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

Preprint ISSN 2198-5855

A Gibbsian model for message routing in highly dense multi-hop networks

Wolfgang König [1,2], András Tóbiás [2]

submitted: April 25, 2017

[1] Weierstraß Institut
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: koenig@wias-berlin.de

[2] Technische Universität Berlin
Institut für Mathematik
Straße des 17. Juni 136
16023 Berlin
Germany
E-Mail: tobias@math.tu-berlin.de

No. 2392
Berlin 2017

2010 *Mathematics Subject Classification.* 60F10, 60G55, 60K30, 82B21, 90B15.

Key words and phrases. Gibbs distribution of trajectories, high-density limit, large deviations, empirical measure, variational formula, multihop ad-hoc network, signal-to-interference ratio, message trajectories, congestion.

The second author acknowledges support by the Berlin Mathematical School (BMS) and the DFG RTG 1845 "Stochastic Analysis with Applications in Biology, Finance and Physics".

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 Gibbsian model for message routing in highly dense multi-hop networks

Wolfgang König , András Tóbiás

Abstract

We investigate a probabilistic model for routing in relay-augmented multihop ad-hoc communication networks, where each user sends one message to the base station. Given the (random) user locations, we weigh the family of random, uniformly distributed message trajectories by an exponential probability weight, favouring trajectories with low interference (measured in terms of signal-to-interference ratio) and trajectory families with little congestion (measured by how many pairs of hops use the same relay). Under the resulting Gibbs measure, the system targets the best compromise between entropy, interference and congestion for a common welfare, instead of a selfish optimization.

We describe the joint routing strategy in terms of the empirical measure of all message trajectories. In the limit of high spatial density of users, we derive the limiting free energy and analyze the optimal strategy, given as the minimizer(s) of a characteristic variational formula. Interestingly, expressing the congestion term requires introducing an additional empirical measure.

1 Introduction

1.1 Background

In spatial wireless telecommunication systems, one of the prominent problems is the question how to conduct a message through the system in an optimal way. Optimality is often measured in terms of determining the shortest path from the transmitter to the recipient, or, if interference is considered, determining the path that yields the least interference. If many messages are considered at the same time, an additional aspect of optimality may be to achieve a minimal amount of congestion. These are problems of optimal routing through a network, a subject of mathematical traffic theory or optimization that is currently very popular and under demand.

Many investigations concern the question just for one single transmitter/recipient pair, which is a question that every single participant faces. However, a strategy found in such a setting may lead to a selfish routing, and it is quite likely that the totality of all these routings for all the individuals is by far not optimal for the community of all the users. Furthermore, the combinatorial or algorithmic efforts required to find all these optimal routings may be huge. Instead, the entire system may work even better if an optimal *compromise* is realized, by which we mean a joint strategy that leads to an optimum for the entire system, though possibly not for every participant. An additional benefit might be that it follows simple rules that are easy to implement and computationally little costly.

In this paper, we present a probabilistic ansatz for describing a jointly optimal routing that takes into account the following three crucial properties of the family of message trajectories: entropy, interference and

DOI 10.20347/WIAS.PREPRINT.2392

congestion. That is, we consider a situation in which all the messages are directed through the system in a random way, such that each hop prefers a low interference, and such that the total amount of congestion is preferred to be low. Parameters control the strengths of influence of the three effects.

Let us describe our model in words. Let the locations of all the users be given randomly as the sites of a Poisson point process, which we fix. Each user sends out precisely one message, which arrives at the (unique) base station, which is located at the origin. We consider the entire collection of possible trajectories of the messages through the system. We employ an ad-hoc relaying system with multiple hops, such that all the users act as relays for the handoffs of the messages. The maximal number of hops is $k_{\max} \in \mathbb{N}$ for each message. Each k-step message trajectory is random and *a priori* uniformly distributed. The family of all trajectories is *a priori* independent.

Now, the probability distribution of this family that we want to study is given in terms of a Gibbs ansatz by introducing two exponential weight terms. The first one weighs the total amount of interference, measured in terms of the signal-to-interference ratio for each hop, and the second one weighs the total congestion, i.e., the number of times that any two trajectories use the same relay. Under the arising measure, there is a competition between all the three decisive effects of the trajectory family: entropy, interference and congestion. Furthermore, the users form a random environment for the family, which not only determines the origins of all the trajectories, but also has a decisive effect on interference and congestion. While the latter has a smoothing effect on the fine details of the spatial distribution of all the trajectories, the effect of the former is not so clear to estimate, as the superposition of signals have a very non-local influence.

We consider this measure an interesting object to study. It describes an idealized situation in which the operator distributes all the message trajectories uniformly randomly and jointly optimizes the interference and the congestion of the entire system at the same time. Our main interest is in understanding the spatial distribution of the totality of all the message trajectories.

In this generality, the measure under consideration is a highly complex object, as it depends on all the user locations and on many detailed properties and quantities. However, we make a substantial step towards a thorough understanding by deriving approximative formulas for the behaviour in the limit of a high spatial density of the users. In this case, the limiting formulas turn out to be deterministic and to depend only on general spatial considerations, not on the individual users. It turns out that the limiting situation is described in terms of a large-deviation rate function and a variational formula, whose minimizers describe the optimal joint choices of the trajectories. These are our main results in this paper.

The main object in terms of which we achieve this description is the empirical measure of message trajectories sent out by the users, disintegrated with respect to their length and rescaled to finite asymptotic size. These measures turn out to converge in the weak topology in the high-density limit that we consider in this paper. The counting complexity of the statistics of the message trajectories can be written in terms of multinomial expressions and afterwards, in the limit of finer and finer decompositions of the space, approximated in terms of relative entropies, using to Stirling's formula. The interference term can also be handled in a standard way [HJKP15], since it is a continuous function of the collection of empirical measures of message trajectories.

However, a key finding of our paper is that the congestion term is a highly discontinuous function of these empirical measures. Indeed, one cannot express it in terms of these measures. Instead, one needs to substantially enlarge the probability space of trajectories and introduce another collection of empirical measures, the ones of the locations of users (relays) who receive given numbers of incoming messages. The congestion expression then turns out to be a lower semi-continuous function of these empirical measures, and hence the limiting congestion term is still expressible in terms of the weak limits of these measures. Again, using explicit combinatorics and Stirling's formula, we arrive at explicit entropic terms describing the statistics of these measures. The two families of these crucial empirical measures together enable us to describe all the properties of the message trajectories that we are interested in. We establish a full large-deviation principle for the tuple of all these measures with an explicit rate function and obtain in particular their convergence towards the minimizer(s) of a characteristic variational formula. We also derive their positivity properties and characterize them in terms of Euler-Lagrange equations.

The purpose of the present paper is to introduce the model, provide a mathematical framework and to establish the main analytical objects. However, there are a number of questions with regard to content about this model, which we do not address here. Here are some of these questions:

1 How does the number of hops of a message depend on the distance of the transmission site to the origin, e.g., in the long-distance limit?

2 Does the density of trajectories increase unboundedly in particularly highly dense areas, or do the messages avoid such areas for the sake of having lower interference?

3 How long is a typical average length of a hop? Does this average length depend much on whether it is one of the first hops or one of the last hops of the trajectory? Does it depend on the denseness of the area that the hop traverses?

4 How do these crucial quantities depend on the parameters of the model, in particular on β and γ, in particular in the limit of large values?

We decided to defer the analysis of such questions to future work, as they have a strongly analytic, rather than probabilistic, nature. Even though we are stressing the applied nature of the model and the questions, certainly the application of the mathematical framework that we introduce to telecommunication is by no means the only source of interest for such a model. Indeed, instead of the very particular choices of the interference and the congestion terms, our results can be easily extended to every other continuous (or at least lower semi-continuous) functionals of the crucial empirical measures, and applications are generally imaginable to other situations, e.g., in biology, chemistry or physics.

Apart from the potential value for the understanding of a new type of message routing models in telecommunication, the present paper provides also some interesting mathematical research on topological fine properties of random paths in random environment in a high-density setting, a subject that received a lot of interest for various types of such processes over the decades. We remind the reader on a number of investigations of the intersection properties of random walks and Brownian motions (both self-intersections and mutual intersections) in highly dense settings, see the monograph [Ch09] and some particular investigations in [KM02, KM13]; in all these works, one is interested in large-deviation properties of suitable empirical measures, and the lack of continuity of the path properties is the main difficulty. Let us mention that the main aspect of the approach in [KM02] is the same as in the present paper: an approximation of combinatorics in finer and finer decompositions of the space by entropic terms. Another line of research in which similar questions arise is a mean-field variant of a spatial version of Bose-Einstein statistics, like in [AK08], where the statistics of the empirical measures of a diverging number of Brownian bridges with symmetrized initial-terminal condition is analyzed in terms of a large-deviation principle in the weak topology. While [AK08] works with the same method as we in the present paper (spatial discretization with limiting fineness), [T08] showed that a method based entirely on the notion of entropy is able to derive such results in a more general framework.

Let us give a short guidance to the organization of the remainder of this paper. We introduce the model and necessary notation in Section 1.2, present our main results in Sections 1.3 (the limiting free energy of the model), 1.4 (the description of the minimizer(s)) and 1.5 (the large deviation principle and the convergence of the empirical measures), and in Section 1.6 we discuss and comment our findings. The remaining sections are devoted to the proofs: in Section 2 we prepare for the proofs by introducing our methods and deriving asymptotic formulas for the probability terms and the functionals, in Section 3 we put all this together to a proof of the limiting free energy, the large deviation principle and the convergence of the empirical measures, and in Section 4 we analyze the minimizer(s) of the characteristic variational formula.

1.2 The Gibbsian model

We introduce now the mathematical setting. For any $n \in \mathbb{N}$ and for any measurable subset V of $\mathbb{R}^n$, let $\mathcal{M}(V)$ denote the set of all finite nonnegative Borel measures on V. We are working in $\mathbb{R}^d$ with some fixed $d \in \mathbb{N}$.

Our model is defined as follows. Let $W \subseteq \mathbb{R}^d$ be compact, the territory of our telecommunication system, containing the origin o of $\mathbb{R}^d$.

1.2.1 Users

Let $\mu \in \mathcal{M}(W)$ be an absolutely continuous measure on W with $\mu(W) > 0$. Note that we do not require that $\operatorname{supp}(\mu) = W$. For $\lambda > 0$, we denote by X^λ a Poisson point process in W with intensity measure $\lambda\mu$. They points $X_i \in X^\lambda$ are interpreted as the *locations of the users* in the system, while the origin o of $\mathbb{R}^d$ is the single *base station*. We assume that $X^\lambda = \{X_i \colon i \in I^\lambda\}$ with $I_\lambda = \{1, \dots, N(\lambda)\}$ and $(N(\lambda))_{\lambda > 0}$ a standard Poisson process on $\mathbb{N}_0$ and $(X_i)_{i \in \mathbb{N}}$ is an i.i.d. sequence of W-distributed random variables with distribution $\mu(\cdot)/\mu(W)$ defined on one probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Since μ has a density, all points X_i are mutually different with probability one. Furthermore, X^λ is increasing in λ, and its *empirical* measure, normalized by $1/\lambda$,

$$L_\lambda = \frac{1}{\lambda} \sum_{i \in I^\lambda} \delta_{X_i}, \qquad (1.1)$$

converges towards μ almost surely as $\lambda \to \infty$.

These assumptions on the users can be relaxed, see Section 1.6.7.

1.2.2 Message trajectories

We now introduce the collection of trajectories sent out from the users to o, i.e., for *uplink communication*. (The *downlink* scenario, that is, communication in the opposite direction, works very similarly and will be described in Section 1.6.3.) For any $i \in I^\lambda$, we call a vector of the form

$$S^i = (S^i_{-1} = K_i, S^i_0 = X_i, S^i_1 \in X^\lambda, \dots, S^i_{K_i - 1} \in X^\lambda, S^i_{K_i} = o) \in \mathbb{N} \times \Big(\bigcup_{k \in \mathbb{N}} W^k \Big) \times \{o\}, \qquad (1.2)$$

a *message trajectory* from X_i to o with K_i hops. That is, S^i starts from X_i and ends in o after K_i hops from user to user $\in X^\lambda$. Hence, the users receive the function of a relay. We fix a number $k_{\max} \in \mathbb{N}$ and write $\mathcal{S}^i_{k_{\max}}(X^\lambda)$ for the set of all possible realizations of the random variable S^i with $K_i \le k_{\max}$, i.e., with no more than $k_{\max}$ hops. Hence, elements $s^i = (s^i_{-1}, s^i_0, s^i_1, \dots, s^i_{s^i_{-1}-1}, s^i_{s^i_{-1}})$ of $\mathcal{S}^i_{k_{\max}}(X^\lambda)$ satisfy $s^i_{-1} \in \{1, \dots, k_{\max}\}$ and $s^i_0 = X_i$. We write $\mathcal{S}_{k_{\max}}(X^\lambda) = \prod_{i \in I^\lambda} \mathcal{S}^i_{k_{\max}}(X^\lambda)$ for the set of all possible realizations of the families $S^\lambda = (S^i)_{i \in I^\lambda}$. We use the notation $[k_{\max}] = \{1, \dots, k_{\max}\}$.

Given $i \in I^\lambda$, we consider each trajectory S^i in (1.2) as an $S^i_{k_{\max}}(X^\lambda)$-valued random variable. Its *a priori* measure is given by the formula

$$s^i \mapsto \frac{1}{N(\lambda)^{s^i_{-1}-1}}, \qquad s^i \in \mathcal{S}^i_{k_{\max}}(X^\lambda). \tag{1.3}$$

That is, its restriction to $\{s^i \in \mathcal{S}^i_{k_{\max}}(X^\lambda) \colon s^i_{-1} = k\}$ is the uniform distribution for any $k \in [k_{\max}]$, and its total mass is equal to $k_{\max}$. Recall that it fixes the starting point X_i and the terminal point o.

Under our joint reference measure, all the trajectories are independent; indeed it gives the value

$$s = (s^i)_{i \in I^\lambda} \mapsto \prod_{i \in I^\lambda} \frac{1}{N(\lambda)^{s^i_{-1}-1}} \tag{1.4}$$

to the configuration $s \in \mathcal{S}_{k_{\max}}(X^\lambda)$. Thus, it gives a total mass of $k_{\max}^{N(\lambda)}$ to $\mathcal{S}_{k_{\max}}(X^\lambda)$.

