\}$ and fix a continuous and compactly supported function $g\colon \mathcal{M}_{\mathbb{N}_0} \to \mathbb{R}$. Recall the identification of $\mathcal{L}_m$ and $\bar{\mathcal{L}}_m$ introduced before Lemma 46. Clearly, we have for any $m \in \mathbb{N}$ that

$$\begin{aligned} \Big| \int g \,\mathrm{d}\lambda^{(n)} - \int g \,\mathrm{d}\lambda \Big| &\le \left| \int g \,\mathrm{d}\lambda^{(n)} - \int g \,\mathrm{d}\bar{\pi}_m(\lambda^{(n)}) \right| \\ &\quad + \left| \int g \,\mathrm{d}\bar{\pi}_m(\lambda^{(n)}) - \int g \,\mathrm{d}\bar{\pi}_m(\lambda) \right| + \left| \int g \,\mathrm{d}\bar{\pi}_m(\lambda) - \int g \,\mathrm{d}\lambda \right| \\ &\le \big(|\lambda^{(n)}| + |\lambda|\big) \|g - g \circ \bar{\pi}_m\|_\infty + \left| \int g \,\mathrm{d}\bar{\pi}_m(\lambda^{(n)}) - \int g \,\mathrm{d}\bar{\pi}_m(\lambda) \right|. \end{aligned} \tag{5.31}$$

Note that $\lambda^{(n)} \in \mathcal{L}$ implies that $|\lambda^{(n)}| \le c_{\lambda^{(n)}}(\mathcal{S}) \le \sup_{N \in \mathbb{N}} \mu_N(\mathcal{S}) \vee \mu(\mathcal{S}) = 1$ for all $n \in \mathbb{N}$, and in the same way $|\lambda| \le 1$. Further, the support of the function g is compact, and thus bounded. So by Lemma 47 we have that the mappings $\bar{\pi}_m$ restricted to the support of g converge uniformly to the identity, as $m \to \infty$. Hence, we can first choose $m \in \mathbb{N}$ sufficently large such that $\big(|\lambda^{(n)}| + |\lambda|\big) \|g - g \circ \bar{\pi}_m\|_\infty$ is arbitrarily small, uniformly in n. Then we can use that, $\bar{\pi}_m(\lambda^{(n)}) \to \bar{\pi}_m(\lambda)$, as $n \to \infty$, holds by assumption, so the second summand on the right-hand side of (5.31) vanishes as $n \to \infty$.

(2) Let $\alpha^{(n)}$, $n \in \mathbb{N}$, be a sequence in $\mathcal{A}$, such that for all $m \in \mathbb{N}$ it holds that $\pi_m(\alpha^{(n)})$ converges to $\pi_m(\alpha)$ vaguely in $\mathcal{A}_m$ as $n \to \infty$. Recall that for all $n \in \mathbb{N}$ the measures $\alpha^{(n)}$ are concentrated on $\mathcal{M}_{\leq 1}(\mathcal{S})$, i.e., on subprobability measures on $\mathcal{S}$, since for any $y \in \operatorname{supp}(\alpha^{(n)})$, we have that $y(\mathcal{S}) \leq c_{\alpha^{(n)}}(\mathcal{S}) \leq \sup_{N\in\mathbb{N}} \mu_N(\mathcal{S}) \vee \mu(\mathcal{S}) = 1$. The same holds for α. Hence, without loss of generality we can show vague convergence by considering any continuous compactly supported test function $g\colon \mathcal{M}_{\leq 1}(\mathcal{S}) \setminus \{0\} \to \mathbb{R}$. By Lemma 37 there exists $\varepsilon > 0$ such that the support of g is contained in $N_\varepsilon := \{\nu \in \mathcal{M}_{\leq 1}(\mathcal{S}) \setminus \{0\}\colon \nu(\mathcal{S}) \geq \varepsilon\}$. Analogously to (5.31) we get

$$\left|\int g \,\mathrm{d}\alpha^{(n)} - \int g \,\mathrm{d}\alpha\right| \leq \\ \big(\alpha^{(n)}(N_\varepsilon) + \alpha(N_\varepsilon)\big) \sup_{y\in N_\varepsilon} |g(y) - g\circ\bar{\pi}_m(y)| + \left|\int g \,\mathrm{d}\bar{\pi}_m(\alpha^{(n)}) - \int g \,\mathrm{d}\bar{\pi}_m(\alpha)\right|. \tag{5.32}$$

To bound the first summand observe the following: Since $\alpha^{(n)} \in \mathcal{A}$ one has that $\varepsilon\alpha^{(n)}(N_\varepsilon) \leq c_{\alpha^{(n)}}(\mathcal{S}) \leq 1$ and hence $\alpha^{(n)}(N_\varepsilon) \leq 1/\varepsilon$. The same holds for α and hence $\alpha^{(n)}(N_\varepsilon) + \alpha(N_\varepsilon) \leq 2/\varepsilon$. Also, by Lemma 47 the supremum in (5.32) vanishes, as $m \to \infty$. So, by first choosing m large enough and then using that $\bar{\pi}_m(\alpha^{(n)}) \to \bar{\pi}_m(\alpha)$, as $n \to \infty$, the right-hand side of (5.32) vanishes. □

5.3 Dawson–Gärtner and identification of the rate function

As an easy consequence of Lemma 40 and the LDP of Theorem 29, we obtain upper and lower LDP bounds for $(\mathrm{Mi}_N^{(m)}, \mathrm{Ma}_N^{(m)})$ with different rate functions. In order to formulate this (in particular, to identify the rate functions) we need to introduce additional notation.

For dealing with the microscopic clusters, we need the discretized version of the connection parameter τ defined in (3.2), which now has to be understood with respect to the discretized kernels. For $* \in \{+,-\}$ and $k \in \mathcal{M}(\mathcal{S}_m) \setminus \{0\}$ we write

$$\tau_m^{(*)}(k) = \sum_{T\in\mathcal{T}(k)} \prod_{\{i,j\}\in E(T)} \kappa^{(m,*)}(r_i, r_j), \tag{5.33}$$

where $(r_i)_{i=1,\dots,|k|} \in \mathcal{S}_m^{|k|}$ is such that $k = \sum_{i=1}^{|k|} \delta_{r_i} \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}_m)$ and $\mathcal{T}(k)$ is the set of spanning trees on $[|k|]$ and we recall that $|k| = k(\mathcal{S}_m)$. Further, we define for $* \in \{+,-\}$

$$C_m^{(*)} = \frac{1}{2}\langle \mu^{(m)}, \kappa^{(m,*)}\mu^{(m)}\rangle, \tag{5.34}$$

$$\hat{I}_{\mathrm{Mi}}^{(m,*)}(\lambda) = \mathbb{H}(\lambda|\mathbb{Q}_{\mu^{(m)}}) - 1 - \langle \lambda, \log \tau_m^{(*)}\rangle + |c_\lambda| - |\lambda|, \quad \lambda \in \mathcal{L}_m, \tag{5.35}$$

$$\hat{I}_{\mathrm{Me}}^{(m,*)}(\nu) = \left\langle \nu, \log \frac{\mathrm{d}\nu}{(\kappa^{(m,*)}\nu)\,\mathrm{d}\mu^{(m)}}\right\rangle, \quad \nu \in \mathcal{M}(\mathcal{S}_m). \tag{5.36}$$

where we recall that $\mathbb{Q}_{\mu^{(m)}}$ is the distribution of a Poisson point process on $\mathcal{S}_m$ with intensity measure $\mu^{(m)}$ and that $\mathbb{H}$ denotes relative entropy between non-normalized measures, defined as in (1.8). Again, we adopt the convention that $I_{\mathrm{Me}}^{(m,*)}(\nu) = \infty$ if $\frac{\mathrm{d}\nu}{(\kappa^{(m,*)}\nu)\,\mathrm{d}\mu^{(m)}}$ does not exist. We have to be more careful in the definition of the macroscopic rate. For $\alpha \in \mathcal{A}_m$ we define

$$\hat{I}_{\mathrm{Ma}}^{(m,*)}(\alpha) = \begin{cases} \int_{\mathcal{M}(\mathcal{S}_m)\setminus\{0\}} \alpha(\mathrm{d}y) \left[\left\langle y, \log \frac{\mathrm{d}y}{(1-\mathrm{e}^{-\kappa^{(m,*)}y})\,\mathrm{d}\mu^{(m)}}\right\rangle - \frac{1}{2}\langle y, \kappa^{(m,*)}y\rangle\right] & \text{if } \alpha \text{ is connectable,} \\ \infty & \text{otherwise,} \end{cases} \tag{5.37}$$

where again we define $I_{\mathrm{Ma}}(\alpha) = \infty$, if it is not true that α-almost everywhere the density in the $\log$-term exists. The definition of $\hat{I}_{\mathrm{Ma}}$ also takes into account the possibility that the discretized kernel might not be irreducible. In particular if κ is irreducible, it is not always true that $\kappa^{(m,-)}$ is also irreducible. In that case we have to additionally assume that α is connectable with respect to $\kappa^{(m,*)}$ to get a finite rate, as it is formulated in the generalized version of Theorem 29, i.e., Theorem 30. We will comment on this in detail below.

When estimating the distribution of the pair $(\mathrm{Mi}^{(m)}, \mathrm{Ma}^{(m)})$ under the measure $\mathbb{P}_N$ by Lemma 40 we will get an additional error term. To deal with that we define for $(\lambda, \alpha) \in \mathcal{L}_m \times \mathcal{A}_m$

$$I^{(m,+)}(\lambda,\alpha) = \begin{cases} \hat{I}^{(m,+)}_{\mathrm{Mi}}(\lambda) + \hat{I}^{(m,+)}_{\mathrm{Ma}}(\alpha) + \hat{I}^{(m,+)}_{\mathrm{Me}}(\mu^{(m)} - c_\lambda - c_\alpha) + C^{(-)}_m & \text{if } c_\lambda + c_\alpha \le \mu^{(m)}, \\ \infty & \text{otherwise,} \end{cases} \tag{5.38}$$

$$I^{(m,-)}(\lambda,\alpha) = \begin{cases} \hat{I}^{(m,-)}_{\mathrm{Mi}}(\lambda) + \hat{I}^{(m,-)}_{\mathrm{Ma}}(\alpha) + \hat{I}^{(m,-)}_{\mathrm{Me}}(\mu^{(m)} - c_\lambda - c_\alpha) + C^{(+)}_m & \text{if } c_\lambda + c_\alpha \le \mu^{(m)}, \\ \infty & \text{otherwise.} \end{cases} \tag{5.39}$$

Corollary 49 (LDP bounds for $(\mathrm{Mi}^{(m)}_N, \mathrm{Ma}^{(m)}_N)$ under $\mathcal{G}_N$)**.** *Assume that μ_N converges to μ as $N \to \infty$. Let κ_N converge to a continuous kernel κ that is irreducible w.r.t. μ. Fix $m \in \mathbb{N}$ and let $\kappa^{(m,-)}_N$ and $\kappa^{(m,+)}_N$, $N \in \mathbb{N}$, be two sequences of kernels on $\mathcal{S}_m$ satisfying* (5.12)*. Then the distribution of $(\mathrm{Mi}^{(m)}_N, \mathrm{Ma}^{(m)}_N)$ under $\mathbb{P}_N$ satisfies, as $N \to \infty$, the upper large-deviations bound with rate function $I^{(m,+)}$ and the lower large-deviations bound with rate function $I^{(m,-)}$.*

Proof. It is easy to verify that

$$\lim_{N\to\infty} \tfrac{1}{N} \log \Delta^{(m)}_N = \tfrac{1}{2}\langle \mu^{(m)}, \kappa^{(m,+)}\mu^{(m)}\rangle - \tfrac{1}{2}\langle \mu^{(m)}, \kappa^{(m,-)}\rangle = C^{(+)}_m - C^{(-)}_m,$$

where $\Delta^{(m)}_N$ is defined as in (5.18). Note that irreducibility of κ implies irreducibility of $\kappa^{(m,+)}$. Hence, for the upper large-deviations bound we can apply Theorem 29 after using the upper bound from Lemma 40. For the lower bound we use the lower bound from Lemma 40 and the lower bound from Theorem 30 that also applies if $\kappa^{(m,-)}$ is reducible. □

Of course, the basic idea is to use Corollary 49 for very large m and the hope is that the discretized rate functions approximate the rate function I given in Theorem 1. However, the problem is that the lower bound can be arbitrarily bad due to the following issue. Observe that although κ is assumed to be irreducible with respect to μ, as a consequence of the definition of the lower approximation $\kappa^{(m,-)}$ given in (5.13) we have to deal with the possibility that $\kappa^{(m,-)}$ is not irreducible with respect to $\mu \circ \pi_m^{-1}$. This might even be the case for all $m \in \mathbb{N}$. To illustrate this, we give a brief example.

Example 50. *Choose $\mathcal{S} = [0,1]$, let μ be the Lebesgue measure and $\kappa_N(x,x') = \kappa(x,x') = xx'$ for $x, x' \in \mathcal{S}$ and all $N \in \mathbb{N}$. Let $\{P_m\}_{m\in\mathbb{N}}$ be any nested partition for $\mathcal{S}$ and let $\mathcal{S}_m$, $m \in \mathbb{N}$, be any choice for the sets of representative points. Then for any $m \in \mathbb{N}$ there exists some set $A_m \in P_m$ such that $0 \in \partial A_m$ and $\mu(A_m) > 0$. Let x_m be the representative of A_m. It can be directly verified that for any $x'_m \in \mathcal{S}_m$ we have that $\kappa^{(m,-)}(x_m, x'_m) = 0$, although $\mu^{(m)}(\{x_m\}) = \mu(A_m) > 0$. On the other hand, assumption* (5.4) *implies that $\mu(A_m) \to 0$ as $m \to \infty$ so there exists m_0 such that for $m \ge m_0$ we also have that $\mu^{(m)}(\mathcal{S}_m \setminus \{x_m\}) > 0$ and hence $\kappa^{(m,-)}$ is reducible with respect to $\mu^{(m)}$ for all $m \ge m_0$.* ◊

Regarding the approximation of the rate function I by its discrete version the problem of (missing) irreducibility enters our analysis only via the function $\hat{I}^{(m,-)}_{\mathrm{Ma}}$. Indeed, for $\hat{I}^{(m,+)}_{\mathrm{Ma}}$ the additional case

distinction given in (5.37) is not necessary, since $\kappa^{(m,+)}$ is always irreducible and hence every α is connectable with respect to $\kappa^{(m,+)}$. In order to make the lower approximation work, we formulate additional assumptions on α. Since they are not satisfied for any α we have to find a way how to deal with the other cases.

Lemma 51 (Identification of the rate function)**.** *Let $\kappa\colon \mathcal{S}\times\mathcal{S}\to[0,\infty)$ be continuous and irreducible with respect to μ. Assume that $\kappa^{(m,+)},\kappa^{(m,-)}\colon \mathcal{S}_m\times\mathcal{S}_m\to[0,\infty)$ are given such that $\kappa^{(m,-)}\le\kappa^{(m,+)}$ and*

$$\lim_{m\to\infty}\kappa^{(m,*)}(\pi_m(x),\pi_m(y))=\kappa(x,y),\qquad \text{uniformly in } x,y\in\mathcal{S},*\in\{+,-\},$$

monotonously decreasing for $*=+$ and monotonously increasing for $*=-$. We abbreviate $\Pi_m(\lambda,\alpha):=(\pi_m(\lambda),\pi_m(\alpha))$ for $m\in\mathbb{N}$, $(\lambda,\alpha)\in\mathcal{L}\times\mathcal{A}$. Then the rate function I introduced in Theorem 1 satisfies the following.

1. *For any $(\lambda,\alpha)\in\mathcal{L}\times\mathcal{A}$ it holds that $I^{(m,+)}(\Pi_m(\lambda,\alpha))\nearrow I(\lambda,\alpha)$ as $m\to\infty$.*
2. *Let $(\lambda,\alpha)\in\mathcal{L}\times\mathcal{A}$ and assume that α satisfies the following: there exists some $m_0\in\mathbb{N}$ such that for all $m\ge m_0$ the measures $\pi_m(\alpha)$ are connectable with respect to $\kappa^{(m,-)}$. Then $I^{(m,-)}(\Pi_m(\lambda,\alpha))\to I(\lambda,\alpha)$ as $m\to\infty$.*

Proof. Fix $\lambda\in\mathcal{L}$ and $\alpha\in\mathcal{A}$. Also, we will denote $\lambda_m=\pi_m(\lambda)$ and $\alpha_m=\pi_m(\alpha)$.

We first assume that $c_\lambda+c_\alpha\le\mu$. It is straight forward to show that this implies that $c_{\lambda_m}+c_{\alpha_m}\le\mu^{(m)}$ for all $m\in\mathbb{N}$. We denote $\nu_m=\mu^{(m)}-c_{\lambda_m}-c_{\alpha_m}$.

We will see in the proof that in each of the terms that we handle, the main part is an entropy between two image measures under π_m plus a perturbation and other terms that will turn out to converge monotonically as $m\to\infty$. Then we will use [Geo88, Prop. 15.6], which says that the named entropy converges, as $m\to\infty$, to its supremum over m.

For this, we will handle each of the four terms in (5.34)–(5.36) separately.

Step 1: term $C_m^{(*)}$. It is easy to deduce that $\lim_{m\to\infty}C_m^{(+)}=\frac{1}{2}\langle\mu,\kappa\mu\rangle=\lim_{m\to\infty}C_m^{(-)}$ and that $C_m^{(*)}$ is decreasing in m for $*=+$ and increasing for $*=-$. Note that $I^{(m,+)}$ is defined with $C_m^{(-)}$, so it has the right direction of monotonicity.