1.2.3 Interference

Now we introduce interference. Let us choose a *path-loss function*, which describes the propagation of signal strength over distance. This is a monotone decreasing, continuous function $\ell \colon [0, \infty) \to (0, \infty)$. A typical example for such ℓ is the one corresponding to isotropic antennas with ideal Hertzian propagation, i.e. $\ell(r) = \min\{1, r^{-\alpha}\}$, for some $\alpha > 0$, see e.g. [GT08, Section II.]. The *signal-to-interference ratio (SIR)* of a transmission from $X_i \in I^\lambda$ to $x \in W$ in the presence of the users in X^λ is given as

$$\mathrm{SIR}(X_i, x, X^\lambda) = \frac{\ell(|X_i - x|)}{\frac{1}{\lambda}\sum_{j \in I^\lambda} \ell(|X_j - x|)}. \tag{1.5}$$

We will call the denominator of the r.h.s of (1.5) the *interference*. See Section 1.6.2 for a discussion about the relevance for telecommunication.

More generally, if $\mu_0 \in \mathcal{M}(W)$, we define for any $x, y \in W$

$$\mathrm{SIR}(x, y, \mu_0) = \frac{\ell(|x - y|)}{\int_W \ell(|z - y|)\mu_0(\mathrm{d}z)}, \tag{1.6}$$

where we call the denominator interference w.r.t. μ_0. Then, in a slight abuse of notation, we have $\mathrm{SIR}(X_i, x, X^\lambda) = \mathrm{SIR}(X_i, x, L_\lambda)$, where we recall the empirical measure L_λ from (1.1).

Now, given a trajectory configuration $s = (s^i)_{i \in I^\lambda} \in \mathcal{S}_{k_{\max}}(X^\lambda)$, we put

$$\mathfrak{S}(s) = \sum_{i \in I^\lambda} \sum_{l=1}^{s^i_{-1}} \mathrm{SIR}(s^i_{l-1}, s^i_l, L_\lambda)^{-1}. \tag{1.7}$$

We provide an interpretation of this in Section 1.6.2.

1.2.4 Congestion

Now we introduce the congestion term. Given a trajectory configuration $s = (s^i)_{i\in I^\lambda} \in \mathcal{S}_{k_{\max}}(X^\lambda)$, we define

$$m_i(s) = \sum_{j\in I^\lambda} \sum_{l=1}^{s^i_{-1}-1} \mathbb{1}\{s^j_l = s^i_0\}, \quad i \in I^\lambda, \tag{1.8}$$

as the number of incoming hops into the user (relay) $s^i_0 = X_i$ of any of the trajectories. Then we take

$$\mathfrak{M}(s) = \sum_{i\in I^\lambda} m_i(s)(m_i(s)-1) \tag{1.9}$$

as the total congestion term that is caused by the trajectory configuration s. Note that $\frac{1}{2}\mathfrak{M}(s)$ is equal to the number of pairs of hops that jump to the same relay.

1.2.5 Gibbsian trajectory distribution

Now we define the central object of this study: a Gibbs distribution on the set of collections of trajectories as follows. For any $s = (s^i)_{i\in I^\lambda} \in \mathcal{S}_{k_{\max}}(X^\lambda)$ put

$$\mathrm{P}^{\gamma,\beta}_{\lambda,X^\lambda}(s) := \frac{1}{Z^{\gamma,\beta}_\lambda(X^\lambda)} \Big(\prod_{i\in I^\lambda} \frac{1}{N(\lambda)^{s^i_{-1}-1}} \Big) \exp\Big\{ -\gamma\mathfrak{S}(s) - \beta\mathfrak{M}(s) \Big\}, \tag{1.10}$$

where $\gamma > 0$ and $\beta > 0$ are parameters. This is the Gibbs distribution with independent reference measure given in (1.4), subject to two exponential weights with the SIR term in (1.7) and the congestion term in (1.9). Here

$$Z^{\gamma,\beta}_\lambda(X^\lambda) = \sum_{r\in\mathcal{S}_{k_{\max}}(X^\lambda)} \Big(\prod_{i\in I^\lambda} \frac{1}{N(\lambda)^{r^i_{-1}-1}} \Big) \exp\Big\{ -\gamma\mathfrak{S}(r) - \beta\mathfrak{M}(r) \Big\} \tag{1.11}$$

is the normalizing constant, which we will refer to as *partition function*. Note that $\mathrm{P}^{\gamma,\beta}_{\lambda,X^\lambda}(\cdot)$ is random conditional on X^λ, and it is a probability measure on $\mathcal{S}_{k_{\max}}(X^\lambda)$. In the jargon of statistical mechanics, it is a *quenched* measure, which we will consider almost surely with respect to the process $(X^\lambda)_{\lambda>0}$. In the *annealed* setting, one would average out over $(X^\lambda)_{\lambda>0}$, see Section 1.6.8.

1.3 The limiting free energy

The main goal of this paper is the description of this model in the limit $\lambda \to \infty$ in the quenched setting. Our first result describes the limiting *free energy*, i.e., the exponential behaviour of the partition function. In order to state this result, we introduce the following notation. For $k \in \mathbb{N}$, elements of the product space $W^k = W^{\{0,1,\dots,k-1\}}$ will be denoted as $(x_0,\dots,x_{k-1})$. For $l = 0,\dots,k-1$, the l-th marginal of a measure $\nu_k \in \mathcal{M}(W^k)$ is denoted by $\pi_l\nu_k \in \mathcal{M}(W)$, i.e., $\pi_l\nu_k(A) = \nu_k(W^{\{0,\dots,l-1\}} \times A \times W^{\{l+1,\dots,k-1\}})$ for any Borel set A of W.

Now we introduce the objects in terms of which we will be able to describe the asymptotics of the entire telecommunication system.

Definition 1.1. *An* admissible trajectory setting *is a collection of measures* $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$ *with* $\nu_k \in \mathcal{M}(W^k)$ *for all* k *and* $\mu_m \in \mathcal{M}(W)$ *for all* m*, satisfying the following properties.*

$$(i) \quad \sum_{k=1}^{k_{\max}} \pi_0 \nu_k = \mu, \qquad (ii) \quad \sum_{m=0}^{\infty} \mu_m = \mu, \qquad (iii) \quad M := \sum_{m=0}^{\infty} m\mu_m = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k. \tag{1.12}$$

The measure ν_k is the measure of the k-step trajectories and μ_m the measure of the users that receive precisely m incoming hops; note that there is no reason that they be normalized (like for μ). Observe that both the length k of the trajectories and the number m of times that a user is used as a relay are random in our model. Condition (i) expresses our assumption that each user transmits precisely one message, (ii) says that each user serves as a relay for precisely m message trajectories for some $m \in \mathbb{N}_0$, and (iii) says that the relays can be calculated in two ways: according to the number of incoming hops and according to the index of the hop of a trajectory that uses it. See Section 1.6 for more explanations and interpretations, moreover for a modified version of our model where the assumption (i) is relaxed. By

$$\mathcal{H}_V(\nu \mid \widetilde{\nu}) = \begin{cases} \int_V \mathrm{d}\nu \, \log \frac{\mathrm{d}\nu}{\mathrm{d}\widetilde{\nu}} - \nu(V) + \widetilde{\nu}(V), & \text{if the density } \frac{\mathrm{d}\nu}{\mathrm{d}\widetilde{\nu}} \text{ exists,} \\ +\infty & \text{otherwise,} \end{cases} \tag{1.13}$$

we denote the relative entropy [GZ93, Section 2.3] of a Borel measure ν with respect to another Borel measure $\widetilde{\nu}$ on a measurable set V.

For an admissible trajectory setting $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$ we define

$$\mathrm{S}(\Psi) = \sum_{k=1}^{k_{\max}} \int_{W^k} \mathrm{d}\nu_k \, f_k, \qquad \text{where } f_k(x_0, \dots, x_{k-1}) = \sum_{l=1}^{k} \frac{\int_W \mu(\mathrm{d}y)\ell(|y - x_l|)}{\ell(|x_{l-1} - x_l|)}, \qquad x_k = o, \tag{1.14}$$

$$\mathrm{M}(\Psi) = \sum_{m=0}^{\infty} m(m-1)\, \mu_m(W) \tag{1.15}$$

and

$$\mathrm{I}(\Psi) = \sum_{k=1}^{k_{\max}} \mathcal{H}_{W^k}\big(\nu_k \mid \mu \otimes M^{\otimes(k-1)}\big) + \sum_{m=0}^{\infty} \mathcal{H}_W(\mu_m \mid \mu c_m) + \mu(W)\Big(2 - \sum_{k=1}^{k_{\max}} M(W)^{k-1}\Big) - 1 - \frac{1}{\mathrm{e}}, \tag{1.16}$$

where we recall $M = \sum_{m \in \mathbb{N}_0} m\mu_m$ from Definition 1.1(iii), and $c_m = \exp(-1/(\mathrm{e}\mu(W))(\mathrm{e}\mu(W))^{-m}/m!$ are the weights of the Poisson distribution with parameter $1/(\mathrm{e}\mu(W))$. Note that according to (i) and (iii) in (1.12), we have $M(W) \le (k_{\max} - 1)\mu(W)$. From the representation in (1.28) below, one easily sees that $\mathrm{I}(\Psi)$ is well-defined as an element of $(-\infty, \infty]$ and $\Psi \mapsto \mathrm{I}(\Psi)$ is a lower semicontinuous function that is bounded from below. A tedious but elementary calculation shows that I is convex. In Section 1.5, I will turn out to govern the large deviations of the trajectory configuration.

We fix all the parameters $W, \mu, \ell, k_{\max}, \gamma$ and β of the model. Our first main result is the following.

Theorem 1.2 (Quenched exponential rate of the partition function)**.** *For* $\mathbb{P}$*-almost all* $\omega \in \Omega$*,*

$$\lim_{\lambda \to \infty} \frac{1}{\lambda} \log Z_\lambda^{\gamma,\beta}(X^\lambda(\omega)) = - \inf_{\Psi \text{ admissible trajectory setting}} \big(\mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)\big). \tag{1.17}$$

See Section 1.6 for a discussion and Section 3.4 for the proof. An analogous result holds for downlink communication, see Section 1.6.3.

1.4 Description of the minimizers

From the variational formula in (1.17), descriptive information about the typical behaviour of the telecommunication system can be deduced, see Sections 1.5 and 1.6. Hence, it is important to derive the Euler-Lagrange equations and to characterize the minimizers in most explicit terms. Our main results in this respect are the following. Note that the case $k_{\max}=1$ is trivial.

Proposition 1.3 (Characterization of the minimizer(s)). *Let $k_{\max}>1$. The infimum in the variational formula in* (1.17) *is attained, and every minimizer $\Psi=((\nu_k)_{k=1}^{k_{\max}},(\mu_m)_{m=0}^{\infty})$ has the following form.*

$$\nu_k(\mathrm{d}x_0,\dots,\mathrm{d}x_{k-1}) = \mu(\mathrm{d}x_0)\,A(x_0)\prod_{l=1}^{k-1}\big(C(x_l)M(\mathrm{d}x_l)\big)\mathrm{e}^{-\gamma f_k(x_0,\dots,x_{k-1})}, \qquad k\in[k_{\max}], \tag{1.18}$$

$$\mu_m(\mathrm{d}x) = \mu(\mathrm{d}x)\,B(x)\frac{C(x)^m}{m!}\mathrm{e}^{-\beta m(m-1)}, \qquad m\in\mathbb{N}_0, \tag{1.19}$$

where $A,B,C\colon W\to[0,\infty)$ are functions such that the conditions in (1.12) *are satisfied.*

The proof of Proposition 1.3 is in Section 4.

While explicit formulas for the functions A and B can, given the function C, easily be derived from (i) and (ii) in (1.12) (see (4.10)), the condition for C coming from (iii) is deeply involved and cannot be easily solved intrinsically; see (4.12) – (4.14). We have no argument for its existence to offer other than via proving the existence of a minimizer Ψ and deriving the Euler-Lagrange equations. By convexity of I, S and M, every solution Ψ to these equations is a minimizer.

In case $k_{\max}=1$, the only admissible trajectory setting is $\Psi=(\nu_1,(\mu_m)_{m\in\mathbb{N}_0})$ with $\mu_0=\nu_1=\mu$ and $\mu_m=0$ otherwise, therefore this Ψ minimizes (1.17). Thus, the limiting free energy is strictly negative, it has value $-\gamma\int_W\mu(\mathrm{d}z)\frac{\int_W\mu(\mathrm{d}y)\ell(|y-o|)}{\ell(|z-o|)}$.

1.5 Large deviations for the empirical trajectory measure

Actually, the minimizers of the variational formula in (1.17) receive a rigorous interpretation in terms of important objects that describe the telecommunication system. Indeed, for fixed $k\in[k_{\max}]$ and for a collection of trajectories $s\in\mathcal{S}_{k_{\max}}(X^\lambda)$, we define

$$R_{\lambda,k}(s)=\frac{1}{\lambda}\sum_{i\in I^\lambda\colon s_{-1}^i=k}\delta_{(s_0^i,\dots,s_{k-1}^i)}, \tag{1.20}$$

the empirical measures of all the k-hop trajectories, which is an element of $\mathcal{M}(W^k)$. The second crucial empirical measure is the one of the users whose number of incoming messages is equal to a fixed number $m\in\mathbb{N}_0$:

$$P_{\lambda,m}(s)=\frac{1}{\lambda}\sum_{i\in I^\lambda\colon m_i(s)=m}\delta_{s_0^i}. \tag{1.21}$$

This is an element of $\mathcal{M}(W)$. Then

$$\Psi_\lambda(s)=\big((R_{\lambda,k}(s))_{k\in[k_{\max}]},(P_{\lambda,m}(s))_{m\in\mathbb{N}_0}\big) \tag{1.22}$$

satisfies the definition of an admissible trajectory setting, apart from the fact that instead of (i), $\sum_{k=1}^{k_{\max}} \pi_0 R_{\lambda,k}(s) = L_\lambda$ holds, and instead of (ii), $\sum_{m=0}^{\infty} P_{\lambda,m}(s) = L_\lambda$, where we recall that L_λ converges to μ almost surely as $\lambda \to \infty$. According to our remarks after Definition 1.1, $R_{\lambda,k}(s)$ and $P_{\lambda,m}(s)$ play the roles of ν_k and μ_m, respectively, in an admissible trajectory setting, which explains this term. Furthermore, for $s \in \mathcal{S}_{k_{\max}}(X^\lambda)$, we can express the congestion term as

$$\mathfrak{M}(s) = \lambda \mathrm{M}(\Psi_\lambda(s)).$$

Moreover, for the interference term we have

$$\mathfrak{S}(s) \approx \lambda \mathrm{S}(\Psi_\lambda(s)), \tag{1.23}$$

where we typically do not have an equality, because the interference terms in $\mathfrak{S}$ are taken w.r.t. L_λ, while the ones in S are taken w.r.t. μ. However, since L_λ tends to μ almost surely, this difference vanishes in the limit, see Proposition 3.2.

We consider now the distribution of $\Psi_\lambda(S)$ with S distributed under the product reference measure introduced in (1.4), normalized to a probability measure, $\mathrm{P}^{0,0}_{\lambda,X^\lambda}$; note that the normalization $Z^{0,0}_\lambda(X^\lambda)$ is equal to $k_{\max}^{N(\lambda)}$. Our next main result is a large-deviation principle (LDP; see (1.25)–(1.26)) and the convergence towards the minimizers of the variational formula.

Theorem 1.4 (LDP and convergence for the empirical measures)**.** *The following statements hold almost surely with respect to $(X^\lambda)_{\lambda>0}$.*

(i) The distribution of $\Psi_\lambda(S)$ under $\mathrm{P}^{0,0}_{\lambda,X^\lambda}$ satisfies an LDP as $\lambda \to \infty$ with scale λ on the set

$$\mathcal{A} = \Big(\prod_{k=1}^{k_{\max}} \mathcal{M}(W^k) \Big) \times \mathcal{M}(W)^{\mathbb{N}_0} \tag{1.24}$$

with rate function given by $\mathcal{A} \ni \Psi \mapsto \mathrm{I}(\Psi) + \mu(W) \log k_{\max}$, which we define as ∞ if Ψ is not an admissible trajectory setting.

(ii) For any $\gamma, \beta \in (0,\infty)$, the distribution of $\Psi_\lambda(S)$ under $\mathrm{P}^{\gamma,\beta}_{\lambda,X^\lambda}$ converges towards the set of minimizers of the variational formula in (1.17)*.*

For the reader's convenience, we recall that the LDP states that the rate function $\mathrm{I} + \mu(W) \log k_{\max}$ is lower semicontinuous and

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in F) \le -\inf_F \big(\mathrm{I} + \mu(W) \log k_{\max}\big), \tag{1.25}$$

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda} \log \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in G) \ge -\inf_G \big(\mathrm{I} + \mu(W) \log k_{\max}\big), \tag{1.26}$$

for any closed set F and any open set G in $\mathcal{A}$. See [DZ98] for more on large deviation theory. On $\mathcal{A}$, we consider the product topology that is induced by weak convergence in each factor; this is equal to coordinatewise weak convergence, see Section 3.3 for more details. Convergence of a distribution towards a set is defined by requiring that for any neighbourhood of the set, the probability of not being in the neighbourhood vanishes.

The proof of Theorem 1.4(i) is in Section 3.5. Assertion (ii) is a simple consequence of (i), since the functionals S and M are bounded and continuous on the set $B_C = \{\Psi \in \mathcal{A} \colon \mathrm{M}(\Psi) \leq C\}$ for any C, and B_C is compact in $\mathcal{A}$ (see Lemma 4.1). Denoting the level sets of the rate function $\mathrm{I} + \mu \log k_{\max}$ by $\Phi_\alpha = \{\Psi \in \mathcal{A} \colon \mathrm{I}(\Psi) + \mu(W) \log k_{\max} \leq \alpha\}$ for $\alpha \in \mathbb{R}$, we see that $\Phi_\alpha \cap B_C$ is compact for any α and C. Thus, Varadhan's lemma can be applied to prove the assertion (ii).

1.6 Discussion

1.6.1 Mathematical essence

Going away from applications in telecommunication and formulating in more abstract terms, this work is about a large-deviation description of a disintegration of the local times of a highly dense family of random trajectories in $\mathbb{R}^d$ according to their number of hits in given sites. More precisely, we register the total number of steps into a given site x coming from all the random trajectories, seen as a measure in x and disintegrated according to this total step number. The reason why one has to introduce the measures μ_m, $m \in \mathbb{N}_0$, is that the number of users receiving a given number of incoming messages cannot be expressed in terms of the trajectory measures in a way that is continuous in the weak topology when taking the limit $\lambda \to \infty$. Indeed, it is possible to write, for each fixed $\lambda > 0$, the empirical measure $P_{\lambda,m}$ as a functional of the collection of the empirical trajectory measures $R_{\lambda,k}$, $k \in [k_{\max}]$, but this functional is highly discontinuous. In the high-density limit, sites standing close to each other are identified with each other in the weak topology, and their distinctness is washed out. On the other hand, after the introduction of the measures $P_{\lambda,m}$, the congestion term is a lower semicontinuous function of them and can be handled in terms of an LDP.

We demonstrate the practical value of our analysis by an application to certain relevant functionals of both the trajectory family and the local time family. It is clear that our results persist to many other choices of these functionals; essentially to all (lower semi-)continuous and bounded ones. Our approach will be fruitful for many other investigations of such mathematical models also in quite different applications.

1.6.2 The SIR term

In a mathematical description of a telecommunication system, one typically requires that the signal-to-interference ratio be larger than a given threshold, in order that the signal can be successfully transmitted. However, our model is designed in the spirit of a common wealth approach, where we do not want to consider any single message, but the total quality of transmission in the entire system. This quantity is the sum of all the reciprocal values of the SIRs of all the (hops of the) messages. It is exponentially weighted with a negative factor, which "softly" keeps all the SIRs at positive values on an average.

The choice of the *reciprocals* of the SIRs comes from the fact that the *bandwidth* used for a transmission is defined [SPW07] as

$$\frac{R}{\log_2(1 + \mathrm{SIR}(\cdot))},$$

where R is the data transmission rate, and SIR is defined as in (1.5) without the factor of $\frac{1}{\lambda}$. In the high-density setting $\lambda \to \infty$ that we study, this quantity can be approached well by (a constant times) the reciprocals of the SIR. [SPW07, Section 3] suggests that in case of multi-hop communication, the used bandwidth equals the sum of the used bandwidth values corresponding to the individual hops, which explains our choice of the sum over l in (1.7).

Note that the conventional definition of interference of a transmission from X_i to x is $\sum_{j\in I^\lambda\setminus\{i\}} \ell(|X_j - x|)$, in contrast to our definition in (1.5), where we added a factor of $\frac{1}{\lambda}$. According to this convention, we should say "total received powerïnstead of "interference", cf. [KB14, Section II.]. As we are interested in the limit $\lambda \to \infty$, where it makes no difference whether or not we add $\frac{1}{\lambda}\ell(|X_i - x|)$ to the denominator, we will stick to our expressions "SIRänd "interference". However, note also our additional factor of $1/\lambda$, which we think is appropriate, at least mathematically, to our setting, in which we consider the high-density limit $\lambda \to \infty$. We actually weight the "usual" SIR term by the density parameter. The interpretation of the appearance of the factor of $1/\lambda$ is that, in order to cope with an enormous number of messages in a system with one base station and a fixed bandwidth, one can either distribute the messages over a longer time stretch or decompose the messages into many smaller ones, and the factor of $1/\lambda$ is a crude approximation of a combination of these two strategies.

1.6.3 Downlink communication

In the downlink scenario, instead of users sending messages to the base station, the base station sends exactly one message to each of the users, using the same relaying rules. One can define a Gibbsian model analogously to the one defined in Section 1.2, now for trajectories from o to X_i instead of the other way around. The SIR term and the congestion term have to be redefined in an obvious way. We are certain that analogues of all our results are true and can be proved in the same way, hence we abstained from spelling them out.

1.6.4 Sending no or multiple messages

All our results can be extended to the possibility that users send no message or multiple messages. This models the standard situation in which large messages are cut into many smaller ones, who independently find their ways through the system.

For this, we have to enlarge the trajectory probability space: to each user $X_i \in X^\lambda$, we attach the number $P_i \in \mathbb{N}_0$ of transmitted messages, and for each $j \in \{1,\dots,P_i\}$, there is an independent trajectory $X_i \to o$. The empirical trajectory measure $R_{\lambda,k}$ must be augmented by these trajectories. The main additional assumption then is that $\sum_{k=1}^{k_{\max}} \pi_0 R_{\lambda,k}$ converges to some measure $\mu_0 \in \mathcal{M}(W)$ with $0 \neq \mu_0 \ll \mu$. Then the Definition 1.1 of an admissible trajectory setting $\Psi = ((\nu_k)_k, (\mu_m)_m)$ changes so that now $\sum_{k=1}^{k_{\max}} \pi_0 \nu_k = \mu_0$ is required instead of (i) of (1.12). ((ii) and (iii) remain unchanged, since they refer only to the number of relaying hops.) Furthermore, in the definition (1.16) of the rate function I, in each summand of the first of the three terms, μ must be replaced by μ_0, while the two others remain unchanged.

The SIR term also has to be changed. The number P_i can be interpreted as a *signal power* of the user X_i. Thus, according to [BB09, Sections 2.3.1, 5.1], the SIR of his transmission of a message to $x \in W$ should be defined as follows

$$\mathrm{SIR}((X_i,P_i),\ x,\ (X_j,P_j)_{j\in I^\lambda}) = \frac{\ell(|X_i - x|)P_i}{\frac{1}{\lambda}\sum_{j\in I^\lambda} \ell(|X_j - x|)P_j}.$$

One could also incorporate (possibly random) sizes of the messages, which would require an additional enlargement of the trajectory space.

1.6.5 Interpretation of the variational formula

The interpretation of an admissible trajectory setting $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$ is given after Definition 1.1; they play the role of the empirical measures introduced in Section 1.5. For each k, the term $f_k(x_0, \dots, x_{k-1})$ describes the SIR-term of the k-step trajectory $(x_0, \dots, x_{k-1})$, and $\int f_k \, \mathrm{d}\nu_k$ is the average SIR-term of the admissible trajectory setting. For each m, the term $\frac{1}{2}m(m-1)\mu_m(W)$ is the linear rate (in λ) of the number of pairs of incoming messages experienced at users who receive precisely m incoming messages. The entropic term I in (1.17) describes the entropy of the choices of the indices i of the users X_i and the indices $j = 1, \dots, K_i - 1$ of the relays S_j^i of the trajectories S^i; it can be understood as the exponential rate of counting complexity.

For a measurable set V and for $\nu, \widetilde{\nu} \in \mathcal{M}(V)$, let us write

$$H_V(\nu \mid \widetilde{\nu}) = \int_V \mathrm{d}\nu \log \frac{\mathrm{d}\nu}{\mathrm{d}\widetilde{\nu}}, \quad \text{if } \nu \ll \widetilde{\nu} \text{ and } \infty \text{ otherwise.} \tag{1.27}$$

Note that $H_V(\nu \mid \widetilde{\nu}) = \mathcal{H}_V(\nu \mid \widetilde{\nu})$ if $\nu(V) = \widetilde{\nu}(V)$. Thus, we have

$$\begin{aligned} \mathrm{I}(\Psi) = {} & \mu(W) H_{[k_{\max}]}\Big(\Big(\frac{\nu_k(W^k)}{\mu(W)}\Big)_{k\in[k_{\max}]} \Big| \, \mathfrak{c}\Big) + \mu(W)\mathcal{H}_{\mathbb{N}_0}\Big(\Big(\frac{\mu_m(W)}{\mu(W)}\Big)_{m\in\mathbb{N}_0} \Big| \, \mathrm{Po}_{1/(\mathrm{e}\mu(W))}\Big) \\ & - M(W)\log\frac{M(W)}{\mu(W)} - \frac{1}{\mathrm{e}} \\ & + \sum_{k\in[k_{\max}]} \nu_k(W^k)\mathcal{H}_{W^k}\big(\overline{\nu_k} \big| \overline{\mu} \otimes \overline{M}^{\otimes(k-1)}\big) + \sum_{m\in\mathbb{N}_0} \mu_m(W)\mathcal{H}_W\big(\overline{\mu_m} \big| \overline{\mu}\big). \end{aligned} \tag{1.28}$$

where we wrote $\overline{N} = N/N(V)$ for the normalized version of a measure N on a set V, Po_α for the Poisson distribution on $\mathbb{N}_0$ with parameter α and $\mathfrak{c}$ for the counting measure on $[k_{\max}]$. The terms on the r.h.s. in the first line are entropies for the trajectory length and the number of incoming messages per relay with respect to natural reference measures. The terms in the last line are entropies for the distribution of the trajectories and of the locations of the relays that receive a given number of incoming messages. From (1.28) it is easy to see that I is bounded from below, using Jensen's inequality and the finiteness of the counting measure on $[k_{\max}]$. (From the LDP in Theorem 1.4(i), one obtains that $\inf \mathrm{I} = -\mu(W)\log k_{\max}$.)

1.6.6 Interpretation of the minimizer(s)

Proposition 1.3 tells us quite some information about the limiting trajectory distribution and the limiting spatial distribution of users with a given number of incoming messages under the measure $\mathrm{P}^{\gamma,\beta}_{\lambda,X^\lambda}$. Indeed, both have densities that are $\mu^{\otimes k}$-almost everywhere positive. It is remarkable that the k-step trajectories follow a distribution that comes from choosing independently all the k sites with measures that do not depend on k (the starting point according to $A(x)\,\mu(\mathrm{d}x)$ and all the other $k-1$ sites each according to another measure), weighted with the SIR-term. Furthermore, all the measures of the users receiving m messages superpose each other on the full set $\mathrm{supp}(\mu)$, and at each space point x, this number m is distributed according to some Poisson distribution, weighted with the congestion term $\mathrm{e}^{-\beta m(m-1)}$.

DOI 10.20347/WIAS.PREPRINT.2392

1.6.7 Non-Poissonian users

In fact, the main results of this paper hold for any collection of (random or non-random) point processes $((X_i)_{i=1,\dots,N(\lambda)})_{\lambda>0}$ on W for which $L_\lambda = \frac{1}{\lambda}\sum_{i=1}^{N(\lambda)} \delta_{X_i}$ converges weakly (almost surely, if random) to μ as $\lambda \to \infty$. Neither the independence or monotonicity in λ, nor the Poissonity of $(N(\lambda))_{\lambda>0}$ is used for the proofs. For example, our results remain also true for the deterministic set $X^\lambda = W \cap (\frac{1}{\lambda}\mathbb{Z}^d)$ and μ the Lebesgue measure on W.