Step 2: term $\hat{I}_{\mathrm{Mi}}^{(m,*)}(\lambda_m)$. Let us turn to the part $\hat{I}_{\mathrm{Mi}}^{(m,*)}(\lambda_m)$. Recall that $\lambda_m=\lambda\circ\pi_m^{-1}$.

Since $\lambda_m=\lambda\circ\pi_m^{-1}$ and $\mathbb{Q}_{\mu^{(m)}}=\mathbb{Q}_\mu\circ\pi_m^{-1}$ by the mapping theorem for Poisson point processes, [Geo88, Prop. 15.6] implies that $\mathbb{H}(\lambda_m|\mathbb{Q}_{\mu^{(m)}})$ converges towards $\mathbb{H}(\lambda|\mathbb{Q}_\mu)$, and $\mathbb{H}(\lambda_m|\mathbb{Q}_{\mu^{(m)}})$ is increasing in m. According to our assumption in (51), using the monotone convergence theorem, we see that the term $-\langle\lambda_m,\log\tau_m^{(*)}\rangle$ converges towards $-\langle\lambda,\log\tau\rangle$ (also if $-\langle\lambda,\log\tau\rangle=\infty$). For $*=+$ the convergence is from below, as desired. It is easy to verify that $|c_{\lambda_m}|-|\lambda_m|=|c_\lambda|-|\lambda|$ holds for all m. Hence, we have shown that $\hat{I}_{\mathrm{Mi}}^{(m,+)}(\pi_m(\lambda))\nearrow\hat{I}_{\mathrm{Mi}}(\lambda)$ and $\hat{I}_{\mathrm{Mi}}^{(m,-)}(\pi_m(\lambda))\to\hat{I}_{\mathrm{Mi}}(\lambda)$ as $m\to\infty$.

Step 3: term $\hat{I}_{\mathrm{Me}}^{(m,*)}(\nu_m)$. Now we turn to the mesoscopic term $\hat{I}_{\mathrm{Me}}^{(m,*)}(\nu_m)$. Note that for $\nu=\mu-c_\lambda-c_\alpha$ the definition of the image measure implies that $\nu_m=\nu\circ\pi_m^{-1}$. Further, we have that

$$\hat{I}_{\mathrm{Me}}^{(m,*)}(\nu_m)=\Big\langle\nu_m,\log\frac{\nu_m}{\mu^{(m)}}\Big\rangle-\langle\nu_m,\log(\kappa^{(m,*)}\nu_m)\rangle.$$

Hence, we may apply [Geo88, Prop. 15.6] to the first term and see that $\langle\nu_m,\log\frac{\mathrm{d}\nu_m}{\mathrm{d}\mu^{(m)}}\rangle$ converges as $m\to\infty$ to its supremum on m, and for the second term we use (51) to see that

$$\lim_{m\to\infty}-\langle\nu_m,\log(\kappa^{(m,*)}\nu_m)\rangle=-\langle\nu,\log(\kappa\nu)\rangle$$

holds by the monotone convergence theorem (also if $-\langle \nu, \log(\kappa\nu)\rangle = \infty$). For $* = +$ the sequence is increasing as desired.

Step 4: term $\hat{I}^{(m,*)}_{\mathrm{Ma}}(\alpha_m)$. Finally, we turn to the macroscopic term $\hat{I}^{(m,*)}_{\mathrm{Ma}}(\alpha_m)$. Note that irreducibility of κ with respect to μ implies irreducibility of $\kappa^{(m,+)}$ with respect to $\mu^{(m)}$. Hence the measures $\pi_m(\alpha)$, $m \in \mathbb{N}$, are connectable. In the setting of statement (2) this is true for $m \geq m_0$ by assumption. Therefore, without loss of generality, we assume $m \geq m_0$. Applying the definition of the image measure we have that

$$\hat{I}^{(m,*)}_{\mathrm{Ma}}(\alpha_m) = \int \alpha(\mathrm{d}y) \left\langle y_m, \log \frac{\mathrm{d}y_m}{\mathrm{d}\mu^{(m)}} \right\rangle - \int \alpha(\mathrm{d}y) \Big[\langle y_m, \log(1 - \mathrm{e}^{-\kappa^{(m,*)} y_m}) \rangle + \frac{1}{2} \langle y_m, \kappa^{(m,*)} y_m \rangle \Big],$$

where we wrote $y_m = y \circ \pi_m^{-1}$. Now the convergence (including the desired monotonicity) can be argued in the same way as for the microscopic and the mesoscopic terms above, using the monotonicity of the entropy and our assumption in (51).

This finishes the proof of (1) and (2) in the case $c_\lambda + c_\alpha \leq \mu$.

Finally, we are considering the case that $\lambda \in \mathcal{L}$ and $\alpha \in \mathcal{A}$ do not satisfy $c_\lambda + c_\alpha \leq \mu$. Then there exists a μ-continuity set $A \subset \mathcal{S}$ such that $c_\lambda(A) + c_\alpha(A) > \mu(A)$. Then with $A_m = \pi_m(A)$ we have that $\pi_m^{-1}(A_m) \supseteq A$ and hence $c_{\lambda_m}(A_m) + c_{\alpha_m}(A_m) \geq c_\lambda(A) + c_\alpha(A)$. Now, let $0 < \varepsilon < c_\lambda(A) + c_\alpha(A) - \mu(A)$. By Lemma 47 we have that $\bar{\pi}_m(\mu) \to \mu$ weakly, as $m \to \infty$. Hence, we can choose m_0 large enough such that $\mu(A) \geq \bar{\pi}_m(\mu)(A) - \varepsilon/2 = \pi_m(\mu)(A_m) - \varepsilon/2$ holds for all $m \geq m_0$. Consequently, $c_{\lambda_m}(A_m) + c_{\alpha_m}(A_m) \geq \pi_m(\mu)(A_m) + \varepsilon/2$. Therefore $I^{(m,*)}(\Pi_m(\lambda, \alpha)) = \infty$ for any $m \geq m_0$ and $* \in \{+, -\}$. □

Now we derive an upper and a lower bound LDP for the distribution of $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ under $\mathcal{G}_N$ by following the two parts of the proof of the Dawson–Gärtner theorem, [DZ10, Theorem 4.6.1] and using the fact that due to Proposition 42 it is sufficient to work with open and closed sets from the projective limit topology.

Lemma 52 (LDP – upper bound)**.** *Suppose that all assumptions of Theorem 1 are satisfied. Then the distribution of $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ under $\mathbb{P}_N$ satisfies the upper bound part of the LDP with rate function I as defined in Theorem 1.*

Proof. Fix a set $F \subset \mathcal{L} \times \mathcal{A}$ that is closed with respect to the vague topology. Then by Proposition 42 the set F is also closed with respect to the projective limit topology. For any $m \in \mathbb{N}$ we use the notation $\Pi_m(\lambda, \alpha) := (\pi_m(\lambda), \pi_m(\alpha))$ and recall that $(\mathrm{Mi}^{(m)}_N, \mathrm{Ma}^{(m)}_N) = \Pi_m(\mathrm{Mi}_N, \mathrm{Ma}_N)$. Therefore,

$$\begin{aligned} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big((\mathrm{Mi}_N, \mathrm{Ma}_N) \in F\big) &\leq \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big((\mathrm{Mi}^{(m)}_N, \mathrm{Ma}^{(m)}_N) \in \Pi_m(F)\big) \\ &\leq -\inf_{\Pi_m(F)} I^{(m,+)} = -\inf_F I^{(m,+)} \circ \Pi_m, \end{aligned}$$

where we used Corollary 49 for the second inequality. Since the left-hand side does not depend on m, we can proceed with the supremum over $m \in \mathbb{N}$ on the right-hand side. By Lemma 51 we have that $\sup_m I^{(m,+)} \circ \Pi_m = I$, which implies the claim. □

It remains to prove the lower bound of the LDP formulated in Theorem 1. Since the approximation of I from below via $I^{(m,-)}$ works only in the case where α satisfies the assumptions given in the second statement of Lemma 51, we have to do some additional work. The idea is the following: given some α that does not fulfill the assumptions, we will first approximate α by some suitable choice for which the assumptions hold. Then we can apply Lemma 51.

Lemma 53. *Let $\alpha \in \mathcal{A}$ with $c_\alpha \leq \mu$ and $I_{\mathrm{Ma}}(\alpha) < \infty$. Then there exists a sequence $(\alpha^{(\delta,\varepsilon)})_{\delta>0,\varepsilon>0}$ in $\mathcal{A}$ such that the following properties hold:*

(1) for fixed $\delta > 0$ and $\varepsilon > 0$ there exists $m_0 = m_0(\delta)$ such that for all $m \geq m_0$ the measures $\pi_m(\alpha^{(\delta,\varepsilon)})$ are connectable with respect to $\kappa^{(m,-)}$;

(2) $\alpha^{(\delta,\varepsilon)} \to \alpha$ as $\delta \to 0$ and $\varepsilon \to 0$ with respect to the vague topology;

(3) for any $\lambda \in \mathcal{L}$ we have that $I(\lambda, \alpha^{(\delta,\varepsilon)}) \to I(\lambda, \alpha)$ as $\delta \to 0$ and $\varepsilon \to 0$.

Proof. We always write $\alpha = \sum_{i\in J} \delta_{y_i}$, where J is a countable set. The idea is to pick some type $x \in \mathcal{S}$ and to restrict the measures $y_i \in \mathcal{M}_{\leq 1}(\mathcal{S}) \setminus \{0\}$, $i \in J$, to a subset S_δ of $\mathcal{S}$ that contains all types $x' \in \mathcal{S}$ that can be connected to x by using a finite sequence of intermediate types x_{h-1}, x_h for which we have $\kappa(x_{h-1}, x_h) \geq \delta$. Then for large enough m connectivity is preserved with respect to $\kappa^{(m,-)}$, which will imply (1). The parameter $\varepsilon > 0$ is only introduced to deal with the fact that J might be infinite and ensures the convergence claimed in (3).

Fix some $x \in \mathrm{supp}(\mu)$. For $\delta > 0$ define

$$S_\delta := S_\delta(x) := \{x' \in \mathcal{S} \colon \exists k \in \mathbb{N},\, \exists x_0, x_1, \ldots, x_k \in \mathcal{S},\, x_0 = x,\, x_k = x',\, \kappa(x_{h-1}, x_h) \geq \delta,\, \forall h \in [k]\}. \tag{5.40}$$

We first show that $\mu(\mathcal{S} \setminus S_\delta) \to 0$ as $\delta \to 0$, which we will need for the proof of (2) and (3). Observe that $S_\delta \subset S_{\delta'}$ if $\delta > \delta'$ and put $S_0 := \bigcup_{\delta>0} S_\delta$. Since κ is irreducible with respect to μ and $x \in \mathrm{supp}(\mu)$ it is easy to see that for δ small enough we have that $\mu(S_\delta) > 0$ and hence $\mu(S_0) > 0$. We now argue that $\kappa = 0$ μ-almost everywhere on $S_0 \times \mathcal{S} \setminus S_0$. Assume the contrary, i.e., $\int_{S_0} \int_{\mathcal{S}\setminus S_0} \kappa(x', x'')\mu(\mathrm{d}x')\mu(\mathrm{d}x'') > 0$. Then by continuity of κ we find sets of positive measure $A \subset S_0$ and $B \subset \mathcal{S} \setminus S_0$ where $\delta' := \inf_{x'\in A, x''\in B} \kappa(x', x'') > 0$, which is a contradiction to the fact that $B \subset \mathcal{S} \setminus S_0$. By the irreducibility of κ with respect to μ, the facts that $\mu(S_0) > 0$ and $\kappa = 0$ holds μ-almost everywhere on $S_0 \times \mathcal{S} \setminus S_0$ imply that $\mu(\mathcal{S} \setminus S_0) = 0$, so by continuity of measures we have that $\mu(\mathcal{S} \setminus S_\delta) \to 0$ as $\delta \to 0$.

Now for any $\delta \geq 0$ and $\varepsilon \geq 0$ we define

$$\alpha^{(\delta,\varepsilon)} := \sum_{i\in J_\varepsilon} \delta_{y_i^{(\delta)}}, \quad \text{where} \quad y^{(\delta)} := y(\cdot \cap S_\delta) \text{ for any } y \in \mathcal{M}_{\leq 1}(\mathcal{S}) \tag{5.41}$$

and where $J_\varepsilon = \{i \in J \colon |y_i| > \varepsilon\}$. Note that $|c_\alpha| \leq 1$ implies that for $\varepsilon > 0$ the set J_ε is finite.

Now we show that $(\alpha^{(\delta,\varepsilon)})_{\delta,\varepsilon>0}$ has the three properties.

(1) Fix $\varepsilon > 0$ and $\delta > 0$. Now take m_0 such that $\|\kappa - \kappa^{(m,-)} \circ \pi_m\|_\infty \leq \delta/2$ holds for all $m \geq m_0$. Let $m \geq m_0$. Then we have that $\kappa^{(m,-)} \circ \pi_m$ is irreducible on S_δ, since $\kappa(x_{i-1}, x_i) \geq \delta$ implies $\kappa^{(m,-)} \circ \pi_m(x_{i-1}, x_i) \geq \delta/2$. With other words, $\kappa^{(m,-)}$ is irreducible on $\pi_m^{-1}(S_\delta)$. For any $i \in J$ we have that $\mathrm{supp}(y_i^{(\delta)}) \subset S_\delta$ by construction and hence $\mathrm{supp}(\pi_m(y_i^{(\delta)})) \subset \pi_m^{-1}(S_\delta)$. Therefore, we get for all $m \geq m_0$ that $\pi_m(\alpha^{(\delta,\varepsilon)})$ is connectable with respect to $\kappa^{(m,-)}$.

(2) It is straight forward to show that for all $i \in J$ and $\delta > 0$ we have $d_{\mathrm{BL}}(y_i, y_i^{(\delta)}) \leq y_i(\mathcal{S} \setminus S_\delta) \leq \mu(\mathcal{S} \setminus S_\delta)$, where d_{BL} is the metric defined in (5.29) that induces the weak topology on $\mathcal{M}(\mathcal{S})$. For

any continuous compactly supported test function $f\colon \mathcal{M}_{\le 1}(\mathcal{S})\setminus\{0\}\to\mathbb{R}$ there exists $\varepsilon_f>0$ such that $f=0$ on $\{y\in\mathcal{M}_{\le 1}(\mathcal{S})\setminus\{0\}\colon |y|\le\varepsilon_f\}$, so for $\varepsilon\le\varepsilon_f$ (including the case $\varepsilon=0$) we have

$$\Big|\int \alpha^{(\delta,\varepsilon)}(\mathrm{d}y)f(y)-\int\alpha(\mathrm{d}y)f(y)\Big| = \sum_{i\in J_{\varepsilon_f}}|f(y_i^{(\delta)})-f(y_i)|\le |J_{\varepsilon_f}|\max_{i\in J_{\varepsilon_f}}|f(y_i^{(\delta)})-f(y_i)|. \tag{5.42}$$

Observe that the right-hand side converges to 0 as $\delta\to 0$. This implies (2).

(3) By lower-semicontinuity of the rate function it is clear that $I(\lambda,\alpha)\le\lim_{\delta,\varepsilon\to 0}I(\lambda,\alpha^{(\delta,\varepsilon)})$. So we only have to deal with the other estimate, i.e., we need to find an upper bound for

$$\begin{aligned} I(\lambda,\alpha^{(\delta,\varepsilon)})-I(\lambda,\alpha) &= \big(I_{\mathrm{Me}}(\mu-c_\lambda-c_\alpha)-I_{\mathrm{Me}}(\mu-c_\lambda-c_{\alpha^{(\delta,\varepsilon)}})\big)\\ &\quad+\big(I_{\mathrm{Ma}}(\alpha^{(\delta,\varepsilon)})-I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})\big)+\big(I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})-I_{\mathrm{Ma}}(\alpha)\big). \end{aligned} \tag{5.43}$$

Let $\gamma>0$. It is straight forward to verify that $c_{\alpha^{(\delta,\varepsilon)}}\to c_\alpha$ weakly as $\delta,\varepsilon\to 0$. Since I_{Me} is continuous we can choose δ_0 and ε_0 such that $|I_{\mathrm{Me}}(\mu-c_\lambda-c_\alpha)-I_{\mathrm{Me}}(\mu-c_\lambda-c_{\alpha^{(\delta,\varepsilon)}})|\le\gamma/3$ for all $\delta\le\delta_0$, $\varepsilon\le\varepsilon_0$.

Recall that

$$I_{\mathrm{Ma}}(\alpha)=\int\alpha(\mathrm{d}y)f_{\mathrm{Ma}}(y),\quad\text{where}\quad f_{\mathrm{Ma}}(y)=\langle y,\log\tfrac{\mathrm{d}y}{(1-\mathrm{e}^{-\kappa y})\mathrm{d}\mu}\rangle-\tfrac12\langle y,\kappa y\rangle, \tag{5.44}$$

and where we interpret the $\log$ term as equal to $+\infty$ if the density does not exist.

We have that

$$I_{\mathrm{Ma}}(\alpha)=I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})+I_{\mathrm{Ma}}(\textstyle\sum_{i\notin J_\varepsilon}\delta_{y_i})\ge I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})+I_{\mathrm{Ma}}(\sum_{i\notin J_\varepsilon}y_i),$$

where the inequality follows from (7.1), which is proved in Lemma 69. Note that $|\sum_{i\notin J_\varepsilon}y_i|\to 0$ as $\varepsilon\to 0$ and $f_{\mathrm{Ma}}(y)\to 0$ as $|y|\to 0$. So we can choose $\varepsilon\le\varepsilon_0$ such that $I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})-I_{\mathrm{Ma}}(\alpha)\le\gamma/3$.