1.6.8 The annealed setting

Of mathematical interest might also be the annealed setting, where we average also over the locations of the users. In order to get an interesting result, we have to assume that L_λ satisfies a large deviation principle on the set $\mathcal{M}(W)$ with some good rate function J. (In the case of a Poisson point process with intensity measure $\lambda\mu$, J would be [HJP16, Proposition 3.6] the relative entropy with respect to μ, see (1.13).) Then the large-λ exponential rate of the annealed free energy should be equal to the negative infimum over $\mu_0 \in \mathcal{M}(W)$ of $J(\mu_0)$ plus the quenched rate function terms from the right-hand side of (1.17) with μ replaced by μ_0 everywhere. Also our other results on the LDP and the form of the minimizer(s) should have some analogue, which we do not spell out.

2 The distribution of the empirical measures

Having seen in Section 1.5 that the Gibbsian model can be entirely described in terms of the trajectory setting $\Psi_\lambda(s)$, i.e., of the crucial empirical measures $R_{\lambda,k}(s)$ and $P_{\lambda,m}(s)$ defined in (1.20)–(1.21), we now consider the question how to describe their distributions. We have to quantify the number of message trajectory families s that give the same family of empirical measures. The plain and short (but wrong) answer is

$$\sum_{s\in\mathcal{S}_{k_{\max}}(X^\lambda)\colon\, R_{\lambda,k}(s)=\nu_k\,\forall k,\; P_{\lambda,m}(s)=\mu_m\,\forall m} \quad \prod_{i\in I^\lambda} \frac{1}{N(\lambda)^{s^i_{-1}-1}} \approx \mathrm{e}^{-\lambda \mathrm{I}(\Psi)}, \tag{2.1}$$

where we recall $\mathrm{I}(\Psi)$ from (1.16) and recall that $\Psi = ((\nu_k)_{k\in[k_{\max}]},(\mu_m)_{m\in\mathbb{N}_0})$. From such an assertion, it is indeed not far to conclude Theorem 1.2, but the problem is that this statement is not true like this. Actually, there are very many Ψ's such that the left-hand side is equal to zero, for example if any of the ν_k's or μ_m's has values outside $\frac{1}{\lambda}\mathbb{N}_0$. However, if we do not consider single Ψ's, but open sets of Ψ's, then the idea behind (2.1) is sustainable. Therefore, we proceed in a standard way by decomposing the area W into finitely many subsets and count the message trajectories only according to the discretization sets that they visit. In Section 2.1 we introduce necessary notation for carrying out this strategy, and in Section 2.2 we derive explicit formulas for the distribution of the empirical measures in this discretization.

For the purpose of the present paper, where we consider the high-density limit $\lambda \to \infty$, we later need to take this limit and afterwards the limit as the fineness parameter δ of the decomposition of W goes to zero. The outcome of these parts of the procedure is formulated in Proposition 3.1. In Proposition 3.2 the consequences for the interference term and for the congestion term are formulated.

2.1 Our discretization procedure

Let us now head towards the formulation of the discretization procedure. We proceed by triadic spatial discretization of the Poisson point process $(X^\lambda)_{\lambda>0}$, similarly to the approach of [HJKP15]. To be more precise, we perform the following discretization argument. Note that we may assume that our telecommunication territory W is taken as $W=[-r,r]^d$, by accordingly extending μ trivially. We write $\mathbb{B}=\{3^{-n}|n\in\mathbb{N}_0\}$. For $\delta\in\mathbb{B}$, we define the set

$$W_\delta=\{[x-r\delta,x+r\delta]^d\colon x\in(2r\delta\mathbb{Z})^d\cap W\}$$

of congruent sub-cubes of W of side length $2r\delta$ and centers in $(2r\delta\mathbb{Z})^d$. Note that W_δ is a finite set, o is a center of an element of W_δ and any intersection of two distinct elements of W_δ has zero Lebesgue measure. Elements of W_δ will be called δ*-subcubes*. We will assume that for all $\delta\in\mathbb{B}$, the δ-subcubes are canonically numbered as $W_1^\delta,\dots,W_{\delta^{-d}}^\delta$, which can be done e.g. according to the increasing lexicographic order of the midpoints of the subcubes. For $j=1,\dots,\delta^{-d}$, let $C(W_j^\delta)$ denote the centre of the δ-subcube W_j^δ. Now, for Lebesgue-almost every $x\in W$, for all $\delta\in\mathbb{B}$ there exists a unique W_j^δ that contains x; let us denote this W_j^δ by W_δ^x, and the set of all $x\in W$ for which W_δ^x is well-defined by $W_\mathbb{B}$. For such x, the δ-discretization operator is defined as $\varrho_\delta\colon x\mapsto C(W_\delta^x)$. We will often use the simplified notation $x^\delta=\varrho_\delta(x)$.

Now, if $\nu\in\mathcal{M}(W)$, then for any $\delta\in\mathbb{B}$, $\nu^\delta=\nu\circ\varrho_\delta^{-1}$ is an element of $\mathcal{M}(W_\delta)$ with the property $\nu^\delta(W_j^\delta)=\nu(W_j^\delta)$, $\forall j=1,\dots,\delta^{-d}$. Note that $\mathcal{M}(W_\delta)=[0,\infty)^{W_\delta}$, which can be embedded in $\mathbb{R}^{W_\delta}$. Thus, weak convergence in $\mathcal{M}(W_\delta)$ is equivalent to norm convergence. On the other hand, if $\nu\in\mathcal{M}(W_\delta)$ for some $\delta\in\mathbb{B}$, then ν defines an atomic measure on W that has the same weights on each W_j^δ as ν and no mass anywhere else. Throughout the rest of this paper, we will denote this measure on W the same way as ν, for simplicity. We proceed analogously for W^k, $k\in[k_{\max}]$ instead of W.

Now we are able to define what a *standard setting* is, the interpretation of which will be given right after the definition. For any set X, let $\mathcal{P}(X)$ denote the power set of X.

Definition 2.1. *A* standard setting *is a collection of measures*

$$\begin{aligned}\underline{\Psi}=\Big(&(\nu_k)_{k=1}^{k_{\max}},((\nu_k^\delta)_{k=1}^{k_{\max}})_{\delta\in\mathbb{B}},((\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}})_{\delta\in\mathbb{B},\lambda>0},\\&(\mu_m)_{m=0}^\infty,((\mu_m^\delta)_{m=0}^\infty)_{\delta\in\mathbb{B}},((\mu_m^{\delta,\lambda})_{m=0}^\infty)_{\delta\in\mathbb{B},\lambda>0},(\mu^{\delta,\lambda})_{\delta\in\mathbb{B},\lambda>0}\Big)\end{aligned}\tag{2.2}$$

with the following properties: For any $\delta,\delta'\in\mathbb{B}$, $\lambda>0$, $k\in[k_{\max}]$, $m\in\mathbb{N}_0$ and $s,s_0,\dots,s_{k-1}=1,\dots,\delta^{-d}$, respectively,

1. $\mu^{\delta,\lambda}\in\mathcal{M}(W)$*, with the property that the event* $\{L_\lambda^\delta=\mu^{\delta,\lambda}\}$ *has positive probability,*
2. $\delta'\le\delta\Longrightarrow\mu^{\delta',\lambda}|_{\mathcal{P}(W_\delta)}=\mu^{\delta,\lambda}$,
3. $\mu^{\delta,\lambda}\overset{\lambda\to\infty}{\Longrightarrow}\mu^\delta$,
4. $\mu^\delta=\mu\circ\varrho_\delta^{-1}$*. In particular,* $\mu^\delta\overset{\delta\downarrow0}{\Longrightarrow}\mu$,
5. $\nu_k^{\delta,\lambda}\in\mathcal{M}(W^k)$*. Further, we have* $\sum_{k=1}^{k_{\max}}\pi_0\nu_k^{\delta,\lambda}=\mu^{\delta,\lambda}$*, moreover* $\lambda\nu_k^{\delta,\lambda}(W_{s_0}^\delta\times\ldots\times W_{s_{k-1}}^\delta)\in\mathbb{N}_0$.
6. $\delta'\le\delta\Longrightarrow\nu_k^{\delta',\lambda}|_{\mathcal{P}(W_\delta^k)}=\nu_k^{\delta,\lambda}$

7 $\nu_k^{\delta,\lambda} \overset{\lambda\to\infty}{\Longrightarrow} \nu_k^\delta$,

8 $\nu_k^\delta = \nu_k \circ (\varrho_\delta, \dots, \varrho_\delta)^{-1}$. *In particular,* $\nu_k^\delta \overset{\delta\downarrow 0}{\Longrightarrow} \nu_k$,

9 $\nu_m^{\delta,\lambda} \in \mathcal{M}(W)$ *with the property that* $\sum_{m=0}^\infty \mu_m^{\delta,\lambda} = \mu^{\delta,\lambda}$, *moreover* $\lambda\mu_m^{\delta,\lambda}(W_s^\delta) \in \mathbb{N}_0$.

10 $\sum_{m=0}^\infty m\mu_m^{\delta,\lambda} = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^{\delta,\lambda}$,

11 $\delta' \le \delta \Longrightarrow \mu_m^{\delta',\lambda}|_{\mathcal{P}(W_\delta)} = \mu_m^{\delta,\lambda}$,

12 $\mu_m^{\delta,\lambda} \overset{\lambda\to\infty}{\Longrightarrow} \mu_m^\delta$,

13 $\mu_m^\delta = \mu_m \circ \varrho_\delta^{-1}$. *In particular,* $\mu_m^\delta \overset{\delta\downarrow 0}{\Longrightarrow} \mu_m$.

Let us introduce also the empirical measure

$$P_\lambda(s) = \sum_{m\in\mathbb{N}_0} P_{\lambda,m}(s) = \frac{1}{\lambda}\sum_{i\in I^\lambda} \delta_{s_0^i}, \qquad s \in \mathcal{S}_{k_{\max}}(X^\lambda). \tag{2.3}$$

The interpretation of a standard setting Ψ is the following:

(i) For $\lambda > 0$ and $\delta \in \mathbb{B}$, $\mu^{\delta,\lambda}$ is the δ-discretized version $P_\lambda^\delta(s)$ of the empirical measure $P_\lambda(s)$ of any configuration $s \in \mathcal{S}_{k_{\max}}(X^\lambda)$; recall that this coincides with the empirical measure L_λ of the Poisson point process X^λ of users defined in (1.1) by means of our assumption that each user is picked precisely once in such a configuration. The consistency criterion (2) ensures that $\mu^{\delta,\lambda} = P_\lambda^\delta(s)$ for the same s. For any $\delta \in \mathbb{B}$, $\mu^{\delta,\lambda}$ converges to the δ-discretized version μ^δ of μ.

(ii) If $\mu^{\delta,\lambda}$ corresponds to the discretized version of the rescaled empirical measure of the transmitters, then $\nu_k^{\delta,\lambda}$ equals the δ-discretized version $R_{\lambda,k}^\delta(s)$ of the rescaled empirical measure $R_{\lambda,k}$ of the k-hop trajectories, related to L_λ via the constraint $\sum_{k=1}^{k_{\max}} \pi_0 \nu_k^{\delta,\lambda} = \mu^{\delta,\lambda}$ in (5), which means that each user sends out exactly one message. Again, we have a consistency relation (6), which ensures that for any $\lambda > 0$ and $k \in [k_{\max}]$, $\nu_k^{\delta,\lambda} = R_{\lambda,k}^\delta(s)$ for the same s for all $\delta \in \mathbb{B}$. For fixed $\delta \in \mathbb{B}$ and $k \in [k_{\max}]$, $\nu_k^{\delta,\lambda}$ converges to ν_k^δ, and the ν_k^δ's are the corresponding δ-discretized versions of a limiting (continuous) measure ν_k describing the asymptotic spatial distribution of k-hop trajectories.

(iii) Finally, for any $m \in \mathbb{N}_0$, $\lambda > 0$ and $\delta \in \mathbb{B}$, $\mu_m^{\delta,\lambda}$ equals the δ-discretized version $(L_\lambda^m)^\delta$ of the rescaled empirical measure

$$L_\lambda^m = \sum_{i\in I^\lambda\colon m_i = m} \delta_{X_i}$$

of the spatial locations of users receiving exactly m incoming messages. The constraint $\sum_{m=0}^\infty \mu_m^{\delta,\lambda} = \mu^{\delta,\lambda}$ in (9) means that each index $i \in I^\lambda$ belongs to exactly one of the sets $\{i \in I^\lambda\colon m_i(s) = m\}$, while the constraint $\sum_{m=0}^\infty m\mu_m^{\delta,\lambda} = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^{\delta,\lambda}$ means that the total number of relaying hops taken by all users equals the total number of incoming messages received by each relay, on any subset of W_δ. The consistency relation (11) ensures that for any $\lambda > 0$ and $n \in \mathbb{N}$, $\mu_m^{\delta,\lambda} = P_{\lambda,m}^\delta(s)$ for the same s for all $\delta \in \mathbb{B}$. For fixed $\delta \in \mathbb{B}$ and $m \in \mathbb{N}_0$, $\mu_m^{\delta,\lambda}$ converges to μ_m^δ, and the μ_m^δ's are the corresponding δ-discretized versions of a limiting (continuous) measure μ_m describing the asymptotic spatial distribution of m-hop trajectories.

Note that the condition (1) in Definition 2.1 in particular implies that for any $\lambda' > \lambda > 0$ and $\delta \in \mathbb{B}$ we have

$$\lambda' \mu^{\delta,\lambda'}(A) \geq \lambda \mu^{\delta,\lambda}(A), \qquad \forall A \subset W_\delta.$$

as a direct consequence of the fact that almost surely, $(X^\lambda)_{\lambda>0}$ is increasing.

Since in the definition of an admissible trajectory setting it is not required that $\mu_m(W) > 0$ holds only for finitely many m, we will often need the following notion of *controlled standard setting* in order to perform our large deviation analysis.

Definition 2.2. *A* controlled standard setting *is a standard setting $\underline{\Psi}$ as in* (2.2) *with the following extra property:*

$$\lim_{\lambda\to\infty} \sum_{m=0}^{\infty} m^2 \mu_m^{\delta,\lambda}(W_\delta) = \sum_{m=0}^{\infty} m^2 \mu_m^{\delta}(W_\delta) < \infty, \quad \textit{for all } \delta \in \mathbb{B}. \tag{2.4}$$

Note that by part (8) of Definition 2.1, we have $\sum_{k=1}^{k_{\max}} k\nu_k^\delta(W_\delta^k) = \sum_{k=1}^{k_{\max}} k\nu_k(W^k)$ for any standard setting. Using this, we have the following lemma.