As a last step we want to choose $\delta=\delta(\varepsilon)\le\delta_0$ such that $|I_{\mathrm{Ma}}(\alpha^{(\delta,\varepsilon)})-I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)})|\le\gamma/3$, so it remains to show that

$$\lim_{\delta\to 0}I_{\mathrm{Ma}}(\alpha^{(\delta,\varepsilon)})=I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)}). \tag{5.45}$$

Notice that the definition of $y^{(\delta)}$ given in (5.41) implies that $|y|-\mu(\mathcal{S}\setminus S_\delta)\le|y^{(\delta)}|\le|y|$. Therefore, we can choose δ small enough such that $\mu(\mathcal{S}\setminus S_\delta)\le\frac12\min_{i\in J_\varepsilon}(|y_i|-\varepsilon)$ to ensure that for all $i\in J_\varepsilon$ we have that $|y_i^{(\delta)}|>\varepsilon$. Now, we choose a function $\chi_\varepsilon\colon\mathcal{M}_{\le 1}(\mathcal{S})\setminus\{0\}\to\mathbb{R}$ that is equal to one on $\{y\in\mathcal{M}_{\le 1}(\mathcal{S})\setminus\{0\}\colon |y|>\varepsilon\}$, equal to zero on $\{y\in\mathcal{M}_{\le 1}(\mathcal{S})\setminus\{0\}\colon |y|\le\varepsilon/2\}$ and continuous. In particular, the function $\chi_\varepsilon f_{\mathrm{Ma}}$ is then compactly supported and we have that for all $\delta'\le\delta$

$$I_{\mathrm{Ma}}(\alpha^{(\delta',\varepsilon)})=\int\alpha^{(\delta',0)}(\mathrm{d}y)\chi_\varepsilon(y)f_{\mathrm{Ma}}(y).$$

Technically, we still have the problem that f_{Ma} can take values in $\mathbb{R}\cup\{+\infty\}$. But our assumption $I_{\mathrm{Ma}}(\alpha)<\infty$ implies that $f_{\mathrm{Ma}}(y_i)<\infty$ for all $i\in J_\varepsilon$, so by continuity of f_{Ma} and the finiteness of J_ε we can tune $\delta'\le\delta$ such that uniformly for all $i\in J_\varepsilon$ we have $f_{\mathrm{Ma}}(y_i^{(\delta)})\le C$ for some constant C. We already showed in (2) that $\alpha^{(\delta',0)}\to\alpha$, so having established continuity and compactly supportedness of the function in the integral we get that

$$\begin{aligned}\lim_{\delta'\to 0}I_{\mathrm{Ma}}(\alpha^{(\delta',\varepsilon)}) &= \lim_{\delta'\to 0}\int\alpha^{(\delta',0)}(\mathrm{d}y)\chi_\varepsilon(y)(f_{\mathrm{Ma}}(y)\wedge C)\\ &=\int\alpha(\mathrm{d}y)\chi_\varepsilon(y)(f_{\mathrm{Ma}}(y)\wedge C)=I_{\mathrm{Ma}}(\alpha^{(0,\varepsilon)}).\end{aligned}$$

Altogether the right-hand side of (5.43) can be bounded by γ, which proves the claim. □

Lemma 54 (LDP – lower bound)**.** *Suppose that all the assumptions of Theorem 1 are satisfied. Then the distribution of* $(\mathrm{Mi}_N, \mathrm{Ma}_N)$ *under* $\mathbb{P}_N$ *satisfies the lower bound part of the LDP with rate function* I*.*

Proof. Fix a set $G \in \mathcal{L} \times \mathcal{A}$ that is open with respect to the vague topology and a point $(\lambda, \alpha) \in G$. We will show that for any $\gamma > 0$ we have that

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big((\mathrm{Mi}_N, \mathrm{Ma}_N) \in G\big) \geq -I(\lambda, \alpha) - \gamma. \tag{5.46}$$

Note that in the case where $I(\lambda, \alpha) = \infty$, the claimed estimate always holds, so we assume that $I(\lambda, \alpha) < \infty$. Let $\gamma > 0$. We will use the approximating sequence $(\alpha^{(\delta,\varepsilon)})_{\delta>0,\varepsilon>0}$ constructed in Lemma 53. Let $\varepsilon > 0$ and $\delta > 0$ be small enough such that $\alpha^{(\delta,\varepsilon)} \in G$ and $|I(\lambda, \alpha) - I(\lambda, \alpha^{(\delta,\varepsilon)})| \leq \gamma/2$. Again, we write $\Pi_m(\tilde{\lambda}, \tilde{\alpha}) = (\pi_m(\tilde{\lambda}), \pi_m(\tilde{\alpha}))$ and also $\Pi_{m,n}(\tilde{\lambda}, \tilde{\alpha}) = (\pi_{m,n}(\tilde{\lambda}), \pi_{m,n}(\tilde{\alpha}))$ for any $n \geq m$. y Proposition 42 the set G is also open with respect to the projective limit topology. It is a general fact that the set

$$\mathcal{B}_1 := \{\Pi_m^{-1}(U_m) \colon m \in \mathbb{N},\, U_m \subset \mathcal{L}_m \times \mathcal{A}_m \text{ open}\} \tag{5.47}$$

is a basis of the projective limit topology. From now on, we fix some large $m_0 \in \mathbb{N}$ which we will specify later. We claim that also the set

$$\mathcal{B}_{m_0} := \{\Pi_m^{-1}(U_m) \colon m \geq m_0,\, U_m \subset \mathcal{L}_m \times \mathcal{A}_m \text{ open}\} \tag{5.48}$$

is a basis of the projective limit topology and argue this as follows: note that for any $m < m_0$ we can take any $n \geq m_0$ and use that $\pi_m = \pi_{m,n} \circ \pi_n$ holds by Lemma 43 to derive that $\Pi_m^{-1}(U_m) = \Pi_n^{-1}(\Pi_{m,n}^{-1}(U_m))$ for any open set U_m. Since the set $\Pi_{m,n}^{-1}(U_m)$ is again open due to the continuity of $\pi_{m,n}$, we get that $\Pi_m^{-1}(U_m) \in \mathcal{B}_{m_0}$. Altogether we have that $\mathcal{B}_1 \subset \mathcal{B}_{m_0}$.

Having established that $\mathcal{B}_{m_0}$ is a basis for the projective limit topology, we may pick $m \geq m_0$ and an open set $U_m \subset \mathcal{L}_m \times \mathcal{A}_m$ such that $(\lambda, \alpha^{(\delta,\varepsilon)}) \in \Pi_m^{-1}(U_m) \subset G$. Therefore, we see that

$$\begin{aligned}\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big((\mathrm{Mi}_N, \mathrm{Ma}_N) \in G\big) &\geq \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big((\mathrm{Mi}_N^{(m)}, \mathrm{Ma}_N^{(m)}) \in U_m\big) \\ &\geq -\inf_{U_m} I^{(m,-)} \geq -I^{(m,-)}(\Pi_m(\lambda, \alpha^{(\delta,\varepsilon)})),\end{aligned} \tag{5.49}$$

where we used Corollary 49 for the second inequality. Using Lemma 51 and Lemma 53 we can pick m_0 large enough such that for all $m \geq m_0$ the measures $\pi_m(\alpha^{(\delta,\varepsilon)})$ are connectable and $|I^{(m,-)}(\Pi_m(\lambda, \alpha^{(\delta,\varepsilon)})) - I(\lambda, \alpha^{(\delta,\varepsilon)})| \leq \gamma/2$. Altogether, this gives (5.46). □

6 The minimizers of I_{Mi}

In this section, we derive an explicit description of the minimizer(s) λ of I_{Mi} under the constraint $c_\lambda = c$ for any $c \in \mathcal{M}(\mathcal{S})$ satisfying $c \leq \mu$. This will allow us to solve the optimization problem in (2.6), i.e., to identify the minimizer(s) of the rate function for the LDP for Mi_N in Theorem 9. It will also be used as an important intermediate step in deriving the full optimization of the rate function I of the LDP in Theorem 7, our main result. Recall the notation that we introduced in Section 2.1, in particular the definition of $\Sigma(\kappa, c)$ from (2.3). Here is the main result of this section.

Proposition 55 (Minimizers of I_{Mi}). *Fix a probability measure μ on $\mathcal{S}$ and a kernel κ on $\mathcal{S}\times\mathcal{S}$ that is nonnegative and continuous.*

Let $c\in\mathcal{M}(\mathcal{S})$ be a measure such that $c\le\mu$.

(i) Assume that $\Sigma(\kappa,c)\le 1$. Then

$$\inf_{\lambda\in\mathcal{L}\colon c_\lambda=c} I_{\mathrm{Mi}}(\lambda)=\Big\langle c,\log\frac{\mathrm{d}c}{\mathrm{d}\mu}\Big\rangle+\frac{1}{2}\langle c,\kappa(\mu-c)\rangle, \tag{6.1}$$

and the infimum is attained in the unique minimizer λ_c defined in (2.1)*.*

(ii) Assume that $\Sigma(\kappa,c)>1$. Then

$$\inf_{\lambda\in\mathcal{L}\colon c_\lambda=c} I_{\mathrm{Mi}}(\lambda)\ge\inf_{\lambda\in\mathcal{L}\colon c_\lambda=b^*} I_{\mathrm{Mi}}(\lambda)+I_{\mathrm{Me}}(c-b^*) \tag{6.2}$$

where $b^*=b^*(c)\in\mathcal{M}(\mathcal{S})$ is the minimal, non-trivial (i.e., not equal to c) solution to (2.9) *and satisfies $\Sigma(\kappa,b^*)=1$.*

It is interesting to notice that one can see the phase transition already from the sole consideration of I_{Mi}. We will refer to (i) and (ii) as to the sub- and supercritical cases, respectively.

In the case where $\mathcal{S}$ is finite we can actually prove an equality in (6.2). In the general case we also expect this to be true, but did not attempt a proof, since the inequality will be enough to prove our main results, Theorems 9 and 7.

The proof is naturally divided into Sections 6.1 – 6.4 according to the distinctions between finite $\mathcal{S}$ (the discrete case) or general compact $\mathcal{S}$ and between the sub- and supercritical cases. In Section 6.1 we construct minimizers for subcritical measures c for finite $\mathcal{S}$; we analyze if the only candidate λ (coming from the Euler–Lagrange equations) satisfies the constraint $c_\lambda=c$, which requires the result about the multivariate power series from Section 3.1. In Section 6.2 we generalize the results to a general compact type space via an approximation argument. In order to deal with supercritical measures c for $\mathcal{S}$ a finite set, we also rely on combinatorial results in Section 6.3. Afterwards, we handle the general supercritical case in Section 6.4.

6.1 The discrete, subcritical case

In this section, we formulate and prove the main assertions about the minimizers of I_{Mi} in the discrete case, i.e., the case of a finite type space $\mathcal{S}$. Here we will be using the notation of linear algebra, i.e., measures $\lambda\in\mathcal{L}$ on $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ will be written as sequences $(\lambda_k)_{k\in\mathbb{N}_0^{\mathcal{S}}}$.

Recall the definition of the rate function I_{Mi} from Theorem 29 as well as the notation for the integrated type-configuration $c_r(\lambda)=\sum_{k\in\mathbb{N}_0^{\mathcal{S}}}\lambda_k\,k_r$ for $r\in\mathcal{S}$ introduced in (4.5). We write $[0,\mu]$ for the set of all $c\in[0,\infty)^{\mathcal{S}}$ satisfying $0\le c_r\le\mu_r$ for any $r\in\mathcal{S}$.

For $c\in[0,\infty)^{\mathcal{S}}$, define $\lambda(c)=(\lambda_k(c))_{k\in\mathbb{N}_0^{\mathcal{S}}}$ by

$$\lambda_k(c)=\tau(k)\prod_{s\in\mathcal{S}}\frac{(c_s\mathrm{e}^{-(\kappa c)_s})^{k_s}}{k_s!},\qquad k\in\mathbb{N}_0^{\mathcal{S}} \tag{6.3}$$

and note that this definition is the discrete analog of the general form of the minimizer in (2.1).

The aim of the present Section 6.1 is to verify the subcritical case of Proposition 55 in the discrete setting, which we restate here quickly.

Proposition 56. *Let $c = (c_s)_{s\in\mathcal{S}}$ be in $[0,\mu]$. Assume that $\Sigma(\kappa, c) \le 1$. Then*

$$\inf_{\lambda\in\mathcal{L}\colon\, c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) = \Big\langle c, \log\frac{c}{\mu}\Big\rangle + \frac{1}{2}\,\langle c, \kappa(\mu - c)\rangle\,, \tag{6.4}$$

and the infimum is attained in the unique minimizer $\lambda(c)$ defined in (6.3)*.*

To derive the form of the minimizer given in (6.3), we will start by giving a short heuristic. First note that I_{Mi} is a strictly convex function and that $\{\lambda\colon c(\lambda) = c\}$ is a convex set, which implies that there is at most one minimizer. Assume that a minimizer λ^* exists in the interior of $\{\lambda\colon c(\lambda) = c\}$. Then by formally writing down the Euler–Lagrange equations, one can see that

$$\lambda^*_k = \tau(k)\prod_{s\in\mathcal{S}}\frac{\theta_s^{k_s}}{k_s!}, \qquad k\in\mathbb{N}_0^{\mathcal{S}}, \tag{6.5}$$

where $\theta = (\theta_s)_{s\in\mathcal{S}}$ is some non-negative real-valued vector. Note that θ has to be chosen in such a way that $c(\lambda^*) = c$, i.e., for every $r\in\mathcal{S}$ the multivariate power series

$$c_r(\lambda^*) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}}\tau(k)k_r\prod_s\frac{\theta_s^{k_s}}{k_s!} \tag{6.6}$$

converges with limit c_r. We already encountered in Section 3.1 that for $\theta = c\mathrm{e}^{-\kappa c}$ the power series on the right-hand side of (6.6) has the right value. The following is just a reformulation of the results from Lemma 16 and Proposition 17 using the notation of the present section.

Corollary 57. *Let $c = (c_s)_{s\in\mathcal{S}}$ in $[0,\mu]$ and assume that $\Sigma(\kappa, c) \le 1$. Then for $\lambda^* = \lambda(c)$ we have that $c(\lambda^*) = c$.*

A rigorous argument showing that this choice uniquely minimizes I_{Mi} can be found at the end of this section. The identification of the optimal rate $I_{\mathrm{Mi}}(\lambda(c))$ needs an additional property of the minimizer, namely a formula of its total mass, which we derive in Lemma 59. For this we use the following recursive formula.

Lemma 58. *Let $k\in\mathbb{N}_0^{\mathcal{S}}$. Then we have the recursion*

$$\sum_{r,s\in\mathcal{S}}\kappa(r,s)\sum_{m,\widetilde{m}\in\mathbb{N}_0^{\mathcal{S}}\colon\, m+\widetilde{m}=k}\Big(\prod_{u\in\mathcal{S}}\frac{k_u!}{m_u!\widetilde{m}_u!}\Big)\tau(m)m_r\tau(\widetilde{m})\widetilde{m}_s = 2(|k|-1)\tau(k). \tag{6.7}$$

Proof. Let $(x_i)_{i\in[|k|]}\in\mathcal{S}^{|k|}$ be a vector compatible to k, i.e., $k = \sum_i\delta_{x_i}$ and recall the definition of $\tau(k)$ from (3.2). For $i,j\in[|k|]$ with $i\neq j$ define

$$W_{i,j} := \sum_{T\in\mathcal{T}(k)\colon\, \{i,j\}\in E(T)}\ \prod_{\{v,w\}\in E(T)}\kappa(x_v,x_w), \tag{6.8}$$

i.e., $W_{i,j}$ is the total weight of trees containing the edge $\{i,j\}$. Observe, that each tree T on $[|k|]$ contains exactly $|k|-1$ edges and, for each edge $\{i,j\}\in E(T)$ the weight of T appears once in $W_{i,j}$ and once in $W_{j,i}$. Thus, the weight of T is counted $2(|k|-1)$ times in the sum $\sum_{i\neq j}W_{i,j}$, which implies that

$$2(|k|-1)\tau(k) = \sum_{i\neq j}W_{i,j}. \tag{6.9}$$

Now, for a fixed pair of types $r, s \in \mathcal{S}$ consider the weights of trees containing an edge connecting some type r with some type s vertex, i.e., consider $\sum_{i\neq j\colon x_i=r, x_j=s} W_{i,j}$. Notice that each tree contributing to this weight can be decomposed into an edge of weight $\kappa(r,s)$ and two trees T_r and T_s with roots of type r and s respectively. (The term 'root' is here only used to mark a certain vertex, not to give some directed structure.) This implies the formula

$$\sum_{i\neq j\colon x_i=r, x_j=s} W_{i,j} = \kappa(r,s) \sum_{m+\widetilde{m}=k} \Big(\prod_{u\in\mathcal{S}} \frac{k_u!}{m_u!\widetilde{m}_u!}\Big)\tau(m)m_r\tau(\widetilde{m})\widetilde{m}_s. \tag{6.10}$$

Here we used formula (3.9) to collect the weight coming from the possible choices of T_r and T_s, which is $\tau(m)m_r$ and $\tau(\widetilde{m})\widetilde{m}_s$ respectively. Formula (6.7) now follows by summing over all possible pairs $r, s \in \mathcal{S}$. □

Lemma 59. *Let $c = (c_s)_{s\in\mathcal{S}}$ be non-negative with $\Sigma(\kappa, c) \leq 1$. Then for $\lambda(c)$ defined as in* (6.3) *we have that*

$$|\lambda(c)| = \sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k(c) = \sum_{r\in\mathcal{S}} c_r - \frac{1}{2}\langle c, \kappa c\rangle. \tag{6.11}$$