Lemma 2.3. *Let $\underline{\Psi}$ be a controlled standard setting as in* (2.2). *Then $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^\infty)$ is an admissible trajectory setting.*

Proof. Part (5) of Definition 2.1 claims that for all $\delta \in \mathbb{B}$ and $\lambda > 0$ we have $\sum_{k=1}^{k_{\max}} \pi_0 \nu_k^{\delta,\lambda} = \mu^{\delta,\lambda}$. By parts (3) and (4) of Definition 2.1, we have $\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty} \nu_k^{\delta,\lambda} = \nu_k$ in the weak topology of $\mathcal{M}(W^k)$, for any fixed $k \in [k_{\max}]$. Similarly, by parts (7) and (8) of Definition 2.1, we have $\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty} \mu^{\delta,\lambda} = \mu$ in the weak topology of $\mathcal{M}(W)$. Moreover, since taking marginals is a continuous operation, also $\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty} \pi_0\nu_k^{\delta,\lambda} = \pi_0\nu_k$ for all k in the weak topology of $\mathcal{M}(W)$. Thus, we have (i) in (1.12) for $(\nu_k)_{k=1}^{k_{\max}}$. In order to see that (ii) holds for $(\mu_m)_{m=0}^\infty$, one can use part (9) of Definition 2.1, together with (2.4) and dominated convergence. Finally, by part (10) of Definition 2.1, (2.4) in Definition 2.2 and dominated convergence, we see that for any controlled setting $\underline{\Psi}$, we also have

$$\sum_{m=0}^{\infty} m\mu_m = \lim_{\delta\downarrow 0} \sum_{m=0}^{\infty} m\mu_m^\delta = \lim_{\delta\downarrow 0}\lim_{\lambda\to\infty} \sum_{m=0}^{\infty} m\mu_m^{\delta,\lambda} = \lim_{\delta\downarrow 0}\lim_{\lambda\to\infty} \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^{\delta,\lambda} = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k \tag{2.5}$$

in the weak topology of $\mathcal{M}(W)$. This implies (iii) in (1.12) for Ψ. Hence, Ψ is an admissible trajectory setting. $\square$

2.2 The distribution of the empirical measures

In this section, we describe the combinatorics of the system. For a standard setting $\underline{\Psi}$ as in Definition 2.1, let us introduce the configuration set

$$J^{\delta,\lambda}(\underline{\Psi}) = \Big\{ s \in \mathcal{S}_{k_{\max}}(X^\lambda) \,\Big|\, R^\delta_{\lambda,k}(s) = \nu_k^{\delta,\lambda}\ \forall k, \quad P^\delta_{\lambda,m}(s) = \mu_m^{\delta,\lambda}\ \forall m \Big\} \tag{2.6}$$

for fixed $\delta \in \mathbb{B}$ and $\lambda > 0$. In words, $J^{\delta,\lambda}(\underline{\Psi})$ is the set of families of trajectories such that the δ-coarsenings of the empirical measures of the trajectories and the hop numbers are given by the respective measures in the setting $\underline{\Psi}$. Note that $J^{\delta,\lambda}(\underline{\Psi})$ depends only on the δ-λ depending measures in the collection $\underline{\Psi}$.

DOI 10.20347/WIAS.PREPRINT.2392

In case $\mu^{\delta,\lambda}(W) > 0$, we will refer to the entity s_0^i, $i = 1, \ldots, \lambda\mu^{\delta,\lambda}(W_\delta)$ as the ith user or ith transmitter, the entity s^i, $i = 1, \ldots, \lambda\mu^{\delta,\lambda}(W_\delta)$ as the trajectory of the ith user, s_{-1}^i as the length (number of hops) of s^i, s_l^i as the l-th relay of s^i (for $l = 1, \ldots, s_{-1}^i - 1$), finally $m_i(s)$ as the number of incoming messages at the relay s_0^i.

The combinatorics of computing $\#J^{\delta,\lambda}(\underline{\Psi})$ is given as follows.

Lemma 2.4 (Cardinality of $J^{\delta,\lambda}(\underline{\Psi})$)**.** *For any $\delta, \lambda > 0$, and for any standard setting $\underline{\Psi}$,*

$$\#J^{\delta,\lambda}(\underline{\Psi}) = N^1_{\delta,\lambda}(\underline{\nu}) \times N^2_{\delta,\lambda}(\underline{\nu}) \times N^3_{\delta,\lambda}(\underline{\nu}), \tag{2.7}$$

where

$$N^1_{\delta,\lambda}(\underline{\Psi}) = \prod_{i=1}^{\delta^{-d}} \binom{\lambda\mu^{\delta,\lambda}(W_i^\delta)}{((\lambda\nu_k^{\delta,\lambda}(W_i^\delta \times W_{i_1}^\delta \times \ldots \times W_{i_{k-1}}^\delta))_{i_1,\ldots,i_{k-1}=1}^{\delta^{-d}})_{k=1}^{k_{\max}}}, \tag{2.8}$$

$$N^2_{\delta,\lambda}(\underline{\Psi}) = \prod_{i=1}^{\delta^{-d}} \binom{\lambda\mu^{\delta,\lambda}(W_i^\delta)}{(\lambda\mu_m^{\delta,\lambda}(W_i^\delta))_{m\in\mathbb{N}_0}}, \tag{2.9}$$

$$N^3_{\delta,\lambda}(\underline{\Psi}) = \prod_{i=1}^{\delta^{-d}} \frac{\left(\lambda \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k(W_i^\delta)\right)!}{\prod_{m=0}^{\infty} m!^{\lambda\mu_m(W_i^\delta)}} = \prod_{i=1}^{\delta^{-d}} \frac{\left(\lambda \sum_{m=0}^{\infty} m\mu_m(W_i^\delta)\right)!}{\prod_{m=0}^{\infty} m!^{\lambda\mu_m(W_i^\delta)}}. \tag{2.10}$$

Proof. We proceed in three steps by counting first the trajectories, registering only the partition sets W_i^δ that they travel through, second, for each $m \in \mathbb{N}_0$, the sets of relays in each partition set that receive precisely m ingoing hops and finally the choices of the relays for each hop in each partition set. Since every choice in the three steps can be freely combined with the other ones, the product of the three cardinalities is equal to the number of all trajectory configurations with the requested coarsened empirical measures.

(A) *Number of the transmitters of trajectories passing through given sequences of δ-subcubes.* For each configuration $s \in J^{\delta,\lambda}(\underline{\Psi})$ defined in (2.6), in each δ-subcube W_i^δ, $i = 1, \ldots, \delta^{-d}$, there are $\lambda\mu^{\delta,\lambda}(W_i^\delta)$ users. Out of them exactly $\lambda\nu_k^{\delta,\lambda}(W_i^\delta \times W_{i_1}^\delta \times \ldots W_{i_{k-1}}^\delta)$ take exactly k hops, having their first relay in $W_{i_1}^\delta$, their second in $W_{i_2}^\delta$ etc. and their $(k-1)$st relay in $W_{i_{k-1}}^\delta$, for any $k \in [k_{\max}]$ and $i_1, \ldots, i_{k-1} = 1, \ldots, \delta^{-d}$. Such choices in different sub-cubes W_i^δ corresponding to the transmitters are independent. Thus, the total number of such choices equals the number $N^1_{\delta,\lambda}(\underline{\Psi})$ defined in (2.8). Note that for $i = 1, \ldots, \delta^{-d}$,

$$\sum_{k=1}^{k_{\max}} \sum_{i_1,\ldots,i_{k-1}=1}^{\delta^{-d}} \nu_k^{\delta,\lambda}(W_i^\delta \times W_{i_1}^\delta \times \ldots \times W_{i_{k-1}}^\delta) = \sum_{k=1}^{k_{\max}} \pi_0 \nu_k^{\delta,\lambda}(W_i^\delta) = \mu^{\delta,\lambda}(W_i^\delta),$$

where we used part (5) of Definition 2.1; hence the multinomial expressions in (2.8) are well-defined.

(B) *Number of incoming messages.* In this step, for any δ-subcube W_i^δ, we count all the possible ways to distribute the incoming messages among the relays ($=$ users) $X_j \in W_i^\delta$, under the two constraints that in W_i^δ there are $\lambda\mu^{\delta,\lambda}(W_i^\delta)$ potential relays, and for any $m \in \mathbb{N}_0$, exactly $\lambda\mu_m^{\delta,\lambda}(W_i^\delta)$ receive precisely m incoming messages. Such choices are clearly independent of each other for different δ-subcubes. Hence, the total number of such choices equals the number $N^2_{\delta,\lambda}(\underline{\Psi})$ defined in (2.9). Again, the constraint (9) from Definition 2.1 implies that the multinomial expression (2.9) is well-defined. Clearly, all choices in this part are independent of the choices in part (A).

(C) *Number of assignments of the hops to the relays.* Assume that we have chosen one possible choice in part (A) and one possible choice in part (B). We now derive the number of possible ways of distributing, for any i, all the incoming hops in W_i^δ among the users in W_i^δ. Let us call this number M_i, then we know from part (A) that $M_i = \lambda \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k(W_i^\delta)$, since each such hop is the l-th of some of the trajectories for some l. The cardinality of the set of relays in W_i^δ is equal to $\lambda \sum_{m=0}^{\infty} \mu_m(W_i^\delta)$, and in part (B) we decomposed it into sets, indexed by m, in which each relay receives precisely m ingoing hops. Let us call such a relay an m-relay. Think of each such relay as being replaced by precisely m copies (in particular those with $m=0$ are discarded), then we have $\lambda \sum_{m=0}^{\infty} m\mu_m(W_i^\delta)$ virtual relays in W_i^δ. (Note that this is equal to M_i by one of our constraints.) Now, if all these m copies of the m-relays were distinguishable, then the number of ways to distribute the M_i ingoing hops to the relays would be simply equal to $M_i!$. However, since these m copies are identical, we overcount by a factor of $m!$ for any m-relay. This means that the number of hops into W_i^δ is equal to $M_i!/\prod_{m=0}^{\infty}(m!)^{\lambda\mu_m(W_i^\delta)}$. Since all these cardinalities can freely be combined with each other, we have deduced that the number of possible choices is equal to the number $N^3_{\delta,\lambda}(\underline{\Psi})$ defined in (2.10).

We also see that all the choices in the three parts are independent of each other, i.e., can be freely combined with each other and yield different combinations. Hence, we arrived at the assertion. $\square$

3 The limiting free energy: proofs of Theorems 1.2 and 1.4

In this section, we prove Theorem 1.2, that is, we derive the variational formula in (1.17) for the high-density (i.e., $\lambda \to \infty$) exponential rate of the partition function. Our first step is to derive the large-λ exponential rate of the combinatorial formulas for the empirical measures of Lemma 2.4 in Section 3.1. Furthermore, in Section 3.2 we formulate and prove how the interference term and the congestion term behave in the limits $\lambda \to \infty$, followed by $\delta \downarrow 0$. In Section 3.3, given an admissible trajectory setting, we construct a standard setting containing it. Using all these, in Section 3.4 we prove Theorem 1.2.

For the rest of this section, we fix the set $\Omega_1 \subset \Omega$ of full $\mathbb{P}$-measure on which we do our quenched investigations:

$$\begin{aligned}\Omega_1 = \Big\{&\omega \in \Omega \colon X_i(\omega) \in W_{\mathbb{B}} \ \forall i \in \mathbb{N},\\ &\lim_{\lambda\to\infty} \frac{\#\{i \in I^\lambda(\omega) \colon X_i(\omega) \in W_j^\delta\}}{\lambda} = \mu(W_j^\delta),\ \forall j = 1,\dots,\delta^{-d},\ \forall \delta \in \mathbb{B}\Big\}.\end{aligned} \tag{3.1}$$

That $\mathbb{P}(\Omega_1) = 1$ holds follows immediately from the Restriction Theorem [K93, Section 2.2] combined with the Poisson Law of Large Numbers [K93, Section 4.2] and the fact that μ is absolutely continuous.

3.1 The asymptotics of the combinatorics

Let us fix a controlled standard setting $\underline{\Psi}$ as in (2.2). Fix any $\omega \in \Omega_1$, and let the quantities I^λ and X^λ refer to this ω. Denote

$$N^0_{\delta,\lambda}(\underline{\Psi}) = \prod_{i=1}^{\delta^{-d}} \prod_{k=1}^{k_{\max}} \prod_{l=1}^{k-1} N(\lambda)^{\lambda \pi_l \nu_k^{\delta,\lambda}(W_i^\delta)}. \tag{3.2}$$

Recall the notation $H_V(\cdot \mid \cdot)$ from (1.27) and $c_m = \exp(-1/(\mathrm{e}\mu(W))(\mathrm{e}\mu(W))^{-m}/m!$ from (1.16). Note that the rate function I defined in (1.16) has also the representation

$$\mathrm{I}(\Psi) = \sum_{k=1}^{k_{\max}} H_{W^k}(\nu_k \mid \mu^{\otimes k}) - H_W\Big(\sum_{m=0}^{\infty} m\mu_m \Big| \mu\Big) + \sum_{m=0}^{\infty} \mathcal{H}_W(\mu_m \mid \mu c_m) + \mu(W) - 1 - \frac{1}{\mathrm{e}}, \quad (3.3)$$

which we are going to use here. We now identify the large-λ exponential rate of the cardinality of $J^{\delta,\lambda}(\underline{\Psi})$ both on the scale $\lambda \log \lambda$ and λ:

Proposition 3.1 (Exponential rates of counting terms)**.** *Let $\underline{\Psi}$ be a controlled standard setting. Let us write $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$. We have*

$$\lim_{\delta \downarrow 0} \lim_{\lambda \to \infty} \frac{1}{\lambda} \log \frac{\# J^{\delta,\lambda}(\underline{\Psi})}{N^0_{\delta,\lambda}(\underline{\Psi})} = -\mathrm{I}(\Psi),$$

as an equality in $[0, \infty]$. Moreover if $\mathrm{I}(\Psi) < \infty$, then

$$\lim_{\delta \downarrow 0} \lim_{\lambda \to \infty} \frac{1}{\lambda \log \lambda} \log \# J^{\delta,\lambda}(\underline{\Psi}) = \sum_{k=1}^{k_{\max}} (k-1)\nu_k(W^k) = \sum_{m=0}^{\infty} m\mu_m(W) < \infty,$$

almost surely.

Proof. Recall that Ψ is an admissible trajectory setting, according to Lemma 2.3. In particular, $\mathrm{I}(\Psi) \in (-\infty, \infty]$ is well-defined.