Proof. Writing $\lambda^* = \lambda(c)$ and using that $c(\lambda^*) = c$ we show the equivalent equation

$$\sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k^*(|k|-1) = \frac{1}{2}\langle c, \kappa c\rangle. \tag{6.12}$$

For fixed $k \in \mathbb{N}_0^{\mathcal{S}}$ the recursive equation (6.7) for $\tau(k)$ easily implies

$$\frac{1}{2} \sum_{m,\widetilde{m}\in\mathbb{N}_0^{\mathcal{S}}\colon m+\widetilde{m}=k} \sum_{r,s\in\mathcal{S}} \lambda_m^* m_r \kappa(r,s) \lambda_{\widetilde{m}}^* \widetilde{m}_s = \lambda_k^*(|k|-1).$$

With the assumption $\Sigma(\kappa, c) \leq 1$ all series in the next equations converge (absolutely) by Corollary 57, so by rearranging terms we get that

$$\sum_{k\in\mathbb{N}_0^{\mathcal{S}}} \lambda_k^*(|k|-1) = \frac{1}{2} \sum_{m\in\mathbb{N}_0^{\mathcal{S}}} \sum_{\widetilde{m}\in\mathbb{N}_0^{\mathcal{S}}} \sum_{r,s\in\mathcal{S}} \lambda_m^* m_r \kappa(r,s) \lambda_{\widetilde{m}}^* \widetilde{m}_s = \frac{1}{2}\langle c, \kappa c\rangle..$$

□

Combining the results from Corollary 57 and Lemma 59 we can now give the proof of Proposition 56:

Proof of Proposition 56. Assume that $\Sigma(\kappa, c) \leq 1$. Define $\lambda^* = \lambda(c)$ as in (6.3). By Corollary 57 we have that $c_r(\lambda^*) = \sum_k \lambda_k^* k_s = c_s$ for all $s \in \mathcal{S}$. Now, take any λ satisfying $c_s(\lambda) = \sum_k \lambda_k k_s = c_s$ for all $s \in \mathcal{S}$. Then from (4.6), using the formula from Lemma 59, we get

$$\begin{aligned} I_{\mathrm{Mi}}(\lambda) &= \sum_k \lambda_k \log\frac{\lambda_k}{\lambda_k^*} + \sum_k \lambda_k \log\Big(\prod\nolimits_{s\in\mathcal{S}} \Big(\frac{c_s \mathrm{e}^{-(\kappa c)_s}}{\mu_s}\Big)^{k_s}\Big) + \sum_k \lambda_k(|k|-1) + \frac{1}{2}\langle c, \kappa\mu\rangle \\ &= \mathbb{H}(\lambda|\lambda^*) - |\lambda^*| + \Big\langle c, \log\frac{c}{\mu}\Big\rangle + |c| + \frac{1}{2}\langle c, \kappa\mu\rangle - \langle c, \kappa c\rangle \\ &\geq \Big\langle c, \log\frac{c}{\mu}\Big\rangle + \frac{1}{2}\langle c, \kappa(\mu - c)\rangle, \end{aligned} \tag{6.13}$$

where we wrote $\mathbb{H}(\lambda|\lambda^*) = \langle \lambda, \log\frac{\lambda}{\lambda^*}\rangle + |\lambda^*| - |\lambda|$ for the entropy and used that $\mathbb{H}(\lambda|\lambda^*) \geq 0$ with equality if and only if $\lambda = \lambda^*$. □

6.2 The general subcritical case

In this section we derive Proposition 55(i). The proof is similar to the proof of the discrete variant in Section 6.1. Again, there is an explicit candidate for the minimizer, but one has to prove that it is admissable, and we need to identify its total mass. This is done in the analogs of Corollary 57 and Lemma 59, see Lemmas 61 and 62, whose proofs proceed via a discrete approximation based on the material of Section 5.1.

In the current case of a general compact metric type space $\mathcal{S}$, the candidate for a minimizer is given in terms of a Poisson point process, see (6.14). Recall that we write $\mathbb{Q}_\theta$ for the distribution of a Poisson point process $\mathbb{X} = (X_i)_{i\in I}$ in $\mathcal{S}$ with intensity measure $\theta \in \mathcal{M}(\mathcal{S})$. We write $k = \sum_i \delta_{X_i} \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ for the measure induced by the random point cloud. Note that the points X_i do not have to be distinct with positive probability, if θ has no Lebesgue density. We start by noting a simple fact about the densities between absolute continuous Poisson point processes.

Lemma 60. *Let $\theta, \hat\theta \in \mathcal{M}(\mathcal{S})$ with $\hat\theta \ll \theta$. Then $\mathbb{Q}_{\hat\theta} \ll \mathbb{Q}_\theta$ and*

$$\frac{\mathrm{d}\mathbb{Q}_{\hat\theta}}{\mathrm{d}\mathbb{Q}_\theta}(k) = \mathrm{e}^{\langle k, \log \frac{\mathrm{d}\hat\theta}{\mathrm{d}\theta}\rangle + \theta(\mathcal{S}) - \hat\theta(\mathcal{S})}, \quad k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S}).$$

Recall the definition of $\tau(k)$ introduced in (1.9). Also recall that for a fixed $c \in \mathcal{M}$ according to definition (2.1) the candidate for the minimizer of I_{Mi} under the constraint $c_\lambda = c$ has the form

$$\lambda_c(\mathrm{d}k) = \mathrm{e}^{\theta_c(\mathcal{S})}\tau(k)\mathbb{Q}_{\theta_c}(\mathrm{d}k), \quad \text{where } \theta_c(\mathrm{d}r) = \mathrm{e}^{-\kappa c(r)}c(\mathrm{d}r). \tag{6.14}$$

We first provide a generalized version of Corollary 57 and Lemma 59. We first impose the stricter condition $\Sigma(\kappa, c) < 1$.

Lemma 61. *Let $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu$. Assume that $\Sigma(\kappa, c) < 1$. Then the following holds.*

1 *For any continuous test function $f\colon \mathcal{S} \to [0,\infty)$ we have that*

$$\int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} \lambda_c(\mathrm{d}k)\langle k, f\rangle = \langle c, f\rangle. \tag{6.15}$$

2 *The total mass of λ_c is given by*

$$|\lambda_c| = \int_{\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})} \lambda_c(\mathrm{d}k) = c(\mathcal{S}) - \frac{1}{2}\langle c, \kappa c\rangle. \tag{6.16}$$

Proof. We focus on showing equation (6.15); the proof of (6.16) is similar (see the end of the proof). Abbreviating $\theta := \theta_c$ and inserting the definition of λ_c we have to prove that

$$\mathbb{Q}_\theta\Big[\tau(k)\langle f, k\rangle \mathrm{e}^{\theta(\mathcal{S})}\Big] = \langle f, c\rangle, \tag{6.17}$$

where we conceive k as an $\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$-valued random variable on the left-hand side. The idea is to deduce the equality from the one that we have in the finite-type case by using the discretization scheme from Section 5.1. Recall the notation from Section 5.1, where we discretized the compact metric space $\mathcal{S}$ into finite spaces $\mathcal{S}_m$, $m \in \mathbb{N}$, and defined the projections π_m, $m \in \mathbb{N}$, on different spaces in equations (5.6)–(5.8). For $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ we will again identify the discretized measure $\pi_m(k)$ with an element of $\mathbb{N}_0^{\mathcal{S}_m} \setminus \{0\}$. Via $\mathcal{S}_m \subset \mathcal{S}$ the function f can be restricted to $\mathcal{S}_m$ and

write $f_m = f|_{\mathcal{S}_m}$. Also, we write $c_m := \pi_m(c)$ and identify it with a vector $(c_m(r))_{r\in\mathcal{S}_m}$. Recall the definitions of the discretized kernels. Let $\kappa_m \in \{\kappa^{(m,+)}, \kappa^{(m,-)}\}$, where $\kappa^{(m,\star)}$ for $\star = \pm$ is defined as in (5.13) and (5.14). Denote $\theta_m(r) := \mathrm{e}^{-(\kappa_m c_m)(r)} c_m(r)$, $r \in \mathcal{S}_m$. For $k \in \mathbb{N}_0^{\mathcal{S}_m}$ let $\tau_m(k)$ be defined as in (3.2), but with respect to κ_m. Fix a continuous function $f\colon \mathcal{S} \to \mathbb{R}$. Our aim is to show that

$$\mathbb{Q}_\theta\Big[\tau(k)\langle f,k\rangle \mathrm{e}^{\theta(\mathcal{S})}\Big] = \lim_{m\to\infty} \mathbb{Q}_{\theta_m}\Big[\tau_m(k)\langle f_m,k\rangle \mathrm{e}^{\theta_m(\mathcal{S}_m)}\Big] = \langle f,c\rangle, \tag{6.18}$$

which finishes the proof of (6.15).

We start with proving the second equality of (6.18). It is straight forward to show that

$$\big|\Sigma(\kappa_m,c_m) - \Sigma(\kappa,c)\big| \le \|\kappa_m\circ\pi_m - \kappa\|_\infty. \tag{6.19}$$

and hence $\Sigma(\kappa_m,c_m) \to \Sigma(\kappa,c)$, as $m \to \infty$ and so we will have for large $m \in \mathbb{N}$ that $\Sigma(\kappa_m,c_m) < 1$. Then we get

$$\begin{aligned}\mathbb{Q}_{\theta_m}\Big[\tau_m(k)\langle f_m,k\rangle \mathrm{e}^{\theta_m(\mathcal{S}_m)}\Big] &= \sum_{r\in\mathcal{S}_m} f_m(r) \sum_{k\in\mathbb{N}_0^{\mathcal{S}_m}} \tau_m(k) k_r \prod_{s\in\mathcal{S}_m} \frac{(\theta_m(s))^{k_s}}{k_s!}\\ &= \sum_{r\in\mathcal{S}_m} f_m(r)c_m(r) = \int_{\mathcal{S}} f(\pi_m(x))\,c(\mathrm{d}x)\\ &\to \int_{\mathcal{S}} f(x)\,c(\mathrm{d}x) = \langle f,c\rangle, \qquad m\to\infty.\end{aligned}$$

where the second equation only holds if m is large enough, and thus $\Sigma(\kappa_m,c_m) \le 1$ due to Lemma 16(i) and Proposition 17.

Now we show the first equation of (6.18). Note that $\mathbb{Q}_{\theta_m}$ is a point process on $\mathcal{S}_m$, whereas $\mathbb{Q}_\theta$ is a point process on $\mathcal{S}$. However, by defining an intensity measure on $\mathcal{S}$ by $\bar\theta_m(\mathrm{d}x) := \mathrm{e}^{-(\kappa_m c_m)(\pi_m(x))} c(\mathrm{d}x)$ we have that $\theta_m = \bar\theta_m \circ \pi_m^{-1}$ and hence $\mathbb{Q}_{\theta_m} = \mathbb{Q}_{\bar\theta_m}\circ\pi_m^{-1}$ holds by the mapping theorem for Poisson point processes. Therefore, according to Lemma 60,

$$\mathbb{Q}_{\theta_m}\Big[\tau_m(k)\langle f_m,k\rangle \mathrm{e}^{\theta_m(\mathcal{S}_m)}\Big] = \mathbb{Q}_{\bar\theta_m}\Big[\tau_m(\pi_m(k))\langle f,\pi_m(k)\rangle \mathrm{e}^{\theta_m(\mathcal{S}_m)}\Big] = \mathbb{Q}_\theta\Big[\Psi_m^f\Big], \tag{6.20}$$

with

$$\Psi_m^f(k) := \tau_m(\pi_m(k))\langle f,\pi_m(k)\rangle \mathrm{e}^{\langle k,\kappa c-(\kappa_m c_m)\circ\pi_m\rangle}\mathrm{e}^{\theta(\mathcal{S})}, \quad k\in\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}). \tag{6.21}$$

Note that Ψ_m^f converges pointwise to Ψ^f, where $\Psi^f(k) = \tau(k)\langle f,k\rangle \mathrm{e}^{\theta(\mathcal{S})}$, $k \in \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$. Hence, the first equation of (6.18) immediately follows as soon as we have given an argument for interchanging the limit as $m\to\infty$ and the integration with respect to $\mathbb{Q}_\theta$. We will be using Lebesgue's theorem about dominated convergence for that. Let us introduce a majorant. Recalling the definition (5.33) of $\tau_m^{(+)}$ we define $\widehat\Psi_m$ by

$$\widehat\Psi_m(k) := \tau_m^{(+)}(\pi_m(k))|\pi_m(k)|\mathrm{e}^{\langle k,\kappa c-(\kappa^{(m,-)}c_m)\circ\pi_m\rangle}\mathrm{e}^{\theta(\mathcal{S})}, \qquad k\in\mathcal{M}_{\mathbb{N}_0}(\mathcal{S}). \tag{6.22}$$

Then, since $\kappa^{(m,-)} \le \kappa_m \le \kappa^{(m,+)}$ we clearly have for any $k\in\mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ that $\Psi_m^f(k) \le \|f\|_\infty \widehat\Psi_m(k)$ and $\widehat\Psi_m(k) \le \widehat\Psi_{m_0}(k)$ if $m \ge m_0$. Hence, $\widehat\Psi_{m_0}$ is a majorant. It remains to show that there exists m_0 such that $\mathbb{Q}_\theta(\widehat\Psi_{m_0}) < \infty$, then the majorant $\widehat\Psi_{m_0}$ is integrable. Arguing as in (6.20) we have that

$$\mathbb{Q}_\theta(\widehat\Psi_m) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}_m}} \tau_m^{(+)}(k)|k| \prod_{s\in\mathcal{S}_m} \frac{\theta_m^{(-)}(s)^{k_s}}{k_s!}, \tag{6.23}$$

where $\theta_m^{(-)}(s) = \mathrm{e}^{-(\kappa^{(m,-)}c_m)(s)}c_m(s)$, for $s \in \mathcal{S}_m$. Let $\chi_m := \chi(\kappa^{(m,+)}, \theta_m^{(-)})$ be defined as in (3.15). We can argue as in the proof of Lemma 20 to get that, for any $n \in \mathbb{N}$,

$$\sum_{k\in\mathbb{N}_0^{\mathcal{S}_m}\,:\,|k|=n} \tau_m^{(+)}(k)|k| \prod_{s\in\mathcal{S}_m} \frac{(\theta_m^{(-)}(s))^{k_s}}{k_s!} \le \mathrm{e}^{o(n)}\mathrm{e}^{-n[\chi_m-1]}, \tag{6.24}$$

(where the $\mathrm{e}^{o(n)}$-term is actually given by $|\mathcal{M}_1^{(n)}(\mathcal{S}_m)| \sum_{r\in\mathcal{S}_m} \Delta_r(n\nu)$). Abbreviating $\delta_m = \|\kappa^{(m,+)} - \kappa^{(m,-)}\|_\infty$, we further have that

$$\begin{aligned}\chi_m &= \inf_{\nu\in\mathcal{M}_1(\mathcal{S}_m)} \left\{ \sum_{r\in\mathcal{S}_m} \nu_r \log \frac{\nu_r}{(\kappa^{(m,+)}\nu)_r(\theta_m^{(+)})_r} - \langle \kappa^{(m,+)}c_m - \kappa^{(m,-)}c_m, \nu\rangle \right\} \\ &\ge \chi(\kappa^{(m,+)}, \theta_m^{(+)}) - \delta_m = \Sigma_m - \log\Sigma_m - \delta_m,\end{aligned}$$

where $\Sigma_m := \Sigma(\kappa^{(m,+)}, c_m)$. Now, choose $\varepsilon > 0$ small enough such that $\Sigma(\kappa, c) + \varepsilon < 1$ and $\Sigma(\kappa, c) - \log(\Sigma(\kappa, c)) > 1 + \varepsilon$. It holds that $\Sigma_m \to \Sigma(\kappa, c)$, as $m \to \infty$, and clearly we have that $\delta_m \to 0$, as $m \to \infty$. Additionally, we use that the function $x \mapsto \phi(x) := x - \log x$ is continuous and decreasing on $[0, 1]$. Hence, we find m_0 such that $\delta_{m_0} \le \varepsilon/2$, as well as $\Sigma_{m_0} \le \Sigma(\kappa, c)+\varepsilon < 1$ and $\phi(\Sigma_{m_0}) \ge \phi(\Sigma(\kappa, c)) - \varepsilon/2$. Consequently,

$$\chi_{m_0} \ge \Sigma_{m_0} - \log\Sigma_{m_0} - \frac{\varepsilon}{2} \ge \Sigma(\kappa, c) - \log\Sigma(\kappa, c) - \varepsilon > 1,$$

which altogether implies that

$$\mathbb{Q}_\theta(\widehat{\Psi}_{m_0}) = \sum_{k\in\mathbb{N}_0^{\mathcal{S}_m}} \tau_m^{(+)}(k)|k| \prod_{s\in\mathcal{S}_m} \frac{(\theta_m^{(-)}(s))^{k_s}}{k_s!} \le \sum_{n\in\mathbb{N}} \mathrm{e}^{o(n)}\mathrm{e}^{-n[\chi_{m_0}-1]} < \infty. \tag{6.25}$$

Thus, Lebesgue's theorem of dominated convergence is applicable and (6.18) follows.

Equation (6.16) can be shown in the same way and relies on the discrete version of the equation, derived in Lemma 59. □

Lemma 62. *The statement of Lemma 61 is also true under the assumption* $\Sigma(\kappa, c) = 1$.

Proof. The idea is to construct a sequence $c^{(n)} \in \mathcal{M}(\mathcal{S})$ with $\Sigma(\kappa, c^{(n)}) < 1$ such that $\theta^{(n)} := \theta_{c^{(n)}} \nearrow \theta_c =: \theta$ monotonically as $n \to \infty$.