We use Stirling's formula $\lambda! = (\lambda/\mathrm{e})^{\lambda}\mathrm{e}^{o(\lambda)}$ in the limit $\lambda \to \infty$, which leads to

$$\lim_{\lambda \to \infty} \frac{1}{\lambda} \log \binom{a^{(\lambda)}}{a_1^{(\lambda)}, \dots, a_n^{(\lambda)}} = -\sum_{i=1}^{n} a_i \log \frac{a_i}{a}, \quad (3.4)$$

for any integers $a_1^{(\lambda)}, \dots, a_n^{(\lambda)}$ that sum up to $a^{(\lambda)}$ and satisfy $\frac{1}{\lambda} a_i^{(\lambda)} \stackrel{\lambda \to \infty}{\to} a_i$ for $i = 1, \dots, n$ with positive numbers $a_1, \dots, a_n$ satisfying $\sum_{i=1}^n a_i = a$.

From (2.8) we obtain that

$$\begin{aligned} I^1_\delta(\underline{\Psi}) &= -\lim_{\lambda \to \infty} \frac{1}{\lambda} \log N^1_{\delta,\lambda}(\underline{\nu}) \\ &= \sum_{i=1}^{\delta^{-d}} \sum_{k=1}^{k_{\max}} \sum_{i_1,\dots,i_{k-1}=1}^{\delta^{-d}} \nu_k^\delta(W_i^\delta \times W_{i_1}^\delta \times \dots \times W_{i_{k-1}}^\delta) \log \frac{\nu_k^\delta(W_i^\delta \times W_{i_1}^\delta \times \dots \times W_{i_{k-1}}^\delta)}{\mu^\delta(W_i^\delta)}, \end{aligned}$$

where we also used that all the measures $\nu_k^{\delta,\lambda}$ and $\mu^{\delta,\lambda}$ converge as $\lambda \to \infty$ to ν_k^δ and μ^δ, respectively.

Now we add the term $\prod_{l=1}^{k-1} \mu^\delta(W_{i_l}^\delta)$ both in the numerator and the denominator under the logarithm and separate these two terms. In the former, we write its logarithm as $\sum_{l=1}^{k-1} \log \mu^\delta(W_{i_l}^\delta)$, interchange this sum on l with all the other sums on the $i_0, \dots, i_{k-1}$ and write the sums over $i_0, \dots, i_{l-1}, i_{l+1}, \dots, i_{k-1}$ in terms of the l-th marginal measure of ν_k^δ. This gives

$$\begin{aligned} I^1_\delta(\underline{\Psi}) &= \sum_{i=1}^{\delta^{-d}} \sum_{k=1}^{k_{\max}} \sum_{i_1,\dots,i_{k-1}=1}^{\delta^{-d}} \nu_k^\delta(W_i^\delta \times W_{i_1}^\delta \times \dots \times W_{i_{k-1}}^\delta) \log \frac{\nu_k^\delta(W_i^\delta \times W_{i_1}^\delta \times \dots \times W_{i_{k-1}}^\delta)}{\mu^\delta(W_i^\delta) \prod_{l=1}^{k-1} \mu^\delta(W_{i_l}^\delta)} \\ &\quad + \sum_{i=1}^{\delta^{-d}} \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^\delta(W_i^\delta) \log \mu^\delta(W_i^\delta). \end{aligned} \quad (3.5)$$

In the same way as for I_1^δ, we obtain

$$I_\delta^2(\underline{\Psi}) = -\lim_{\lambda\to\infty}\frac{1}{\lambda}\log N_{\delta,\lambda}^2(\underline{\Psi}) = \sum_{i=1}^{\delta^{-d}}\sum_{m=0}^{\infty}\mu_m^\delta(W_i^\delta)\log\frac{\mu_m^\delta(W_i^\delta)}{\mu^\delta(W_i^\delta)}. \tag{3.6}$$

Using (3.1), on Ω_1 we have that the asymptotic behaviour of (3.2) is the following

$$N_{\delta,\lambda}^0(\underline{\Psi}) = N(\lambda)^{\lambda\sum_{i=1}^{\delta^{-d}}\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^{\delta,\lambda}(W_i^\delta)} = (\lambda\mu(W))^{\lambda(1+o(1))\sum_{i=1}^{\delta^{-d}}\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^{\delta,\lambda}(W_i^\delta)}.$$

On the other hand, also by Stirling's formula, we can identify the large-λ rate of the quotient of the counting terms in (2.10) and (3.2) as follows:

$$\begin{aligned} I_\delta^{3,0}(\underline{\Psi}) &= -\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\frac{N_{\delta,\lambda}^3(\underline{\Psi})}{N_{\delta,\lambda}^0(\underline{\Psi})} \\ &= -\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\prod_{i=1}^{\delta^{-d}}\frac{\left(\frac{1}{\mathrm{e}\mu(W)}\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^{\delta,\lambda}(W_i^\delta)\right)^{\lambda\sum_{k'=1}^{k_{\max}}\sum_{l'=1}^{k'-1}\pi_{l'}\nu_{k'}^{\delta,\lambda}(W_i^\delta)}}{\prod_{m=0}^{\infty}m!^{\lambda\mu_m(W_i^\delta)}} \\ &= -\sum_{i=1}^{\delta^{-d}}\sum_{k'=1}^{k_{\max}}\sum_{l'=1}^{k'-1}\pi_{l'}\nu_{k'}^\delta(W_i^\delta)\left(\log\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^\delta(W_i^\delta) - (1+\log\mu(W))\right) \\ &\quad + \sum_{i=1}^{\delta^{-d}}\sum_{m=0}^{\infty}\mu_m^\delta(W_i^\delta)\,\log(m!), \end{aligned} \tag{3.7}$$

where for the last term we used the fact that $\underline{\Psi}$ is controlled (see also Lemma 2.3), together with dominated convergence. We can summarize the sum of the terms in (3.5), (3.6) and (3.7) as

$$\begin{aligned} -\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\frac{\#J^{\delta,\lambda}(\underline{\Psi})}{N_{\delta,\lambda}^0(\underline{\Psi})} &= I_\delta^1(\underline{\Psi}) + I_\delta^2(\underline{\Psi}) + I_\delta^{3,0}(\underline{\Psi}) \\ &= \sum_{k=1}^{k_{\max}}\sum_{i_0,\dots,i_{k-1}=1}^{\delta^{-d}}\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)\log\frac{\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)}{\prod_{l=0}^{k-1}\mu^\delta(W_{i_l}^\delta)} \\ &\quad + \sum_{i=1}^{\delta^{-d}}\sum_{m=0}^{\infty}\mu_m^\delta(W_i^\delta)\log\frac{\mu_m^\delta(W_i^\delta)}{\mu^\delta(W_i^\delta)} \\ &\quad - \sum_{i=1}^{\delta^{-d}}\left(\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^\delta(W_i^\delta)\right)\log\frac{\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^\delta(W_i^\delta)}{\mu^\delta(W_i^\delta)} \\ &\quad + \sum_{m=0}^{\infty}\mu_m^\delta(W_\delta)[m(1+\log\mu(W)) + \log(m!)]. \end{aligned} \tag{3.8}$$

where in the first line on the right-hand side we changed the summing index i into i_0. Since we have

$$\sum_{m=0}^{\infty}\mu_m^\delta(W_\delta) = \sum_{m=0}^{\infty}\mu_m(W) = \mu(W),$$

and thus

$$\sum_{i=1}^{\delta^{-d}}\sum_{m=0}^{\infty}\mu_m^\delta(W_i^\delta)\log\frac{\mu_m^\delta(W_i^\delta)}{\mu^\delta(W_i^\delta)}+(m(1+\log\mu(W))+\log(m!))=\sum_{m=0}^{\infty}\mathcal{H}_{W_\delta}\big(\mu_m^\delta\mid\mu^\delta c_m\big)+\mu(W)-1-\frac{1}{\mathrm{e}},$$

we obviously arrived at the discrete version of the entropy terms in (3.3).

Now we argue that taking the limit as $\delta\downarrow 0$ through $\delta\in\mathbb{B}$, yields the desired entropy terms in (3.3). Let us begin with the first line on the right-hand side of (3.8). For $\delta\in\mathbb{B}$, let us define $(\nu_k'^\delta)_{k=1}^{k_{\max}}$ with $\nu_k'^\delta\in\mathcal{M}(W^k)$ as follows,

$$\nu_k'^\delta=\mu^{\otimes k}\sum_{i_0,\dots,i_{k-1}=1}^{\delta^{-d}}\mathbb{1}_{W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta}\frac{\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)}{\mu^{\otimes k}(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)},$$

so that for all k,

$$H_{W^k}(\nu_k'^\delta\mid\mu^{\otimes k})=\sum_{i_0,\dots,i_{k-1}=1}^{\delta^{-d}}\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)\log\frac{\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)}{\prod_{l=0}^{k-1}\mu^\delta(W_{i_l}^\delta)}.$$

Now, $\nu_k'^\delta$ also converges to ν_k in the weak topology of $\mathcal{M}(W^k)$, for all k. Therefore, by lower semicontinuity of the relative entropy (cf. [DZ98, Lemma 6.2.12 and Theorem D.12])

$$\liminf_{\delta\downarrow 0}\sum_{k=1}^{k_{\max}}\sum_{i_0,\dots,i_{k-1}=1}^{\delta^{-d}}\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)\log\frac{\nu_k^\delta(W_{i_0}^\delta\times\dots\times W_{i_{k-1}}^\delta)}{\prod_{l=0}^{k-1}\mu^\delta(W_{i_l}^\delta)}\geq\sum_{k=1}^{k_{\max}}H_{W^k}(\nu_k\mid\mu^{\otimes k}).\qquad(3.9)$$

On the other hand, by part (6) of Definition 2.1, for any $\delta',\delta\in\mathbb{B}$, $\delta'<\delta$, we have

$$\nu_k^\delta(W_i^\delta)=\nu_k^{\delta'}(W_i^\delta)=\sum_{j\in\{1,\dots,\delta^{-d}\}:\,W_j^{\delta'}\subseteq W_i^\delta}\nu_k^{\delta'}(W_j^{\delta'}),\quad\forall i=1,\dots,\delta^{-d}.$$

Therefore by Jensen's inequality, the complementary bound for $\limsup_{\delta\downarrow 0}$ follows, such that the limit exists with '$=$' instead of '$\geq$'. Similarly, we have the convergence of all the other terms on the right-hand side of (3.8) to their continuous counterparts. Indeed, using that by Lemma 2.3, Ψ satisfies (1.12)(iii), we conclude that

$$\lim_{\delta\downarrow 0}\sum_{i=1}^{\delta^{-d}}\left(\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^\delta(W_i^\delta)\right)\log\frac{\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^\delta(W_i^\delta)}{\mu^\delta(W_i^\delta)}=H_W\Big(\sum_{m=0}^{\infty}m\mu_m\mid\mu\Big).$$

Finally, we have

$$\lim_{\delta\downarrow 0}\sum_{m=0}^{\infty}\mathcal{H}_{W_\delta}\big(\mu_m^\delta\mid\mu^\delta c_m\big)+\mu(W)-1-\frac{1}{\mathrm{e}}=\sum_{m=0}^{\infty}\mathcal{H}\big(\mu_m\mid\mu c_m\big)+\mu(W)-1-\frac{1}{\mathrm{e}}.\qquad(3.10)$$

The first part of Proposition 3.1 follows.

Moreover, if $\mathrm{I}(\Psi)<\infty$, then we have by continuity

$$\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty}\frac{1}{\lambda\log\lambda}\log\#J^{\delta,\lambda}(\nu)=\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty}\frac{1}{\lambda\log\lambda}\log N^0_{\delta.\lambda}(\underline{\nu})$$

$$=\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty}\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\sum_{i=1}^{\delta^{-d}}\pi_l\nu_k^{\delta,\lambda}(W_i^\delta)=\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k(W)=\sum_{k=1}^{k_{\max}}(k-1)\nu_k(W^k)\in[0,\infty),$$

where in the last equality we used that by Fubini's theorem, $\pi_0\nu_k(W) = \nu_k(W^k)$ for all k. Hence, the second part of Proposition 3.1 follows. □

3.2 Approximations for the interference and the congestion terms

The limiting relations between the congestion terms in (1.9) and (1.15), and between the SIR terms in (1.7) and (1.14) are given as follows.

Proposition 3.2. *Let $\underline{\Psi}$ be a controlled standard setting. Let us write $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$ for the admissible trajectory setting contained in $\underline{\Psi}$. Then*

$$\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty}\sup_{s\in J^{\delta,\lambda}(\underline{\Psi})}\Big|\frac{1}{\lambda}\mathfrak{M}(s)-\mathrm{M}(\Psi)\Big| = 0, \tag{3.11}$$

and

$$\lim_{\delta\downarrow 0}\lim_{\lambda\to\infty}\sup_{s\in J^{\delta,\lambda}(\underline{\Psi})}\Big|\frac{1}{\lambda}\mathfrak{S}(s)-\mathrm{S}(\Psi)\Big| = 0. \tag{3.12}$$

Proof. First, we consider the congestion term. Consider some $s \in J^{\delta,\lambda}(\underline{\Psi})$ for $\lambda > 0$ and $\delta \in \mathbb{B}$. Additionally assume that $s_l^i \in W_{\mathbb{B}}$ for all $i \in I^\lambda$ and $l = 0, \dots, k$ (which is always the case for $s = (S^i)_{i\in I^\lambda}$ on Ω_1).

Then $P_\lambda^\delta(s) = \mu^{\delta,\lambda}$ and $P_{\lambda,m}^\delta(s) = \mu_m^{\delta,\lambda}$ for all $m \in \mathbb{N}_0$, see the definition (2.6) of $J^{\delta,\lambda}(\underline{\Psi})$ and (2.2). Recall that $m_i(s)$ is the number of ingoing messages at relay X_i for the trajectory configuration s. Hence we have

$$\begin{aligned}\mathfrak{M}(s) &= \sum_{i\in I^\lambda} m_i(s)(m_i(s)-1) = \sum_{m=0}^{\infty} m(m-1)\#\{i\in I^\lambda\colon m_i(s)=m\} = \sum_{m=0}^{\infty} m(m-1)P_{\lambda,m}(W)\\ &= \sum_{m=0}^{\infty} m(m-1)P_{\lambda,m}^\delta(W_\delta) = \lambda\sum_{m=0}^{\infty} m(m-1)\mu_m^{\delta,\lambda}(W_\delta),\end{aligned}$$

for all such s. Now, (2.4) in Definition 2.2, together with the fact that the total mass of μ_m^δ equals the one of μ_m for any m, implies the assertion in (3.11).