Recall that $\mathcal{S}$ is compact and κ is continuous, hence a standard argument (see e.g. [BJR07, Lemma 5.15]) shows that the operator $T_{\kappa,c}$ is a positive Hilbert–Schmidt operator and therefore has a non-negative eigenfunction corresponding to the eigenvalue $\Sigma(\kappa, c)$. By the assumption $\Sigma(\kappa, c) = 1$ we can find a function $g\colon \mathcal{S} \to (0, \infty)$ such that $T_{\kappa,c}g = g$. For any $n \in \mathbb{N}$ define $c^{(n)} \in \mathcal{M}(\mathcal{S})$ via $\frac{\mathrm{d}c^{(n)}}{\mathrm{d}c} := 1 - \frac{1}{n}g$. Then for n large enough $c^{(n)}(A) < c(A)$ for any measurable $A \subset \mathcal{S}$ with $\int_A g\,\mathrm{d}c > 0$. In particular $\Sigma(\kappa, c^{(n)}) < 1$. (An ad-hoc argument in the case that κ is irreducible is as follows: Pick an $L^2(c^{(n)})$-normalized positive eigenfunction g_n of $T_{\kappa,c^{(n)}}$ corresponding to the eigenvalue $\Sigma(\kappa, c^{(n)})$ and observe that $\widetilde{g}_n(x) = g_n(x)(1 - \frac{1}{n}g(x))^{1/2}$ is $L^2(c)$ normalized and that $\Sigma(\kappa, c^{(n)}) = \|T_{\kappa,c^{(n)}}g_n\|_{L^2(c^{(n)})} < \|T_{\kappa,c}\widetilde{g}_n\|_{L^2(c)} \le \|T_{\kappa,c}\| = \Sigma(\kappa, c)$. If κ is reducible, then apply this argument to the irreducible components.) Now, observe that for any $n \in \mathbb{N}$ we have

$$\frac{\mathrm{d}\theta^{(n)}}{\mathrm{d}\theta} = \mathrm{e}^{\kappa(c-c^{(n)})}\frac{\mathrm{d}c^{(n)}}{\mathrm{d}c} = \mathrm{e}^{T_{\kappa,c}(\frac{1}{n}g)}\Big(1 - \frac{1}{n}g\Big) = \mathrm{e}^{\frac{1}{n}g}\Big(1 - \frac{1}{n}g\Big),$$

and the right-hand side converges pointwise monotonically to 1, as $n \to \infty$.

Now, fix any continuous test function $f\colon \mathcal{S} \to [0,\infty)$. Then, by monotone convergence and the fact that we can apply Lemma 61 to $c^{(n)}$ for all $n \in \mathbb{N}$, we get that

$$\begin{aligned}\mathbb{Q}_\theta\big[\tau(k)\langle k,f\rangle \mathrm{e}^{\theta(\mathcal{S})}\big] &= \lim_{n\to\infty} \mathbb{Q}_\theta\big[\tau(k)\langle k,f\rangle \mathrm{e}^{\theta(\mathcal{S})}\mathrm{e}^{\langle k,\log\frac{\mathrm{d}\theta^{(n)}}{\mathrm{d}\theta}\rangle}\big] = \lim_{n\to\infty} \mathbb{Q}_{\theta^{(n)}}\Big[\tau(k)\langle k,f\rangle \mathrm{e}^{\theta^{(n)}(\mathcal{S})}\Big]\\ &= \lim_{n\to\infty}\langle c^{(n)},f\rangle = \langle c,f\rangle.\end{aligned}$$

The same argument shows that $\mathbb{Q}_\theta\Big[\tau(k)e^{\theta(\mathcal{S})}\Big] = c(\mathcal{S}) - \frac{1}{2}\langle c,\kappa c\rangle$. □

Now we can identify the minimizers of I_{Mi}. The following is a variant of Proposition 56 in the general setting. Once having established Lemmas 61 and 62, the proof in the general setting follows the ones in the discrete setting. Recall that μ is the reference probability measure on $\mathcal{S}$.

Lemma 63 (Minimizers of I_{Mi})**.** *Assume that $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu$ satisfying $\Sigma(\kappa,c) \le 1$. Then the unique minimizer λ of I_{Mi} under the assumption $c_\lambda = c$ is equal to λ_c defined in* (2.1) *and*

$$\min_{\lambda\in\mathcal{L}\colon c_\lambda = c} I_{\mathrm{Mi}}(\lambda) = I_{\mathrm{Mi}}(\lambda_c) = -\frac{1}{2}\langle c,\kappa c\rangle + \langle c,\log\tfrac{\mathrm{d}c}{\mathrm{d}\mu}\rangle.$$

Proof. Note that λ_c is admissible, according to Lemmas 61 and 62, since $c_{\lambda^*} = c$.

Using Lemma 60 and the fact that μ is a probability measure we can rewrite the measure $\mathbb{Q}_\mu$ as

$$\mathbb{Q}_\mu(\mathrm{d}k) = \mathrm{e}^{\theta(c)(\mathcal{S})}\mathbb{Q}_{c\mathrm{e}^{-\kappa*c}}(\mathrm{d}k)\mathrm{e}^{\langle k,\kappa*c\rangle}\mathrm{e}^{-\langle k,\log\frac{\mathrm{d}c}{\mathrm{d}\mu}\rangle}\mathrm{e}^{-1}$$

Now, writing $\mathcal{M} = \mathcal{M}_{\mathbb{N}_0}(\mathcal{S})$ we get for any $\lambda \in \mathcal{L}$ satisfying $c_\lambda = c$ that

$$\begin{aligned}I_{\mathrm{Mi}}(\lambda) &= \Big\langle \lambda, \log\frac{\mathrm{d}\lambda}{\tau\mathrm{d}\mathbb{Q}_\mu}\Big\rangle + c(\mathcal{S}) - 2\lambda(\mathcal{M}) + \frac{1}{2}\langle c,\kappa\mu\rangle\\ &= \Big\langle \lambda, \log\frac{\mathrm{d}\lambda}{\mathrm{d}\lambda_c}\Big\rangle + \Big\langle c, \log\frac{\mathrm{d}c}{\mathrm{d}\mu}\Big\rangle - \langle c,\kappa c\rangle + c(\mathcal{S}) - \lambda(\mathcal{M}) + \frac{1}{2}\langle c,\kappa\mu\rangle\\ &= \mathbb{H}(\lambda|\lambda_c) - \lambda_c(\mathcal{M}) + \Big\langle c, \log\frac{\mathrm{d}c}{\mathrm{d}\mu}\Big\rangle - \langle c,\kappa c\rangle + c(\mathcal{S}) + \frac{1}{2}\langle c,\kappa\mu\rangle\\ &= \mathbb{H}(\lambda|\lambda_c) + \Big\langle c, \log\frac{\mathrm{d}c}{\mathrm{d}\mu}\Big\rangle + \frac{1}{2}\langle c,\kappa(\mu - c)\rangle.\end{aligned}$$

We used the fact that $\lambda_c(\mathcal{M}) = c(\mathcal{S}) - \frac{1}{2}\langle c,\kappa c\rangle$, which was derived in the last statement of Lemma 61. Since $\mathbb{H}(\lambda|\lambda_c) \ge 0$ and $\mathbb{H}(\lambda|\lambda_c) = 0$ if and only if $\lambda = \lambda_c$, the claim follows. □

6.3 The discrete, supercritical case

In this section, we assume again that $\mathcal{S}$ is a finite space and investigate the case where the measure c (the one that formulates the constraint) is supercritical, meaning $\Sigma(\kappa,c) > 1$. The aim of this section is to verify the following result.

Proposition 64 (Discrete, supercritical case)**.** *Let $c \in [0,\mu]$ with $\Sigma(\kappa,c) > 1$. Then*

$$\inf_{\lambda\colon c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) = \inf_{\lambda\colon c(\lambda)=b^*} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Me}}(c-b^*) \tag{6.26}$$

where $b^* = b^*(c)$ is the minimal non-trivial (i.e., not equal to c) solution to

$$\kappa(c-b^*)b^* = c - b^*, \qquad b^* \le c, \tag{6.27}$$

and satisfies $\Sigma(\kappa, b^*) = 1$.

Indeed, one possible realization of the rate (6.26) is given by constructing a minimizer as in formula (6.3) with respect to the (sub-)critical parameter b^* and realizing the remaining part $c - b^*$ by means of a diverging sequence $k^{(n)}$, such that the mesoscopic rate term appears.

The proof will be a consequence of the next lemmas. In Lemma 65 we derive an upper and a lower bound for $\inf_{\lambda\colon c_\lambda = c} I_{\mathrm{Mi}}(\lambda)$ and in Lemma 66 we show that they coincide, if there are solutions to the fixed point equation (6.27). We will postpone the proof about existence of solutions to Section 6.4, Lemma 68.

Lemma 65. *Let $c \in [0,\mu]$ with $\Sigma(\kappa, c) > 1$. For $b \in [0,c]$ with $\Sigma(\kappa, b) \le 1$ we put*

$$F_c(b) = \Big\langle c, \log\frac{b}{\mu}\Big\rangle + |c-b| + \frac{1}{2}\langle b, \kappa b\rangle - \langle c, \kappa b\rangle + \frac{1}{2}\langle c, \kappa\mu\rangle, \tag{6.28}$$

$$G_c(b) = \Big\langle b, \log\frac{b}{\mu}\Big\rangle - \frac{1}{2}\langle b, \kappa b\rangle + \Big\langle c-b, \log\frac{c-b}{(\kappa(c-b))\mu}\Big\rangle + \frac{1}{2}\langle c, \kappa\mu\rangle. \tag{6.29}$$

Then

$$\sup_{b\in[0,c]\colon \Sigma(\kappa,b)\le 1} F_c(b) \le \inf_{\lambda\colon c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) \le \inf_{b\in[0,c]\colon \Sigma(\kappa,b)\le 1} G_c(b). \tag{6.30}$$

Proof. We first show the first inequality in (6.30). Fix $b \in [0,c]$ with $\Sigma(\kappa, b) \le 1$. Let $\lambda^* := \lambda(b)$ be given as in (6.3). We proceed in the same way as in the proof of Proposition 56, but use that this time $|\lambda^*| = |b| - \frac{1}{2}\langle b, \kappa b\rangle$. Then for any λ with $c(\lambda) = c$ we have

$$I_{\mathrm{Mi}}(\lambda) = \mathbb{H}(\lambda|\lambda^*) - |\lambda^*| + \Big\langle c, \log\frac{b}{\mu}\Big\rangle - \langle c, \kappa b\rangle + |c| + \frac{1}{2}\langle c, \kappa\mu\rangle \ge F_c(b).$$

We now prove the upper bound in (6.30). Fix $b \in [0,c]$ with $\Sigma(\kappa, b) \le 1$.

Case 1: First, we assume that κ is irreducible with respect to $c - b$. Let $\lambda^* := \lambda(b)$ be defined as in (6.3). For $n \in \mathbb{N}$ define $k^{(n)} := \lfloor n(c-b)\rfloor$ and write $R_n := |k^{(n)}|$ and $b^{(R_n)} := \sum_{|k|\le R_n} \lambda^*_k k$. We define

$$\lambda^{(n)}_k := \begin{cases} \lambda^*_{\mathbf{e}_s} + \varepsilon_s & \text{if } k = \mathbf{e}_s \text{ for some } s \in \mathcal{S} \\ \lambda^*_k & \text{if } 1 < |k| < R_n \\ \frac{1}{n} & \text{if } k = k^{(n)} \\ 0 & \text{else} \end{cases} \tag{6.31}$$

with $\varepsilon := c - b^{(R_n)} - \frac{1}{n}k^{(n)}$, which ensures that $c(\lambda^{(n)}) = c$ holds for all $n \in \mathbb{N}$, but is negligible in the limit, i.e., $\lim_{n\to\infty} \lambda^{(n)}_{\mathbf{e}_s} = \lambda^*_{\mathbf{e}_s}$ for all $s \in \mathcal{S}$. Note that due to the irreducibility assumption we have that $\tau(k^{(n)}) > 0$ if n is large enough, which ensures that $I_{\mathrm{Mi}}(\lambda^{(n)})$ is finite.

Using the notation $I^{(R)}_{\mathrm{Mi}}(\lambda)$ introduced in (4.16) we get

$$I_{\mathrm{Mi}}(\lambda^{(n)}) = I^{(R_n-1)}_{\mathrm{Mi}}(\lambda^{(n)}) + \frac{1}{2}\langle c - b^{(R_n)}, \kappa\mu\rangle + \frac{1}{n}\log\frac{\frac{1}{n}\prod_s k^{(n)}_s!}{\tau(k^{(n)})\mathrm{e}^{1-|k^{(n)}|}\prod_s \mu_s^{k^{(n)}_s}}$$

Denote the last summand as A_n. By using the formula (3.10) from Lemma 18 for some $r \in \mathrm{supp}(c)$ as well as Stirling's formula for the factorial terms, we have that, as $n \to \infty$

$$A_n = o(1) + \sum_{s\in\mathcal{S}} (c_s - b_s) \log \frac{c_s - b_s}{(\kappa(c-b)_s)\mu_s} + \frac{1}{n} \log \frac{n(c_r - b_r) \prod_s \kappa(c-b)_s}{\mathrm{e}\Delta_r(c-b)}, \tag{6.32}$$

where Δ_r is defined in (3.11), which can be easily extended to arguments in $[0,\infty)^{\mathcal{S}}$. Note that by construction $\Delta_r(c-b) > 0$. Clearly, the last summand in (6.32) is of order $o(1)$. By the construction of $\lambda^{(n)}$ it is immediate that $\lim_{n\to\infty} I_{\mathrm{Mi}}^{(R_n-1)}(\lambda^{(n)}) = \langle b, \log(b/\mu)\rangle + \frac{1}{2}\langle b, \kappa(\mu - b)\rangle$, so altogether we get that

$$\lim_{n\to\infty} I_{\mathrm{Mi}}(\lambda^{(n)}) = \Big\langle b, \log\frac{b}{\mu}\Big\rangle + \frac{1}{2}\langle b, \kappa(\mu - b)\rangle + \langle c - b, \log\frac{c-b}{(\kappa(c-b))\mu}\rangle + \frac{1}{2}\langle c - b, \kappa\mu\rangle = G_c(b).$$

Case 2: If κ is reducible with respect to $c - b$, we can find a decomposition of $\mathrm{supp}(c-b)$ into disjoint sets $\mathcal{S}_j$ such that κ restricted to $\mathcal{S}_j \times \mathcal{S}_j$ is irreducible and $\kappa|_{\mathcal{S}_i\times\mathcal{S}_j} = 0$ for $i \neq j$. Then we have to modify the construction of $\lambda^{(n)}$ given above by putting mass $\frac{1}{n}$ on each of the meso-particles $k^{(j,n)} := \lfloor n(c-b)\rfloor \mathbb{1}_{\mathcal{S}_j}$. We omit the details. □

The following lemma completes the proof of Proposition 64. Its assumptions are verified later and in more generality in Lemma 68.

Lemma 66. *Let $c \in [0,\mu]$ with $\Sigma(\kappa, c) > 1$. If there exists a non-trivial solution $b^* \in [0,c]$ to* (6.27) *and $\Sigma(\kappa, b^*) = 1$, then $F(b^*) = G(b^*)$. Consequently, equation* (6.26) *holds.*

Proof. Using the fixed point equation (6.27) we can substitute $|c - b^*| = \langle b^*, \kappa(c-b^*)\rangle$ to rewrite $F_c(b^*)$ and $\langle c - b^*, \log[(c-b^*)/\kappa(c-b^*)] = \langle c - b^*, \log b^*\rangle$ to rewrite $G_c(b^*)$. Then $F_c(b^*) = G_c(b^*)$. □

6.4 The general supercritical case

Building on the results of the previous subsection we derive a slightly weaker result than Proposition 64 for the general case, which will still be enough to derive the optimal rates for the contraction principle as well as to fully optimize the rate function I.

Lemma 67. *Fix $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu$ and $\Sigma(\kappa, c) > 1$. Then*

$$\inf_{\lambda\in\mathcal{L}:\, c_\lambda = c} I_{\mathrm{Mi}}(\lambda) \ge \inf_{\lambda\in\mathcal{L}:\, c_\lambda = b^*} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Me}}(c - b^*), \tag{6.33}$$

where $b^* = b^*(c) \in \mathcal{M}(\mathcal{S})$ is the minimal, non-trivial (i.e., not equal to c) solution to (2.9) *and satisfies $\Sigma(\kappa, b^*) = 1$.*

Sketch of proof. We can generalize the proof of the lower bound of Lemma 65 and the definition of F_c to obtain

$$\inf_{\lambda:\, c_\lambda = c} I_{\mathrm{Mi}}(\lambda) \ge \Big\langle c, \log\frac{\mathrm{d}b}{\mathrm{d}\mu}\Big\rangle + c(\mathcal{S}) - b(\mathcal{S}) + \frac{1}{2}\langle b, \kappa b\rangle - \langle c, \kappa b\rangle =: F_c(b)$$

for any $b \in \mathcal{M}(\mathcal{S})$ with $b \leq c$ and $\Sigma(\kappa, b) \leq 1$. This relies on the admissibility of the (auxiliary) minimizers λ_b proved in Lemmas 61 and 62. The lower bound is obtained by writing everything with entropies as in the proof of Lemma 63.

Observe that if b^* is a solution of (2.9), then one can argue as in the proof of Lemma 66 to see that

$$F_c(b^*) = \Big\langle b^*, \log \frac{\mathrm{d}b^*}{\mathrm{d}\mu} \Big\rangle - \frac{1}{2}\langle b^*, \kappa b^* \rangle + I_{\mathrm{Me}}(c - b^*) + \frac{1}{2}\langle c, \kappa\mu \rangle =: G_c(b^*)$$

□

Lemma 68 (Solutions to (2.9))**.** *Fix $c \in \mathcal{M}(\mathcal{S})$.*

(i) Assume that κ is irreducible w.r.t. c and $\Sigma(\kappa, c) > 1$. Then there exists exactly one solution b^* to (2.9) *that satisfies $b^* \neq c$. Further, it holds that $\Sigma(\kappa, b^*) = 1$.*

(ii) If $\Sigma(\kappa, c) \leq 1$, then the only solution b^* to (2.9) *is given by the trivial solution $b^* = c$.*

(iii) Assume that κ is reducible w.r.t. c and $\Sigma(\kappa, c) > 1$, then there exists at least one solution b^* to (2.9) *with $b^* \neq c$. Moreover, there exists a unique minimal solution b_* to* (2.9) *(which is minimal in the sense that $b_* \leq b^*$ holds c-almost everywhere for all solutions b^* of* (2.9)*). Further, we have that $\Sigma(\kappa, b^*) > 1$ for all solutions b^* with $b^* \neq b_*$ and $\Sigma(\kappa, b_*) = 1$.*

Proof. We will study the existence and uniqueness of non-trivial solutions $f^* \colon \mathcal{S} \to [0, 1)$ to

$$T_{\kappa,c} f^* = \frac{f^*}{1 - f^*}. \tag{6.34}$$

By substituting $b^* = (1 - f^*)c$ it is easily seen that solving (6.34) is equivalent to solving (2.9).

(i) *Existence*: We once more reformulate (6.34) by substituting $g^* = f^*/(1 - f^*)$ (which is equivalent to $f^* = g^*(1 + g^*)$). Then (6.34) is equivalent to

$$U(g^*) := T_{\kappa,c}\Big(\frac{g^*}{1 + g^*}\Big) = g^*, \tag{6.35}$$

i.e., we are searching for a fixed point of U. Note that $g^*(s)/(1 + g^*(s)) \leq 1$ for all $s \in \mathcal{S}$. Together with the fact that κ is non-negative this implies that any solution g^* of (6.35) satisfies $g^* \leq T_{\kappa,c}\mathbb{1}$. Hence, it suffices to study the operator U on the domain $D = \{g \colon \mathcal{S} \to \mathbb{R} \colon 0 \leq g \leq T_{\kappa,c}\mathbb{1}\}$. We construct a solution iteratively by defining $g_0 := T_{\kappa,c}\mathbb{1}$ and $g_m := U(g_{m-1})$ for $m \in \mathbb{N}$. Since the function $x \mapsto x/(1 + x)$ is strictly increasing on $[0, \infty)$ and κ is non-negative, we have that $g \leq g'$ implies $U(g) \leq U(g')$. Since $g_1 \leq g_0$ we can iterate this argument to show that $g_m \leq g_{m-1}$ holds for any $m \in \mathbb{N}$. Therefore the limit $g^* := \lim_{m\to\infty} g_m \in D$ exists and by the continuity of U it satisfies (6.35). We claim that our assumptions on κ and c imply that $g^* > 0$: By the assumptions that κ is irreducible w.r.t. c, $\mathcal{S}$ is compact and κ is continuous, $T_{\kappa,c}$ is a positive, irreducible and compact operator. Therefore there exists a strictly positive eigenfunction v of $T_{\kappa,c}$ with eigenvalue $\Sigma(\kappa, c) > 1$. Note that the function $T_{\kappa,c}v$ is continuous (by compactness of $\mathcal{S}$ and continuity of κ), hence v is continuous and by compactness of $\mathcal{S}$ it is also bounded. So without loss of generality we can pick v such that $v(s) \in (0, 1]$ for any $s \in \mathcal{S}$. Observe that by the irreducibility assumption $g_0 = T_{\kappa,c}\mathbb{1} > 0$. Now, pick $\delta > 0$ such that $\Sigma(\kappa, c) \geq 1 + \delta$ and $g_0 \geq \delta v$. Observe that for any g with $g \geq \delta v$ we have that

$$U(g) = T_{\kappa,c}\Big(\frac{g}{1 + g}\Big) \geq T_{\kappa,c}\Big(\frac{\delta v}{1 + \delta v}\Big) \geq \frac{\delta}{1 + \delta} T_{\kappa,c} v \geq \delta v.$$

Hence $g_m \geq \delta v$ holds for all $m \in \mathbb{N}$. Consequently, $g^* \geq \delta v > 0$.

Additionally, we claim that g^* is the maximal solution to (6.35). Let $\tilde{g}^*$ be any other solution to (6.35), then we necessarily have that $\tilde{g}^* \leq T_{\kappa,c}\mathbb{1} = g_0$. It follows by the monotonicity of U that $\tilde{g}^* = U(\tilde{g}^*) \leq g_1$ and, iteratively, $\tilde{g}^* \leq g_m$ for any $m \in \mathbb{N}$. Hence $\tilde{g}^* \leq g^*$. By the equivalence of (6.34) and (6.35) and monotonicity of $x \mapsto x/(1+x)$ we have that $f^* = g^*/(1+g^*)$ is the maximal solution to (6.34).

Uniqueness: Assume towards a contradiction that f^* and $\tilde{f}^*$ are non-trivial solutions of (6.34) and $f^* \neq \tilde{f}^*$ on a set $A \subset \mathcal{S}$ with $c(A) > 0$. Without loss of generality we can assume that f^* is the maximal solution (as constructed in the existence part), i.e., $\tilde{f}^* \leq f^*$ on $\mathcal{S}$ and $\tilde{f}^* < f^*$ everywhere on A. For any $h\colon \mathcal{S} \to [0,1]$ put $\Psi(h) := (1-h)T_{\kappa,c}h$. Then $\Psi(f^*) = f^*$ and $\Psi(\tilde{f}^*) = \tilde{f}^*$ and we have that

$$\begin{aligned}\Psi\Big(\frac{f^*}{2} + \frac{\tilde{f}^*}{2}\Big) &= \frac{1}{2}\Psi(f^*) + \frac{1}{2}\Psi(\tilde{f}^*) + \frac{1}{4}(f^* - \tilde{f}^*)T_{\kappa,c}(f^* - \tilde{f}^*) \\ &\geq \frac{1}{2}\Psi(f^*) + \frac{1}{2}\Psi(\tilde{f}^*) = \frac{1}{2}f^* + \frac{1}{2}\tilde{f}^*,\end{aligned}$$

where the inequality relies on the fact that κ is non-negative and $\tilde{f}^* \leq f^*$. Note that we even have a strict inequality on the set A. Now, define $h := (f^* - \tilde{f}^*)/2$, then $\tilde{f}^* + h = (\tilde{f}^* + f^*)/2$ and we already argued that $\Psi(\tilde{f}^* + h) \geq \tilde{f}^* + h$ where the inequality is strict on A. On the other hand, as $h \geq 0$, we get

$$\Psi(\tilde{f}^* + h) = (1 - \tilde{f}^* - h)T_{\kappa,c}(\tilde{f}^* + h) \leq (1 - \tilde{f}^*)T_{\kappa,c}(\tilde{f}^* + h),$$

so altogether $(1 - \tilde{f}^*)T_{\kappa,c}(\tilde{f}^* + h) \geq \tilde{f}^* + h$ with strict inequality on A. Using (6.34) for $\tilde{f}^*$ and the symmetry of κ we get

$$\langle c, \tilde{f}^* T_{\kappa,c}(\tilde{f}^* + h)\rangle > \Big\langle c, \frac{\tilde{f}^*}{1 - \tilde{f}^*}(\tilde{f}^* + h)\Big\rangle = \langle c, (T_{\kappa,c}\tilde{f}^*)(\tilde{f}^* + h)\rangle = \langle c, \tilde{f}^* T_{\kappa,c}(\tilde{f}^* + h)\rangle,$$

which is a contradiction. Hence the solution to (6.34) is unique up to sets that have measure zero w.r.t. c, which implies uniqueness of b^*.

We now argue that $\Sigma(\kappa, b^*) = 1$. Our procedure is very similar to the one used in the proof of Lemma 6.6 in [BJR07]. Let $w\colon \mathcal{S} \to \mathbb{R}$ be an eigenfunction of T_{κ,b^*} with eigenvalue a. Then using (6.34) and the symmetry of κ, we get that

$$\begin{aligned}\langle c, wf^*\rangle &= \langle c, w(1-f^*)T_{\kappa,c}f^*\rangle = \langle c, f^* T_{\kappa,c}(w(1-f^*))\rangle \\ &= \langle c, f^* T_{\kappa,b^*}w\rangle = a\langle c, f^* w\rangle.\end{aligned} \tag{6.36}$$

Hence, we either have that $\langle c, wf^*\rangle = 0$ or $a = 1$. By the Krein–Rutman Theorem (the extension of the Perron–Frobenius Theorem to positive compact operators) the eigenfunction w corresponding to the largest eigenvalue of T_{κ,b^*} is non-negative and non-trivial, so $\langle c, wf^*\rangle > 0$ and hence $\Sigma(\kappa, b^*) = 1$. (Interestingly, we have constructed b^* in such a way that all other eigenfunctions $\widetilde{w}$ of T_{κ,b^*} satisfy $\langle c, \widetilde{w}f^*\rangle = 0$.)

(ii) Let $\Sigma(\kappa, c) \leq 1$. Assume towards a contradiction that f^* is a solution to (6.34) and $\int_A f^* \,\mathrm{d}c > 0$ for some open set $A \subset \mathcal{S}$ such that $c(A) > 0$. Using the substitution $f^* = g/(1+g)$ we have that equation (6.34) is equivalent to $T_{\kappa,c}(g/(1+g)) = g$ and since κ is continuous and $\mathcal{S}$ is compact the left-hand side $T_{\kappa,c}(g/(1+g))$ is a continuous function, which implies that both g and f^* are

continuous functions. So we can find an $\varepsilon > 0$ and a set $A_\varepsilon \subset A$ such that $f^* \geq \varepsilon$ on A_ε and $c(A_\varepsilon) > 0$. Therefore,

$$T_{\kappa,c} f^* = \frac{f^*}{1-f^*} \geq \frac{f^*}{1-\varepsilon}$$

holds on A_ε and $T_{\kappa,c} f^* \geq f^*$ holds on $\mathcal{S}$. This implies that $\|T_{\kappa,c} f^*\|_{L^2(c)} > \|f^*\|_{L^2(c)}$ and hence $\Sigma(\kappa, c) > 1$ in contradiction to our assumption.

(iii) Since κ is reducible w.r.t. c, we find a decomposition of $\mathrm{supp}(c)$ into (countable many) disjoint sets S_j, $j \in J$, such that $\kappa^{(j)} = \kappa|_{S_j \times S_j}$ is irreducible with respect to $c^{(j)}$, the restriction of c to S_j for any $j \in J$, and $\kappa|_{S_i \times S_j} = 0$ holds c-almost everywhere, if $i \neq j$. Let $J' := \{j \in J \colon \Sigma(\kappa^{(j)}, c^{(j)}) > 1\}$ and note that $J' \neq \emptyset$. By (i) we get that for any $j \in J'$ there exists a function $f^{(j)} \colon S_j \to [0,1)$ that solves (6.34) on S_j and $f^{(j)} > 0$. By (ii) we get that for any $j \in J \backslash J'$ the only function $f \colon S_j \to [0,1)$ that solves (6.34) on S_j is equal to 0 $c^{(j)}$-almost everywhere. Now for $\sigma = (\sigma_j)_{j \in J'} \in \{0,1\}^{J'}$ define $f^{(\sigma)} \colon \mathcal{S} \to [0,1)$ by $f^{(\sigma)}(s) = \sigma_j f^{(j)}(s)$, if $s \in S_j$ for some $j \in J'$ and $f^{(\sigma)}(s) = 0$ for $s \in \mathcal{S} \setminus \bigcup_{j \in J'} S_j$. It can now be easily checked that all solutions to (6.34) are given by

$$\mathcal{F} := \left\{ f^{(\sigma)} \colon \sigma = (\sigma_j)_{j \in J'} \in \{0,1\}^{J'} \right\} \tag{6.37}$$

and that $\mathcal{F}$ contains at least one non-trivial solution. Write $b^{(\sigma)} = (1 - f^{(\sigma)})c$ and note that all possible solutions of (2.9) are of this form. Clearly, the minimal solution b^* of (2.9) is given via the maximal solution in $\mathcal{F}$, i.e., by choosing $\sigma \equiv 1$.

We will now investigate the quantities $\Sigma(\kappa, b^{(\sigma)})$ for any choice of σ. First, let σ be such that there exists some $j \in J'$ with $\sigma_j = 0$. Then for $r \in S_j$ and any function $h \colon \mathcal{S} \to \mathbb{R}$

$$T_{\kappa,b^{(\sigma)}} h(r) = \int_{S_j} \kappa(r,s) h(s) (1 - \sigma_j f^{(j)}(s))\, c(\mathrm{d}s) = T_{\kappa^{(j)},c^{(j)}} h(r).$$

Therefore, given an eigenfunction $g^{(j)}$ of $T_{\kappa^{(j)},c^{(j)}}$ that corresponds to the eigenvalue $\Sigma(\kappa^{(j)}, c^{(j)})$, we can construct an eigenfunction g for $T_{\kappa,b^{(\sigma)}}$ with the same eigenvalue by choosing $g = g^{(j)}$ on S_j and $g = 0$ on $\mathcal{S} \setminus S_j$. Hence, $\Sigma(\kappa, b^{(\sigma)}) = \Sigma(\kappa^{(j)}, c^{(j)}) > 1$. Now, consider $\sigma \equiv 1$. Then we can argue as in (6.36) to show that $\Sigma(\kappa, b^{(\sigma)}) = 1$. □

7 Analysis of minimizers of the rate function in Theorem 1

In this section we provide the final steps needed for the optimization of the rate function and prove Theorems 9 and 7. Since the arguments for the remaining steps do not rely on discrete combinatorics (as it was the case in Section 6), we will immediately work in the general setting. In Section 7.1 we study a constrained optimization problem for the functions I_{Me} and I_{Ma}. In Section 7.2 we prove the explicit form of the rate functions that is derived by applying the contraction principle and formulated in Theorem 9. Section 7.3 presents the last step for a full optimization of the rate function I that gives Theorem 7. In Section 7.4 we derive the Flory equation that we formulated in Proposition 13.

7.1 The minimizers of I_{Ma} and I_{Me}

Complementary to what was done in Section 6 for the function I_{Mi} we will solve the analogous optimization problems for the functions I_{Me} and I_{Ma} defined in (1.12) and (1.13). To optimize the function

I_{Me} it is beneficial to combine it with I_{Ma}. We will again fix a measure $c \in \mathcal{M}(\mathcal{S})$ to formulate the constraint. In contrast to the result of Proposition 55 it will turn out that we do not have to distinguish between the cases $\Sigma(\kappa, c) \leq 1$ and $\Sigma(\kappa, c) > 1$.

Lemma 69 (Minimizers of I_{Me} and I_{Ma})**.** *Let $c \in \mathcal{M}(\mathcal{S})$ be such that $c \leq \mu$. Then*

$$\inf\{I_{\mathrm{Ma}}(\alpha) \colon \alpha \in \mathcal{A}, c(\alpha) = c\} = I_{\mathrm{Ma}}(\delta_c), \tag{7.1}$$

$$\inf\{I_{\mathrm{Ma}}(\alpha) + I_{\mathrm{Me}}(\nu) \colon \alpha \in \mathcal{A}, \nu \in \mathcal{M}(\mathcal{S}), c(\alpha) + \nu = c\} = I_{\mathrm{Ma}}(\delta_c) + I_{\mathrm{Me}}(0). \tag{7.2}$$

If κ is irreducible with respect to c, then the minimizers are unique.