We continue with the SIR term. We start with defining discretized versions of SIR-related quantities. Let $\delta \in \mathbb{B}$ and $\mu_0 \in \mathcal{M}(W_\delta)$ be arbitrary. Then one can define a δ-discretized analogue of the definition (1.6) of SIR with a discrete interference term taken with respect to some measure μ_0 as follows

$$\mathrm{SIR}_\delta(\xi,\ \eta,\ \mu_0) = \frac{\ell(|\xi-\eta|)}{\int_{W_\delta}\ell(|\zeta-\eta|)\mu_0(\mathrm{d}\zeta)} = \frac{\ell(|\xi-\eta|)}{\sum_{i=1}^{\delta^{-d}}\mu_0(W_i^\delta)\ell(|C(W_i^\delta)-\eta|)}, \qquad \xi,\eta\in W_\delta,$$

where we recall that $C(W_i^\delta)$ denotes the centre of the δ-subcube W_i^δ. Furthermore, we define a δ-discretized version of the function $f_k(\mu, x_0, \dots, x_{k-1}) = f_k(x_0, \dots, x_{k-1})$ defined in (1.14) by

$$f_k^\delta(\mu_0,\xi_0,\dots,\xi_{k-1}) = \sum_{l=1}^{k}\mathrm{SIR}_\delta(\xi_{l-1},\ \xi_l,\mu_0^\delta)^{-1}, \qquad \mu_0\in\mathcal{M}(W),\xi_0,\dots,\xi_{k-1}\in W_\delta,$$

where we used the convention that all i_k-indexed sites are equal to the origin o, i.e., $C(W_{i_k}^\delta) = o = \xi_k$.

Towards the proof of Proposition 3.2, let us fix an arbitrary controlled standard setting $\underline{\Psi}$. Our goal is to prove that (3.12) holds for this $\underline{\Psi}$. Note that for an admissible trajectory setting $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$, $\mathrm{S}(\Psi)$ depends only on $(\nu_k)_{k=1}^{k_{\max}}$; observe that all SIR-related quantities in this paper depend only on the trajectories, but not on the numbers of incoming messages at the users.

Now for $\delta \in \mathbb{B}$ and $\lambda > 0$, we define the following discretized analogue of $\mathfrak{S}(\cdot)$, which corresponds to the case $P_\lambda^\delta(\cdot) = \mu^{\delta,\lambda}$, and $R_{\lambda,k}^\delta(\cdot) = \nu_k^{\delta,\lambda}$, $\forall k \in [k_{\max}]$, i.e., to configurations with empirical measure of users corresponding to $\mu^{\delta,\lambda}$ and empirical measure of trajectories with exactly k hops corresponding to $\nu_k^{\delta,\lambda}$ for all $k \in [k_{\max}]$:

$$S_{\delta,\lambda}(\underline{\Psi}) = \sum_{k=1}^{k_{\max}} \int_{W^k} \nu_k^{\delta,\lambda}(\mathrm{d}\xi_0, \dots, \mathrm{d}\xi_{k-1}) f_k^\delta(\nu_k^{\delta,\lambda}, \xi_0, \dots, \xi_{k-1}). \tag{3.13}$$

One easily sees that if $s \in \mathcal{S}_{k_{\max}}(X^\lambda)$ is such that $s_l^i \in W_{\mathbb{B}}$ for all $i \in I^\lambda$ and $l = 0, \dots, k$, and it holds that $s \in J^{\delta,\lambda}(\underline{\Psi})$, in particular $P_\lambda^\delta(s) = \mu^{\delta,\lambda}$ and $R_{\lambda,k}^\delta(s) = \nu_k^{\delta,\lambda}$ for all $k \in [k_{\max}]$, then we have

$$S_{\delta,\lambda}(\underline{\Psi}) = \frac{1}{\lambda} \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k} \sum_{i \in I^\lambda \colon s_{-1}^i = k} \frac{\frac{1}{\lambda} \sum_{m \in I^\lambda} \ell(|(s_0^m)^\delta - (s_l^i)^\delta|)}{\ell(|(s_{l-1}^i)^\delta - (s_l^i)^\delta|)} = \frac{1}{\lambda} \sum_{i \in I^\lambda} \sum_{l=1}^{s_{-1}^i} \mathrm{SIR}_\delta((s_{l-1}^i)^\delta,\ (s_l^i)^\delta,\ P_\lambda^\delta(s))^{-1}. \tag{3.14}$$

where we recall the notation $x^\delta = \varrho_\delta(x)$ for $x \in W_{\mathbb{B}}$.

Now, since (3.14) is true for all $s \in J^{\delta,\lambda}(\underline{\Psi})$, further ℓ is continuous and bounded from below, moreover $\nu_k^{\delta,\lambda}$ converges weakly to ν_k as first $\lambda \to \infty$ and then $\delta \downarrow 0$ (what one easily sees using parts (7) and (8) of Definition 2.1), we conclude that the following holds.

Lemma 3.3. *Let $\underline{\Psi}$ be a controlled standard setting. Then,*

$$\lim_{\delta \downarrow 0} \lim_{\lambda \to \infty} \sup_{s \in J^{\delta,\lambda}(\underline{\Psi})} \left| S_{\delta,\lambda}(\underline{\Psi}) - \frac{1}{\lambda} \mathfrak{S}(s) \right| = 0.$$

Having Lemma 3.3, the proof of Proposition 3.2 reduces to proving that $\lim_{\delta \downarrow 0} \lim_{\lambda \to \infty} S_{\delta,\lambda}(\underline{\Psi}) = \mathrm{S}(\Psi)$ for $\underline{\Psi}$ satisfying the assumptions of the Proposition. Now, for fixed $\delta \in \mathbb{B}$ and $k \in [k_{\max}]$, by the continuity of $W_\delta \to \mathbb{R}$, $\xi \mapsto \int_{W_\delta} \xi(\mathrm{d}y) \ell(|y - x|)$ and part (3) of Definition 2.1, we have

$$\lim_{\lambda \to \infty} f^\delta(\mu^{\delta,\lambda}, \xi_0, \dots, \xi_{k-1}) = f^\delta(\mu^\delta, \xi_0, \dots, \xi_{k-1}),$$

uniformly in $\xi_0, \dots, \xi_{k-1} \in W$. We thus conclude that

$$\sum_{k=1}^{k_{\max}} \int_{W_\delta^k} \nu_k^{\delta,\lambda}(\mathrm{d}\xi_0, \dots, \mathrm{d}\xi_{k-1}) f^\delta(\mu^{\delta,\lambda}, \xi_0, \dots, \xi_{k-1}) \underset{\lambda \to \infty}{\to} \sum_{k=1}^{k_{\max}} \int_{W_\delta^k} \nu_k^\delta(\mathrm{d}\xi_0, \dots, \mathrm{d}\xi_{k-1}) f^\delta(\mu^\delta, \xi_0, \dots, \xi_{k-1}). \tag{3.15}$$

Using this assumption and also part (8) of Definition (2.1) together with the boundedness and continuity properties of ℓ, it follows that we have

$$\lim_{\delta \downarrow 0} \lim_{\lambda \to \infty} \sum_{k=1}^{k_{\max}} \nu_k^{\delta,\lambda}(\mathrm{d}\xi_0, \dots, \mathrm{d}\xi_{k-1}) f_k^\delta(\mu^\delta, \xi_0, \dots, \xi_{k-1}) = \sum_{k=1}^{k_{\max}} \int_{W^k} \nu_k(\mathrm{d}x_0, \dots, \mathrm{d}x_{k-1}) f_k(\mathrm{d}x_0, \dots, \mathrm{d}x_{k-1}). \tag{3.16}$$

Thus, the proof of Proposition 3.2 is finished. □

3.3 Existence of standard settings

Recall that we equip $\mathcal{A}$ defined in (1.24) with the product topology of the weak topologies of the factors $\mathcal{M}(W^k)$ and that this is the topology of coordinatewise weak convergence. For $k \in \mathbb{N}$, let $d_k(\cdot,\cdot)$ be a metric on $\mathcal{M}(W^k)$ that generates the weak topology on this space. Then,

$$d_0(\Psi^1, \Psi^2) = \sum_{k=1}^{k_{\max}} d_k(\nu_k^1, \nu_k^2) + \sum_{m=0}^{\infty} 2^{-m} d_1(\mu_m^1, \mu_m^2), \quad \Psi^1, \Psi^2 \in \mathcal{A} \tag{3.17}$$

is a metric on $\mathcal{A}$ that generates the product topology. For $\varrho > 0$ and $\Psi \in \mathcal{A}$, let us write $B_\varrho(\Psi) = \{\Psi' \in \mathcal{A}\colon d_0(\Psi', \Psi) < \varrho\}$ for the open ϱ-ball around Ψ. It turns out to be convenient to choose d_k to be the *Lipschitz-bounded metric* [DZ98, Section D.2] on $\mathcal{M}(W^k)$, that is,

$$d_k(\nu_k^1, \nu_k^2) = \sup\{|\langle f, \nu_k^1\rangle - \langle f, \nu_k^2\rangle|\colon f \in \mathrm{Lip}_1(W^k)\}$$

for all k, where $\mathrm{Lip}_1(W^k)$ is the set of Lipschitz continuous functions taking W^k to $\mathbb{R}$ with Lipschitz parameter less than or equal to 1 and with uniform bound 1.

We have the following.

Proposition 3.4. *On Ω_1, for any admissible trajectory setting (see Definition 1.1), $\Psi = ((\nu_k)_k, (\mu_m)_m)$, there exists a standard setting $\underline{\Psi}$ containing it. If $\sum_m m(m-1)\mu_m(W) < \infty$, then $\underline{\Psi}$ can be chosen to be a controlled standard setting.*

Proof. We fix an admissible trajectory setting Ψ and construct $\underline{\Psi}$ as follows. As is required in Definition 2.1, the measures μ^δ, ν_k^δ for $k \in [k_{\max}]$ and μ_m^δ for $m \in \mathbb{N}_0$ are the δ-coarsenings of the measures μ, ν_k and μ_m, respectively, and $\mu^{\delta,\lambda} = L_\lambda^\delta$. Now for $\delta \in \mathbb{B}$ and $\lambda > 0$, pick some measures $\nu_k^{\delta,\lambda}$ and $\mu_m^{\delta,\lambda}$ with values in $\frac{1}{\lambda}\mathbb{N}_0$ such that the requirements (5) $\sum_{k=1}^{k_{\max}} \pi_0 \nu_k^{\delta,\lambda} = \mu^{\delta,\lambda}$, (9) $\sum_{m=0}^\infty \mu_m^{\delta,\lambda} = \mu^{\delta,\lambda}$ and (10) $\sum_{m=0}^\infty m\mu_m^{\delta,\lambda} = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^{\delta,\lambda}$ of Definition 2.1 are met, such that $\nu_k^{\delta,\lambda} \Longrightarrow \nu_k^\delta$ and $\mu_m^{\delta,\lambda} \Longrightarrow \mu_m^\delta$ as $\lambda \to \infty$ and such that the collection $\underline{\Psi}$ of all these measures is a standard setting containing Ψ, which is controlled if $\sum_m m(m-1)\mu_m(W) < \infty$.

We claim that this can be done by taking suitable up- and downroundings of the numbers

$$\nu_k'^{\delta,\lambda}(W_{s_0}^\delta \times \ldots \times W_{s_{k-1}}^\delta) = \nu_k^\delta(W_{s_0}^\delta \times \ldots \times W_{s_{k-1}}^\delta)\frac{L_\lambda^\delta(W_{s_0}^\delta)}{\mu^\delta(W_{s_0}^\delta)}\mathbb{1}\{\mu^\delta(W_{s_0}^\delta) > 0\}, \quad k \in [k_{\max}], \tag{3.18}$$

for all $s_0, \ldots, s_{k-1} = 1, \ldots, \delta^{-d}$, and dividing by λ, analogously for the μ_m's. Now, using the d-metric defined in (3.17), we prove that the convergences required in Definition 2.1 hold for such $\underline{\Psi}$.

First, we prove the convergence of the δ-coarsenings $\Psi^\delta = ((\nu_k^\delta)_k, (\mu_m^\delta)_m)$ to Ψ in the d_0-metric. We claim that for any $\varrho > 0$, there exists $\delta_0 \in \mathbb{B}$ such that $\Psi^\delta \in B_\varrho(\Psi)$ for all $\mathbb{B} \ni \delta \le \delta_0$. Indeed, for $k \in [k_{\max}]$, $\nu_k \in \mathcal{M}(W^k)$ and $\delta \in \mathbb{B}$ we see that the distance between ν_k and its δ-coarsening is of order δ:

$$\begin{aligned} d_k(\nu_k, \nu_k^\delta) &= \sup_{f \in \mathrm{Lip}_1(W^k)} \sum_{j_0,\ldots,j_{k-1}=1}^{\delta^{-d}} \int_{W_{j_0}^\delta \times \ldots \times W_{j_{k-1}}^\delta} |f(x) - f(C(W_{j_0}^\delta \times \ldots \times W_{j_{k-1}}^\delta))|\, \nu_k(\mathrm{d}x) \\ &\le \sum_{j_0,\ldots,j_{k-1}=1}^{\delta^{-d}} \int_{W_{j_0}^\delta \times \ldots \times W_{j_{k-1}}^\delta} |x - C(W_{j_0}^\delta \times \ldots \times W_{j_{k-1}}^\delta)|\, \nu_k(\mathrm{d}x) \le \nu_k(W^k)\frac{\sqrt{dk}\delta}{2}, \end{aligned}$$

where we wrote $x = (x_0, \dots, x_{k-1})$; and analogously for μ_m. Thus, we have

$$d_0(\Psi, \Psi^\delta) \le \delta \frac{\sqrt{d}}{2} \Big[\sum_{k=1}^{k_{\max}} \nu_k(W^k)\sqrt{k} + \sum_{m=0}^{\infty} \mu_m(W) 2^{-m} \Big].$$

Since $\sum_{m=0}^{\infty} \mu_m(W) < \infty$ by (ii) in (1.12), there exists a constant C, only depending on Ψ, such that $\Psi^\delta \in B_\varrho(\Psi)$ for any $\delta \le C\varrho$.