Proof. Writing $\widehat{I}_{\mathrm{Ma}}(\alpha) = I_{\mathrm{Ma}}(\alpha) - \frac{1}{2}\langle c(\alpha), \kappa\mu\rangle$ and $\widehat{I}_{\mathrm{Me}}(\nu) = I_{\mathrm{Me}}(\nu) - \frac{1}{2}\langle \nu, \kappa\mu\rangle$ it suffices to prove equations (7.1) and (7.2) for $\widehat{I}_{\mathrm{Ma}}$ and $\widehat{I}_{\mathrm{Me}}$ since the difference does not depend on c. Observe that $\widehat{I}_{\mathrm{Ma}}(\alpha) = A(\alpha) + B(\alpha)$, where

$$A(\alpha) = \int_{\mathcal{S}} \mu(\mathrm{d}r) \int \alpha(\mathrm{d}y) \frac{\mathrm{d}y}{\mathrm{d}\mu}(r) \log \frac{\frac{\mathrm{d}y}{\mathrm{d}\mu}(r)}{\kappa y(r)}, \tag{7.3}$$

$$B(\alpha) = \int_{\mathcal{S}} \mu(\mathrm{d}r) \int \alpha(\mathrm{d}y) \frac{\mathrm{d}y}{\mathrm{d}\mu}(r) \log \frac{\kappa y(r)}{\mathrm{e}^{\frac{1}{2}\kappa y(r)} - \mathrm{e}^{-\frac{1}{2}\kappa y(r)}}. \tag{7.4}$$

Let $\alpha \in \mathcal{A}$ with $c(\alpha) = c$. Note that $\int \alpha(\mathrm{d}y)\, \kappa y(r) = \kappa c(r)$ for $r \in \mathcal{S}$. With $\phi(u) = u \log u$ we use Jensen's inequality to get that

$$\begin{aligned} A(\alpha) &= \int_{\mathcal{S}} \mu(\mathrm{d}r) \kappa c(r) \int \alpha(\mathrm{d}y) \frac{\kappa y(r)}{\kappa c(r)} \phi\left(\frac{\frac{\mathrm{d}y}{\mathrm{d}\mu}(r)}{\kappa y(r)}\right) \geq \int_{\mathcal{S}} \mu(\mathrm{d}r) \kappa c(r) \phi\left(\int \alpha(\mathrm{d}y) \frac{\frac{\mathrm{d}y}{\mathrm{d}\mu}(r)}{\kappa c(r)}\right) \\ &= \int_{\mathcal{S}} \mu(\mathrm{d}r) \frac{\mathrm{d}c}{\mathrm{d}\mu}(r) \log \frac{\frac{\mathrm{d}c}{\mathrm{d}\mu}(r)}{\kappa c(r)} = A(\delta_c), \end{aligned} \tag{7.5}$$

where we used that $\int \alpha(\mathrm{d}y) \frac{\mathrm{d}y}{\mathrm{d}\mu}(r) = \frac{\mathrm{d}c}{\mathrm{d}\mu}(r)$. We now derive a corresponding lower bound for $B(\alpha)$. Note that the function $u \mapsto u/\sinh(u) =: \psi(u)$ is strictly decreasing on $[0, \infty)$ and that for any $y \in \operatorname{supp}(\alpha)$ we have that $\kappa y \leq \kappa c$. Therefore

$$\begin{aligned} B(\alpha) &= \int_{\mathcal{S}} \mu(\mathrm{d}r) \int \alpha(\mathrm{d}y) \frac{\mathrm{d}y}{\mathrm{d}\mu}(r) \log \psi\left(\tfrac{1}{2}\kappa y(r)\right) \geq \int_{\mathcal{S}} \mu(\mathrm{d}r) \int \alpha(\mathrm{d}y) \frac{\mathrm{d}y}{\mathrm{d}\mu}(r) \log \psi\left(\tfrac{1}{2}\kappa c(r)\right) \\ &= \int_{\mathcal{S}} \mu(\mathrm{d}r) \frac{\mathrm{d}c}{\mathrm{d}\mu}(r) \log \psi\left(\tfrac{1}{2}\kappa c(r)\right) = B(\delta_c). \end{aligned} \tag{7.6}$$

This implies equation (7.1). Now, let $\alpha \in \mathcal{A}$ and $\nu \in \mathcal{M}(\mathcal{S})$ be such that $c(\alpha) + \nu = c$. Note that

$$\widehat{I}_{\mathrm{Ma}}(\alpha) + \widehat{I}_{\mathrm{Me}}(\nu) = A(\alpha + \delta_\nu) + B(\alpha).$$

Since $c(\alpha + \delta_\nu) = c$ holds, we can use the estimate from before to get that $A(\alpha + \delta_\nu) \geq A(\delta_c)$. To get the estimate for $B(\alpha)$, observe that we still have that $\kappa y \leq \kappa c$ for any $y \in \operatorname{supp}(\alpha)$. So,

$$B(\alpha) \geq \int_{\mathcal{S}} \mu(\mathrm{d}r) \frac{\mathrm{d}c(\alpha)}{\mathrm{d}\mu}(r) \log \psi\left(\tfrac{1}{2}\kappa c(r)\right) \geq \int_{\mathcal{S}} \mu(\mathrm{d}r) \frac{\mathrm{d}c}{\mathrm{d}\mu}(r) \log \psi\left(\tfrac{1}{2}\kappa c(r)\right) = B(\delta_c),$$

where the second estimate is due to the the fact that $\psi(u) \in [0, 1]$ for $u \geq 0$ and thus $\log \psi(\frac{1}{2}\kappa c(r)) \leq 0$. Combining the estimates gives equation (7.2).

For both uniqueness claims we rely on Lemma 70 below.

To show uniqueness of the minimizer in equation (7.1) assume that $\alpha \in \mathcal{A}$ with $c(\alpha) = c$ and $\alpha \neq \delta_c$. Without loss of generality we can assume that $I_{\mathrm{Ma}}(\alpha) < \infty$. Now, we only have to note that the inequality in the estimate (7.6) is a strict inequality, since by Lemma 70 we have that $\kappa y < \kappa c$ holds on some set A, for which $y(A) > 0$ and the function ψ is strictly decreasing.

To show uniqueness of the minimizer in equation (7.2) assume that $\alpha \in \mathcal{A}$ and $\nu \in \mathcal{M}(\mathcal{S})$ with $\nu \neq 0$ and $c(\alpha) + \nu = c$. Then by the same arguments as before

$$B(\alpha) \geq B(\delta_{c(\alpha)}) = \int c(\alpha)(\mathrm{d}r) \log \psi\left(\kappa c(\alpha)(r)\right) > \int c(\alpha)(\mathrm{d}r) \log \psi\left(\kappa c(r)\right) \geq B(\delta_c)$$

holds, if $\kappa c > \kappa c(\alpha)$ on some measurable set $A \subset \mathcal{S}$ for which $c(\alpha)(A) > 0$. To see that the latter condition is satisfied, we apply Lemma 70 to the measure $\delta_{c(\alpha)} + \delta_\nu$. This proves the claim. □

Lemma 70. *Let $c \in \mathcal{M}(\mathcal{S})$ and let κ be irreducible with respect to c. Let $\alpha \in \mathcal{A}$ be such that $c(\alpha) = c$ and assume that $\alpha = \sum_{i \in I} \delta_{y^{(i)}}$ with $|I| \geq 2$ and $I_{\mathrm{Ma}}(\alpha) < \infty$. Then for any fixed $i \in I$ there is a measurable set $A \subset \mathcal{S}$ such that $y^{(i)}(A) > 0$ and $\kappa(c - y^{(i)})(x) > 0$ for all $x \in A$.*

Proof. Denote $y := y^{(i)}$, $S_1 = \mathrm{supp}(y)$ and $S_2 = \mathrm{supp}(c(\alpha) - y)$. We first study the case where the sets S_1 and S_2 are disjoint. Assume towards a contradiction that for y-almost every r we have that $\kappa(c-y)(r) = 0$. Then $\kappa|_{S_1 \times S_2} = 0$ which is equivalent to saying that $\kappa|_{S_1 \times \mathcal{S} \setminus S_1} = 0$ holds c-almost everywhere. Since κ is irreducible with respect to c we get that either $c(S_1) = 0$ or $c(\mathcal{S} \setminus S_1) = 0$. Consequently, either $y(S_1) = 0$ or $(c-y)(\mathcal{S} \setminus S_1) = 0$, i.e., either $y = 0$ or $c - y = 0$ in contradiction to our assumptions.

Now, assume that S_1 and S_2 are not disjoint. In that case we can pick $j \in I \setminus \{i\}$ in such a way that with $\hat{y} := y^{(j)}$ there exists $r_0 \in \mathrm{supp}(y) \cap \mathrm{supp}(\hat{y})$, which implies that for any open neighborhood $A_0 \subset \mathcal{S}$ of r_0 we have that $y(A_0) > 0$ and $\hat{y}(A_0) > 0$. By our assumption $I_{\mathrm{Ma}}(\alpha) < \infty$, we have that $-\infty < \langle \hat{y}, \log(1 - \mathrm{e}^{-\kappa \hat{y}}) \rangle$. This, together with the uniform continuity of κ implies that we can find a neighborhood A_0 of r_0 such that $\kappa \hat{y}|_{A_0} > 0$. Now the claim follows, since $\kappa(c - y) \geq \kappa \hat{y}$ and $y(A_0) > 0$. □

7.2 Minimization for the contraction principles

Here, we will exploit the work of Sections 6 and 7.1 to prove Theorem 9, which is an application of the contraction principle but also provides an explicit solution for the optimization problem. When studying the optimization problem in the large deviation principle for Ma_N, we encounter a functional that combines rates coming from the microscopic and the mesoscopic part. Its optimization is derived in the following lemma.

Lemma 71. *Fix $c \in \mathcal{M}(\mathcal{S})$ with $c \leq \mu$. For $b \in \mathcal{M}(\mathcal{S})$ with $b \leq c$ and $\Sigma(\kappa, b) \leq 1$ let G_c be as in* (6.29)

$$G_c(b) := \left\langle b, \log \frac{\mathrm{d}b}{\mathrm{d}\mu} \right\rangle - \frac{1}{2}\langle b, \kappa b \rangle + \left\langle c - b, \log \frac{\mathrm{d}(c-b)}{\kappa(c-b)\mathrm{d}\mu} \right\rangle + \frac{1}{2}\langle c, \kappa \mu \rangle.$$

Then the following holds.

1 *If $\Sigma(\kappa, c) \leq 1$, then*

$$\min\left\{G_c(b) \colon b \in \mathcal{M}(\mathcal{S}), b \leq c, \Sigma(\kappa, b) \leq 1\right\} = G_c(c), \tag{7.7}$$

and c is the unique minimizer.

2 If $\Sigma(\kappa, c) > 1$*, then*

$$\min\{G_c(b)\colon b \in \mathcal{M}(\mathcal{S}), b \le c, \Sigma(\kappa, b) \le 1\} = G_c(b^*(c)) \tag{7.8}$$

and $b^*(c)$ *is the unique minimizer, which is given as the minimal, non-trivial (i.e., not equal to* c*) solution of* (2.9)*, and it satisfies* $\Sigma(\kappa, b^*(c)) = 1$*.*

Proof. (1) We use that $\frac{1}{2}\langle c, \kappa c\rangle - \frac{1}{2}\langle b, \kappa b\rangle = \langle b, \kappa(c-b)\rangle + \frac{1}{2}\langle c-b, \kappa(c-b)\rangle$ holds by the symmetry of κ. Therefore,

$$\begin{aligned} G_c(b) - G_c(c) &= \Big\langle b, \log\frac{\mathrm{d}b}{\mathrm{d}c}\Big\rangle - \frac{1}{2}\langle b, \kappa b\rangle + \Big\langle c-b, \log\frac{1-\frac{\mathrm{d}b}{\mathrm{d}c}}{\kappa(c-b)}\Big\rangle + \frac{1}{2}\langle c, \kappa c\rangle \\ &= \Big\langle b, \log\frac{\frac{\mathrm{d}b}{\mathrm{d}c}}{\mathrm{e}^{-\kappa(c-b)}}\Big\rangle + \Big\langle c-b, \log\frac{1-\frac{\mathrm{d}b}{\mathrm{d}c}}{\kappa(c-b)\mathrm{e}^{-\frac{1}{2}\kappa(c-b)}}\Big\rangle. \end{aligned}$$

We claim that $\kappa(c-b)\mathrm{e}^{-\frac{1}{2}\kappa(c-b)} \le 1 - \mathrm{e}^{-\kappa(c-b)}$ holds pointwise. Indeed, the function $\psi(z) := 1 - \mathrm{e}^{-z} - z\mathrm{e}^{-\frac{z}{2}}$, $z \ge 0$, satisfies that $\psi(0) = 0$ and $\psi'(z) = \mathrm{e}^{-\frac{z}{2}}(\mathrm{e}^{-\frac{z}{2}} - (1 - \frac{z}{2})) \ge 0$ for any $z \ge 0$, implying that $\psi(z) \ge 0$ for any $z \ge 0$. So the claim holds, since $\kappa(c-b) \ge 0$ holds pointwise. Therefore, we can estimate

$$\begin{aligned} G_c(b) - G_c(c) &\ge \Big\langle b, \log\frac{\frac{\mathrm{d}b}{\mathrm{d}c}}{\mathrm{e}^{-\kappa(c-b)}}\Big\rangle + \Big\langle c-b, \log\frac{1-\frac{\mathrm{d}b}{\mathrm{d}c}}{1-\mathrm{e}^{-\kappa(c-b)}}\Big\rangle \\ &= \int_{\mathcal{S}} c(\mathrm{d}r)\Big[\frac{\mathrm{d}b}{\mathrm{d}c}(r)\log\frac{\frac{\mathrm{d}b}{\mathrm{d}c}(r)}{\mathrm{e}^{-\kappa(c-b)(r)}} + \Big(1 - \frac{\mathrm{d}b}{\mathrm{d}c}(r)\Big)\log\frac{1-\frac{\mathrm{d}b}{\mathrm{d}c}(r)}{1-\mathrm{e}^{-\kappa(c-b)(r)}}\Big] \ge 0, \end{aligned}$$

with equality if and only if $b = c$. The last inequality can be seen by applying Jensen's inequality to the function $x \mapsto x\log x$ or by noting the following: For any point $r \in \mathcal{S}$ the term in brackets in the last line is an entropy between the Bernoulli distribution with (success) parameter $\frac{\mathrm{d}b}{\mathrm{d}c}(r)$ and the Bernoulli distribution with parameter $\mathrm{e}^{-\kappa(c-b)(r)}$ and therefore non-negative.

(2) Define F_c as the generalized analog of (6.28), i.e., for $b \le c$

$$F_c(b) = \Big\langle c, \log\frac{\mathrm{d}b}{\mathrm{d}\mu}\Big\rangle + (c-b)(\mathcal{S}) + \frac{1}{2}\langle b, \kappa b\rangle - \langle c, \kappa b\rangle + \frac{1}{2}\langle c, \kappa\mu\rangle.$$

Let $b, b' \le c$ with $\Sigma(\kappa, b) \le 1$ and $\Sigma(\kappa, b') \le 1$. We want to show that $F_c(b') \le G_c(b)$. By rearranging terms one can see that

$$G_c(b) - F_c(b') = \mathbb{H}(c-b|\kappa(c-b)b') + \mathbb{H}(b|b') - \frac{1}{2}\langle b-b', \kappa(b-b')\rangle. \tag{7.9}$$

Now, given the signed measure $b - b'$ we use the Hahn decomposition theorem to decompose $\mathcal{S}$ into two disjoint sets S_+, S_- with $S_+ \cup S_- = \mathcal{S}$ such that $\delta_+(\cdot) := (b-b')(\cdot \cap S_+)$ and $-\delta_-(\cdot) := -(b-b')(\cdot \cap S_-)$ are non-negative measures and $b - b' = \delta_+ - \delta_-$. Observe that

$$\frac{1}{2}\langle b-b', \kappa(b-b')\rangle = \frac{1}{2}\int_{S_+}(b-b')(\mathrm{d}s)\,\kappa(b-b')(s) + \frac{1}{2}\int_{S_-}(b'-b)(\mathrm{d}s)\,\kappa(b'-b)(s)..$$

We write $f_{b',b} := \frac{\mathrm{d}b'}{\mathrm{d}b}\mathbb{1}_{S_+}$ and denote by $\langle\cdot,\cdot\rangle_b$ the inner product on $L^2(b)$, i.e. $\langle f, g\rangle_b = \int f(s)g(s)\,b(\mathrm{d}s)$. Note that by the symmetry of κ we have $\langle f, T_{\kappa,b}g\rangle_b = \langle g, T_{\kappa,b}f\rangle_b$, so we have that $\Sigma(\kappa, b) =$

$\sup_{f\neq 0}\frac{\langle f,T_{\kappa,b}f\rangle_b}{\langle f,f\rangle_b}\leq 1$. Then

$$\begin{aligned}\frac{1}{2}\int_{S_+}(b-b')(\mathrm{d}s)\,\kappa(b-b')(s)&=\frac{1}{2}\langle \mathbb{1}_{S_+}-f_{b',b},T_{\kappa,b}(\mathbb{1}_{S_+}-f_{b',b})\rangle_b\\&\leq\frac{1}{2}\Sigma(\kappa,b)\langle \mathbb{1}_{S_+}-f_{b',b},\mathbb{1}_{S_+}-f_{b',b}\rangle_b\leq\frac{1}{2}\int_{S_+}b(\mathrm{d}s)\,(1-f_{b',b}(s))^2.\end{aligned}\tag{7.10}$$

An elementary analysis shows that $\frac{1}{2}(1-x)^2\leq -\log x+x-1$ for $x\in(0,1]$ with equality if and only if $x=1$, and since $b'\leq b$ on S_+ implies that $f_{b',b}(s)\in(0,1]$ for $s\in S_+$, we get that

$$\frac{1}{2}\int_{S_+}b(\mathrm{d}s)\,(1-f_{b',b}(s))^2\leq\int_{S_+}b(\mathrm{d}s)\Big(-\log f_{b',b}(s)+f_{b',b}(s)-1\Big)\tag{7.11}$$

Denote $f_{b,b'}:=\frac{\mathrm{d}b}{\mathrm{d}b'}\mathbb{1}_{S_-}$. Interchanging the roles of b and b' and replacing S_+ by S_- one can argue as in (7.10) to show that

$$\frac{1}{2}\int_{S_-}(b'-b)(\mathrm{d}s)\,\kappa(b'-b)(s)\leq\frac{1}{2}\int_{S_-}b'(\mathrm{d}s)\,(1-f_{b,b'}(s))^2.$$

An elementary analysis shows that $\frac{1}{2}(1-x)^2\leq x\log x+1-x$ for $x\in(0,1]$ with equality if and only if $x=1$, and since $b\leq b'$ on S_- implies that $f_{b,b'}(s)\in[0,1]$ for $s\in S_-$, we get that

$$\frac{1}{2}\int_{S_-}b'(\mathrm{d}s)\,(1-f_{b,b'}(s))^2\leq\int_{S_+}b'(\mathrm{d}s)\Big(f_{b,b'}(s)\log f_{b,b'}(s)+1-f_{b,b'}(s)\Big).\tag{7.12}$$

Note that the two expressions on the right-hand sides of (7.11) and (7.12) sum up to $\mathbb{H}(b|b')$, hence we have shown that

$$\mathbb{H}(b|b')-\frac{1}{2}\langle b-b',\kappa(b-b')\rangle\geq 0$$

and we have equality if and only if $b=b'$. Using this in equation (7.9) and the fact that the first entropy term in (7.9) is always non-negative, we get that

$$\sup_{b\leq c\colon\,\Sigma(\kappa,b)\leq 1}F_c(b)\leq\inf_{b\leq c\colon\,\Sigma(\kappa,b)\leq 1}G_c(b).$$

Note that $G_c(b)-F_c(b')=0$ if and only if the following conditions are satisfied: (i) $b=b'$, (ii) b is a solution of (2.9) and (iii) $\Sigma(\kappa,b)=1$. By Lemma 68 the only choice is given by $b=b'=b_*$, where b_* is the unique minimal solution of (2.9). Hence, the uniqueness claim holds. □

Proof of Theorem 9. The projection $(\lambda,\alpha)\mapsto\lambda$ is continuous with respect to the vague topology, so the contraction principle gives that the LDP for Mi_N holds with rate function

$$\begin{aligned}\mathcal{I}_{\mathrm{Mi}}(\lambda)&=\inf_{\alpha\colon\,c(\alpha)\leq\mu-c(\lambda)}I(\lambda,\alpha)\\&=I_{\mathrm{Mi}}(\lambda)+\inf_{\alpha\in\mathcal{A},\nu\in\mathcal{M}(\mathcal{S})\colon\,c(\alpha)+\nu=\mu-c(\lambda)}\Big(I_{\mathrm{Ma}}(\alpha)+I_{\mathrm{Me}}(\nu)\Big)\end{aligned}\tag{7.13}$$

assuming $c(\lambda)\leq\mu$ (the other case is trivial). By equation (7.2) of Lemma 69 we immediately get the representation for $\mathcal{I}_{\mathrm{Mi}}$ claimed in equation (2.6).

The projection $(\lambda, \alpha) \mapsto \alpha$ is continuous with respect to the chosen topology, so the contraction principle gives that the LDP for Ma_N holds with rate function

$$\begin{aligned}\mathcal{I}_{\mathrm{Ma}}(\alpha) &= \inf_{\lambda\colon c(\lambda)\le \mu - c(\alpha)} I(\lambda,\alpha)\\ &= I_{\mathrm{Ma}}(\alpha) + \inf_{c\in[0,\mu-c(\alpha)]}\Big(\inf_{\lambda\colon c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Me}}(\mu - c(\alpha) - c)\Big),\end{aligned} \tag{7.14}$$

assuming $c(\alpha) \le \mu$ (the other case is trivial). Define $\mu_\alpha = \mu - c(\alpha)$.

Now, assume that $\Sigma(\kappa, \mu_\alpha) \le 1$. Then for any fixed $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu_\alpha$ we have that $\Sigma(\kappa, c) \le 1$, so according to equation (6.1) of Proposition 55 we have

$$\inf_{\lambda\colon c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Me}}(\mu_\alpha - c) = G_{\mu_\alpha}(c)$$

with $G_{\mu_\alpha}(c)$ defined as in (6.29). By Lemma 71 we have that $\min_{c\le\mu_\alpha} G_{\mu_\alpha}(c) = G_{\mu_\alpha}(\mu_\alpha)$, which implies equation (2.7) under the assumption $\Sigma(\kappa, \mu_\alpha) \le 1$.

Now, assume that $\Sigma(\kappa, \mu_\alpha) > 1$. Then by Proposition 55 we have

$$\inf_{\lambda\colon c(\lambda)=c} I_{\mathrm{Mi}}(\lambda) + I_{\mathrm{Me}}(\mu_\alpha - c) = \begin{cases} G_{\mu_\alpha}(c) & \text{if } \Sigma(\kappa,c)\le 1,\\ G_c(b^*(c)) + \langle \mu_\alpha - c, \log\frac{\mu_\alpha - c}{\kappa(\mu_\alpha - c)}\rangle & \text{if } \Sigma(\kappa,c) > 1.\end{cases}$$

In the case $\Sigma(\kappa, c) > 1$, one can use the same argument as in (7.5) to show that $G_c(b^*(c)) + \langle \mu_\alpha - c, \log\frac{\mu_\alpha - c}{\kappa(\mu_\alpha - c)}\rangle > G_{\mu_\alpha}(b^*(c))$, where $b^* = b^*(c)$ is given as in (2.9). In particular, $\Sigma(\kappa, b^*(c)) = 1$ holds, so any possible minimizer has to be in the set $\{c\colon \Sigma(\kappa, c) \le 1\}$. Now, recall that due to Lemma 71

$$\inf_{c\colon \Sigma(\kappa,c)\le 1} G_{\mu_\alpha}(c) = G_{\mu_\alpha}(b^*(\mu_\alpha))$$

This proves the claim of equation (2.7) under the assumption $\Sigma(\kappa, \mu_\alpha) > 1$. □

7.3 The minimizers of I

Proof of Theorem 7. Note that $\inf I(\lambda, \alpha) = \inf_{c\in\mathcal{M}(\mathcal{S})\colon c\le\mu} \mathcal{J}(c)$ where for fixed $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu$ we define

$$\mathcal{J}(c) = \inf_{\lambda,\alpha: c(\lambda)=c} I(\lambda,\alpha) =: \begin{cases}\mathcal{J}_{\le 1}(c) & \text{if } \Sigma(\kappa,c)\le 1\\ \mathcal{J}_{>1}(c) & \text{otherwise.}\end{cases} \tag{7.15}$$

By Proposition 55 and Lemma 69 we have that

$$\begin{aligned}\mathcal{J}_{\le 1}(c) &= \Big\langle c, \log\frac{\mathrm{d}c}{\mathrm{d}\mu}\Big\rangle + \frac{1}{2}\langle c, \kappa(\mu - c)\rangle + I_{\mathrm{Ma}}(\delta_{\mu-c})\\ \mathcal{J}_{>1}(c) &\ge \Big\langle b^*, \log\frac{\mathrm{d}b^*}{\mathrm{d}\mu}\Big\rangle + \frac{1}{2}\langle b^*, \kappa(\mu - b^*)\rangle + I_{\mathrm{Me}}(c - b^*) + I_{\mathrm{Ma}}(\delta_{\mu-c})\end{aligned}$$

with $b^* = b^*(c)$ characterized in (2.9).

We will start by minimizing the function $\mathcal{J}_{\le 1}$ over all $c \in \mathcal{M}(\mathcal{S})$ with $c \le \mu$. Rearranging terms, we get that

$$\begin{aligned}\mathcal{J}_{\le 1}(c) &= \int_{\mathcal{S}} \mu(\mathrm{d}r)\Big(\frac{\mathrm{d}c}{\mathrm{d}\mu}(r)\log\frac{\frac{\mathrm{d}c}{\mathrm{d}\mu}(r)}{\mathrm{e}^{-\kappa(\mu-c)(r)}} + \Big(1 - \frac{\mathrm{d}c}{\mathrm{d}\mu}(r)\Big)\log\frac{1 - \frac{\mathrm{d}c}{\mathrm{d}\mu}(r)}{1 - \mathrm{e}^{-\kappa(\mu-c)(r)}}\Big)\\ &= \int_{\mathcal{S}} \mu(\mathrm{d}r) H(\beta^{(r)}|\gamma^{(r)}),\end{aligned}$$

where for $r \in \mathcal{S}$ we defined $\beta^{(r)}$ and $\gamma^{(r)}$ to be Bernoulli distributions with success rate $\frac{\mathrm{d}c}{\mathrm{d}\mu}(r)$ and $\mathrm{e}^{-\kappa(\mu-c)(r)}$, respectively (note that $c \leq \mu$ implies that $\frac{\mathrm{d}c}{\mathrm{d}\mu}(r) \leq 1$ for all $r \in \mathcal{S}$). By Jensen's inequality we have that $H(\beta^{(r)}|\gamma^{(r)}) \geq 0$ for all choices of $\beta^{(r)}$ and $\gamma^{(r)}$ and $H(\beta^{(r)}|\gamma^{(r)}) = 0$ if and only if $\beta^{(r)} = \gamma^{(r)}$, that is if and only if $\frac{\mathrm{d}c}{\mathrm{d}\mu}(r) = \mathrm{e}^{-\kappa(\mu-c)(r)}$. The minimizer of $\mathcal{J}_{\leq 1}$ is therefore characterized by (2.4).

In the case $\Sigma(\kappa, \mu) \leq 1$ (which implies $\Sigma(\kappa, c) \leq 1$ for all $c \in \mathcal{M}(\mathcal{S})$ with $c \leq \mu$), Lemma 16 states that μ is the only solution to (2.4).

Now assume that $\Sigma(\kappa, \mu) > 1$. Then for any $c \in \mathcal{M}(\mathcal{S})$ with $c \leq \mu$ and $\Sigma(\kappa, c) > 1$, Lemma 69 implies that $I_{\mathrm{Me}}(c - b^*) + I_{\mathrm{Ma}}(\delta_{\mu-c}) > I_{\mathrm{Ma}}(\delta_{\mu-b^*})$, where $b^* = b^*(c)$ is given as in (2.9) and satisfies $\Sigma(\kappa, b^*) = 1$. Therefore, $\mathcal{J}_{>1}(c) > \mathcal{J}_{\leq 1}(b^*(c))$, which implies that the minimizer of $\mathcal{J}$ lies in the set $\{c\colon \Sigma(\kappa, c) \leq 1\}$. (Note, that in this way, μ is ruled out as a minimizer, although it solves equation (2.4)). By the analysis of $\mathcal{J}_{\leq 1}$ above, the minimizer of $\mathcal{J}$ is given by a solution to (2.4) satisfying $\Sigma(\kappa, c) \leq 1$. Applying the second part of Lemma 16 finishes the proof. □

Remark 72. (Reducibility) *Theorem 7 and 9 are proved under the assumptions of Theorem 1, in particular when κ is irreducible with respect to μ. We see however that this condition does not play any role in the minimization of Proposition 55 and of Lemma 69. It is indeed Theorem 12 that excludes the admissibility of a minimizer of the form $(\lambda_{c^*}, \delta_{\mu-c^*})$ when $\Sigma(\kappa, \mu) > 1$ for $\widetilde{I}$, as $\alpha^* = \delta_{\mu-c^*}$ may not be connectable. It is straightforward to see that the optimal macroscopic mass in this case takes the form $\widetilde{\alpha}^* = \sum_n \delta_{y^{(n)}}$, with $y^{(n)}(\cdot) = (\mu - c^*)(\cdot \cap S^{(n)})$, for each irreducible class $S^{(n)}$.*

7.4 The Flory equation

As we explained in Section 2.4 the graph model studied in this paper has an important connection with a certain inhomogeneous coagulation process. In this section we prove that the statistics of the limiting microscopic cluster distribution, i.e., the minimizer of the rate function I, satisfy the Flory equation (2.14), the related deterministic PDE, over the entire time interval $[0, \infty)$, before and after the gelation time $t_{\mathrm{c}} = 1/\Sigma(\kappa, \mu)$. We prove it in the case of a finite type set $\mathcal{S}$.

We fix an irreducible symmetric matrix κ on the finite type space $\mathcal{S}$. Recall from Proposition 56 the explicit formula

$$\lambda_k(c; \kappa) = \tau(k; \kappa) \prod_{r \in \mathcal{S}} \frac{(c_r \mathrm{e}^{-(\kappa c)_r})^{k_r}}{k_r!}, \qquad k \in \mathbb{N}_0^{\mathcal{S}}, c \in (0, \infty)^{\mathcal{S}}, \tag{7.16}$$

for the minimizer of the microscopic rate function I_{Mi} (see also (6.3)). With this notation we stressed the dependence on κ, since we consider now $t\kappa$ instead of κ, where $t \in [0, \infty)$ is a time instant, writing $\lambda(\mu; t\kappa)$. Note that $\lambda(\mu; t\kappa)$ is the minimizer both for $t < t_{\mathrm{c}}$ and for $t \geq t_{\mathrm{c}}$, as the characteristic equation $c_r^*(t)\mathrm{e}^{-t(\kappa c^*(t))_r} = \mu_r \mathrm{e}^{-t(\kappa\mu)_r}$ (where $c^*(t)$ is the c^* of Proposition 56 for $t\kappa$ instead of κ) ensures that $\lambda(c^*(t); t\kappa) = \lambda(\mu; t\kappa)$. Note that $c(\lambda(c^*(t); t\kappa)) = c^*(t)$. We now show that the function $t \mapsto \lambda(\mu; t\kappa)$ solves the Flory equation that corresponds to our model.

Lemma 73. *The function $t \mapsto \lambda(\mu; t\kappa)$ is a solution of*

$$\frac{\mathrm{d}}{\mathrm{d}t} l_k(t) = \frac{1}{2} \sum_{m+\widetilde{m}=k} l_m(t) l_{\widetilde{m}}(t) \langle m, \kappa \widetilde{m} \rangle - l_k(t) \sum_m l_m(0) \langle k, \kappa m \rangle, \quad \textit{for } k \in \mathbb{N}_0^{\mathcal{S}}, \tag{7.17}$$

with initial condition $\lambda(\mu; 0) = \sum_{r \in \mathcal{S}} \mu_r \mathbb{1}_{\{\mathbf{e}_r\}}$.

Proof. The initial condition is easily checked.

Since any tree contributing to the term $\tau(k;t\kappa)$ has exactly $|k|-1$ edges, we can rewrite $\lambda(\mu;t\kappa)$ for any $t\geq 0$ as

$$\lambda_k(\mu;t\kappa)=t^{|k|-1}\tau(k;\kappa)\mathrm{e}^{-t\langle k,\kappa\mu\rangle}\prod_{r\in\mathcal{S}}\frac{\mu_r^{k_r}}{k_r!},\quad \text{for } k\in\mathbb{N}_0^{\mathcal{S}}. \tag{7.18}$$

Abreviating $\lambda(t):=\lambda(\mu;t\kappa)$, we get that

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_k(t)=(|k|-1)\,\frac{1}{t}\lambda_k(t)-\langle k,\kappa\mu\rangle\lambda_k(t). \tag{7.19}$$

Now, we study the first summand of the r.h.s. of (7.17). By first inserting (7.18) and then using the recursive equation (6.7) from Lemma 18, we have that

$$\begin{aligned}
&\frac{1}{2}\sum_{m+\widetilde{m}=k}\lambda_m(t)\lambda_{\widetilde{m}}(t)\langle m,\kappa\widetilde{m}\rangle\\
&\quad=\frac{1}{2}t^{|k|-2}\mathrm{e}^{-t\langle k,\kappa\mu\rangle}\Big(\prod_u\frac{\mu_u^{k_u}}{k_u!}\Big)\sum_{r,s}\kappa(r,s)\sum_{m+\widetilde{m}=k}\Big(\prod_u\frac{k_u!}{m_u!\widetilde{m}_u!}\Big)\tau(m)m_r\tau(\widetilde{m})\widetilde{m}_s\\
&\quad=(|k|-1)\,\frac{1}{t}\lambda_k(t).
\end{aligned}$$

Furthermore, we have that

$$\lambda_k(t)\sum_m\lambda_m(0)\langle k,\kappa m\rangle=\lambda_k(t)\langle k,\kappa\mu\rangle,$$

which implies the claim. □

References

[AKP19] Luisa Andreis, Wolfgang König, and Robert IA Patterson. A large-deviations principle for all the cluster sizes of a sparse erdős–rényi graph. *Random Structures & Algorithms*, 2019.

[BJR07] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

[BM14] Olivier Bernardi and Alejandro H. Morales. Counting trees using symmetries. *J. Comb. Theory Ser. A*, 123(1):104–122, April 2014.

[BvdHvL10] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, 15:no. 54, 1682–1703, 2010.

[BvdHvL12] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.*, 40(6):2299–2361, 2012.

[CCH20] Arijit Chakrabarty, Sukrit Chakraborty, and Rajat Subhra Hazra. Eigenvalues outside the bulk of inhomogeneous Erdős-Rényi random graphs. *J. Stat. Phys.*, 181(5):1746–1780, 2020.

[CRY21] Edward Crane, Balázs Ráth, and Dominic Yeo. Age evolution in the mean field forest fire model via multitype branching processes. *The Annals of Probability*, 49(4):2031–2075, 2021.

[DF14] Luc Devroye and Nicolas Fraiman. Connectivity of inhomogeneous random graphs. *Random Structures & Algorithms*, 45(3):408–420, 2014.

[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.

[Geo88] Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[Ges87] Ira M. Gessel. A combinatorial proof of the multivariable lagrange inversion formula. *Journal of Combinatorial Theory, Series A*, 45(2):178–195, 1987.

[Gil59] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30(4):1141–1144, 12 1959.

[JKT19] Sabine Jansen, Tobias Kuna, and Dimitrios Tsagkarogiannis. Virial inversion and density functionals. *arXiv preprint arXiv:1906.02322*, 2019.

[JKT21] Sabine Jansen, Tobias Kuna, and Dimitrios Tsagkarogiannis. Lagrange inversion and combinatorial species with uncountable color palette. In *Annales Henri Poincaré*, pages 1–36. Springer, 2021.

[MN14] Mathieu Merle and Raoul Normand. Self-organized criticality in a discrete model for smoluchowski's equation. *arXiv preprint arXiv:1410.8338*, 2014.

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

[NZ11] Raoul Normand and Lorenzo Zambotti. Uniqueness of post-gelation solutions of a class of coagulation equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(2):189–215, 2011.

[RT$^+$09] Balázs Ráth, Bálint Tóth, et al. Erdos-rényi random graphs+ forest fires= self-organized criticality. *Electronic Journal of Probability*, 14:1290–1327, 2009.

[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.

[vdH13] Remco van der Hofstad. Critical behavior in inhomogeneous random graphs. *Random Structures Algorithms*, 42(4):480–508, 2013.

[Yeo18] Dominic Yeo. Frozen percolation on inhomogeneous random graphs. *arXiv preprint arXiv:1810.02750*, 2018.