Second, we ignore the up- or downroundings in the construction of Ψ and prove the following. For $\delta \in \mathbb{B}$ and $\lambda > 0$, let $\Psi'^{\delta,\lambda}$ be the collection of the measures introduced in (3.18). We claim that on Ω_1, we have

$$\limsup_{\lambda \to \infty} d_0(\Psi^\delta, \Psi'^{\delta,\lambda}) = 0.$$

Indeed, for any $k \in [k_{\max}]$ and $s_0, \dots, s_{k-1} = 1, \dots, \delta^{-d}$, $d_k(\nu_k^\delta, \nu_k'^{\delta,\lambda})$ is bounded from above by

$$\sup_{f \in \mathrm{Lip}_1(W^k)} \sum_{s_0,\dots,s_{k-1}=1}^{\delta^{-d}} \nu_k^\delta(W_{s_0}^\delta \times \dots \times W_{s_{k-1}}^\delta) \Big| \frac{L_\lambda^\delta(W_{s_0}^\delta)}{\mu^\delta(W_{s_0}^\delta)} - 1 \Big| \|f\|_\infty \le \nu_k^\delta(W_\delta^k) \max_{s_0=1}^{\delta^{-d}} \Big| \frac{L_\lambda^\delta(W_{s_0}^\delta)}{\mu^\delta(W_{s_0}^\delta)} - 1 \Big|. \tag{3.19}$$

Thus,

$$d_0(\Psi^\delta, \Psi'^{\delta,\lambda}) \le \Big(\sum_{k=1}^{k_{\max}} \nu_k^\delta(W_\delta^k) + \sum_{m=0}^{\infty} 2^{-m} \mu_m^\delta(W_\delta) \Big) \max_{s_0=1}^{\delta^{-d}} \Big| \frac{L_\lambda^\delta(W_{s_0}^\delta)}{\mu^\delta(W_{s_0}^\delta)} - 1 \Big|,$$

which tends to 0 on Ω_1 as $\lambda \to \infty$, according to (3.1).

Now, if we add the suitable up- and downroundings, we only change distances in the d-metric by an error term of order $1/\lambda$, which vanishes as $\lambda \to \infty$. This implies that $\underline{\Psi}$ is a standard setting. It also follows easily that if $\sum_m m(m-1)\mu_m(W) < \infty$, then $\underline{\Psi}$ is controlled.

□

3.4 Proof of Theorem 1.2

Abbreviate

$$\mathfrak{Y}(r) = \Big(\prod_{i \in I^\lambda} N(\lambda)^{-(r_{-1}^i - 1)} \Big) \exp\Big\{ -\gamma \mathfrak{S}(r) - \beta \mathfrak{M}(r) \Big\}, \qquad \lambda > 0, r \in \mathcal{S}_{k_{\max}}(X^\lambda),$$

and note that the partition function is given as

$$Z_\lambda^{\gamma,\beta}(X^\lambda) = \sum_{r \in \mathcal{S}_{k_{\max}}(X^\lambda)} \mathfrak{Y}(r). \tag{3.20}$$

Then Theorem 1.2 says that its large-λ negative exponential rate is given as the infimum of $\mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)$, taken over all admissible trajectory settings Ψ. Throughout the proof, we assume that the configuration $X^\lambda = X^\lambda(\omega)$ comes from some $\omega \in \Omega_1$ defined in (3.1).

Having proved Propositions 3.1, 3.2 and 3.4, our strategy to prove Theorem 1.2 is the following. First, Proposition 3.4 gives a standard way how to construct from an admissible trajectory setting Ψ a standard setting $\underline{\Psi}$ that contains Ψ. Then the lower bound for the partition function is easily given in terms of the objects that are contained in any such $\underline{\Psi}$ and using the logarithmic asymptotics for their combinatorics from Propositions 3.1 and 3.2 and finally taking the infimum over all such Ψ, respectively $\underline{\Psi}$. The upper bound needs more care, since the entire sum over r has to be handled. First of all, we show that the sum can be restricted for all $\lambda > 0$, modulo some error term that is negligible on the exponential scale, to the sum of those configurations whose congestion exponent is at most $C\lambda$ for some appropriate large constant $C > 0$. This sum can be decomposed, for any $\delta \in \mathbb{B}$, to sums on configurations coming from a particular choice of empirical measures on the δ-partitions of W. The number of these empirical measures and the sum on the partitions is negligible in the limit $\lambda \to \infty$, and the asymptotics of the sums on r in these partitions can be evaluated with the help of our spatial discretization procedure, using arguments of the proofs of Propositions 3.1 and 3.2 in the limit $\lambda \to \infty$, followed by $\delta \downarrow 0$. Using these, we arrive at the said formula.

Let us give the details. We start with the proof of the lower bound. For any admissible trajectory setting Ψ, we pick $\underline{\Psi}$ as in Proposition 3.4 and recall the configuration class $J^{\delta,\lambda}(\underline{\Psi})$ from (2.6). Then, for any $\lambda > 0$ and $\delta \in \mathbb{B}$,

$$Z_\lambda^{\gamma,\beta}(X^\lambda) \geq \sum_{r \in J^{\delta,\lambda}(\underline{\Psi})} \mathfrak{Y}(r) \geq \frac{\# J^{\delta,\lambda}(\underline{\Psi})}{\sup_{r \in J^{\delta,\lambda}(\underline{\Psi})} \prod_{i \in I^\lambda} N(\lambda)^{-(r^i_{-1}-1)}} \exp\Big\{ - \sup_{r \in J^{\delta,\lambda}(\underline{\Psi})} \big(\gamma \mathfrak{S}(r) + \beta \mathfrak{M}(r)\big)\Big\}. \tag{3.21}$$

Hence,

$$\begin{aligned} \liminf_{\lambda\to\infty} \frac{1}{\lambda} \log Z_\lambda^{\gamma,\beta}(X^\lambda) &\geq \liminf_{\delta\downarrow 0} \liminf_{\lambda\to\infty} \frac{1}{\lambda} \log \frac{\# J^{\delta,\lambda}(\underline{\Psi})}{\sup_{r \in J^{\delta,\lambda}(\underline{\Psi})} \prod_{i \in I^\lambda} N(\lambda)^{-(r^i_{-1}-1)}} \\ &\quad - \gamma \limsup_{\delta\downarrow 0} \limsup_{\lambda\to\infty} \sup_{r \in J^{\delta,\lambda}(\underline{\Psi})} \frac{1}{\lambda} \mathfrak{S}(r) - \beta \limsup_{\delta\downarrow 0} \limsup_{\lambda\to\infty} \sup_{r \in J^{\delta,\lambda}(\underline{\Psi})} \frac{1}{\lambda} \mathfrak{M}(r) \\ &= -\mathrm{I}(\Psi) - \gamma \mathrm{S}(\Psi) - \beta \mathrm{M}(\Psi). \end{aligned} \tag{3.22}$$

In the last step we also used Propositions 3.1 and 3.2 together with the fact that $\underline{\Psi}$ is controlled. Now take the supremum over all such Ψ on the r.h.s. of (3.22) to conclude that the lower bound in (1.17) holds.

The upper bound of Theorem 1.2 requires some additional work. We start from (3.20). For $C > 0$ we have

$$Z_\lambda^{\gamma,\beta}(X^\lambda) = \sum_{r \in \mathcal{S}_{k_{\max}}(X^\lambda) \colon \mathfrak{M}(r) \leq \lambda C} \mathfrak{Y}(r) + \sum_{r \in \mathcal{S}_{k_{\max}}(X^\lambda) \colon \mathfrak{M}(r) > \lambda C} \mathfrak{Y}(r). \tag{3.23}$$

Since the total mass of our *a priori* measure has a bounded large-λ exponential rate (see Section 1.2.2), we see that

$$\limsup_{C\to\infty} \limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \sum_{r \in \mathcal{S}_{k_{\max}}(X^\lambda) \colon \mathfrak{M}(r) > \lambda C} \mathfrak{Y}(r) = -\infty.$$

Thus, for C sufficiently large, the exponential rate of $Z_\lambda^{\gamma,\beta}(X^\lambda)$ is equal to the one of the first term on the right-hand side of (3.23). We additionally require C so large that

$$\inf_{\Psi \text{ adm. traj. setting},\ \mathrm{M}(\Psi)\leq C} (\mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)) = \inf_{\Psi \text{ adm. traj. setting}} (\mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)). \tag{3.24}$$

Let us write $\mathcal{S}_{k_{\max},C}(X^\lambda) = \{r \in \mathcal{S}_{k_{\max}}(X^\lambda) \colon \mathfrak{M}(r) \leq \lambda C\}$ and $Z_\lambda^{\gamma,\beta,C}(X^\lambda) = \sum_{r \in \mathcal{S}_{k_{\max},C}(X^\lambda)} \mathfrak{Y}(r)$. The upper bound of Theorem 1.2 follows as soon as we show that

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log Z_\lambda^{\gamma,\beta,C}(X^\lambda) \leq - \inf_{\Psi \text{ admissible trajectory setting},\ \mathrm{M}(\Psi)\leq C} (\mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)). \tag{3.25}$$

For fixed $\lambda > 0$ and $\delta \in \mathbb{B}$, let us say that a collection of measures $\Psi^{\delta,\lambda} = ((\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}}, (\mu_m^{\delta,\lambda})_{m=0}^{\infty})$ lies in $G(\delta,\lambda) = G(\delta,\lambda)(X^\lambda)$ if all these measures take values in $\frac{1}{\lambda}\mathbb{N}_0$ only and satisfy the constraints $\sum_{k=1}^{k_{\max}} \pi_0 \nu_k^{\delta,\lambda} = L_\lambda^\delta$, $\sum_{m=0}^{\infty} \mu_m^{\delta,\lambda} = L_\lambda^\delta$ and $\sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \nu_k^{\delta,\lambda} = \sum_{m=0}^{\infty} m\mu_m^{\delta,\lambda}$. We will write $J^{\delta,\lambda}(\Psi^{\delta,\lambda})$ for the set $J^{\delta,\lambda}(\underline{\Psi})$ defined in (2.6). Then the union of $J^{\delta,\lambda}(\Psi^{\delta,\lambda})$ over all $\Psi^{\delta,\lambda}$ with $\sum_{m=0}^{\infty} m(m-1)\mu_m^{\delta,\lambda}(W_\delta) \le C$ is equal to

$$\big\{(R_{\lambda,k}^\delta(r))_{k\in[k_{\max}]}, (P_{\lambda,m}^\delta(r))_{m\in\mathbb{N}_0})\colon r \in \mathcal{S}_{k_{\max},C}(X^\lambda)\big\},$$

since these three equations characterize the tuple of the measures $(R_{\lambda,k}^\delta(S))_{k=1}^{k_{\max}}$ and $(P_{\lambda,m}^\delta(S))_{m=0}^{\infty}$ if $(S^i)_{i\in I^\lambda} \in \mathcal{S}_{k_{\max},C}(X^\lambda)$.

Using this, we can estimate, for any $\delta \in \mathbb{B}$,

$$Z_\lambda^{\gamma,\beta,C}(X^\lambda) = \sum_{\Psi^{\delta,\lambda}\in G(\delta,\lambda)\colon \mathrm{M}(\Psi^{\delta,\lambda})\le C} \; \sum_{r\in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \mathfrak{Y}(r) \le \#G(\delta,\lambda) \sup_{\Psi^{\delta,\lambda}\in G(\delta,\lambda)\colon \mathrm{M}(\Psi^{\delta,\lambda})\le C} \; \sum_{r\in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \mathfrak{Y}(r). \tag{3.26}$$

Hence,

$$\begin{aligned}
&\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log Z_\lambda^{\gamma,\beta,C}(X^\lambda) \\
&\qquad \le \limsup_{\delta\downarrow 0} \limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \#G(\delta,\lambda) \\
&\qquad\quad + \limsup_{\delta\downarrow 0} \limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \sup_{\Psi^{\delta,\lambda}\in G(\delta,\lambda)\colon \mathrm{M}(\Psi^{\delta,\lambda})\le C} \Big[\frac{\#J^{\delta,\lambda}(\Psi^{\delta,\lambda})}{\inf_{r\in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \prod_{i\in I^\lambda} N(\lambda)^{-(r_{-1}^i-1)}} \\
&\qquad\quad - \gamma \liminf_{\delta\downarrow 0} \liminf_{\lambda\to\infty} \inf_{r\in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \frac{1}{\lambda} \mathfrak{S}(r) - \beta \liminf_{\delta\downarrow 0} \liminf_{\lambda\to\infty} \inf_{r\in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \frac{1}{\lambda} \mathfrak{M}(r)\Big].
\end{aligned} \tag{3.27}$$

According to Lemma 3.5 below, the first term on the right-hand side is equal to zero. Now pick a sequence $(\delta_n)_n$ and for each n a sequence $(\lambda_{n,j})_j$ along which the superior limits as $n \to \infty$, respectively $j \to \infty$, are realized. Now pick, for any n and j, a maximizer $\widetilde{\Psi}^{\delta_n,\lambda_{n,j}}$. Pick λ_0 so large that $N(\lambda) \le 2\mu(W)\lambda$ for all $\lambda \ge \lambda_0$. Hence,

$$\bigcup_{\lambda>\lambda_0,\delta\in\mathbb{B}} G(\delta,\lambda) \subseteq \Big(\prod_{k=1}^{k_{\max}} \mathcal{M}_{\le 2\mu(W)}(W^k)\Big) \times \mathcal{M}_{\le 2\mu(W)}(W)^{\mathbb{N}_0}, \tag{3.28}$$

where we wrote $\mathcal{M}_{\le\alpha}(V)$ for the set of measures on a space V with total mass $\le \alpha$. (We recall from Section 2.1 that we conceive all measures on W_δ^k as measures on W^k.) Note that $\mathcal{M}_{\le 2\mu(W)}(W^k)$ is compact in the weak topology of $\mathcal{M}(W^k)$ for any k, according to Prohorov's theorem.

Without loss of generality (using two diagonal sequence arguments), we can assume that for all $n \in \mathbb{N}$, $\widetilde{\Psi}^{\delta_n,\lambda_{n,j}}$ converges coordinatewise weakly to a collection of measures $\widetilde{\Psi}^{\delta_n} = ((\widetilde{\nu}_k^{\delta_n})_{k=1}^{k_{\max}}, (\widetilde{\mu}_m^{\delta_n})_{m=0}^{\infty})$ as $j \to \infty$, and $\widetilde{\Psi}^{\delta_n}$ converges coordinatewise weakly to a collection of measures $\widetilde{\Psi}$ as $n \to \infty$. Then, it is clear that $\widetilde{\Psi}$ satisfies (i) from (1.12), and also that

$$\lim_{n\to\infty} \lim_{j\to\infty} \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \widetilde{\nu}_k^{\delta_n,\lambda_{n,j}} = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \widetilde{\nu}_k.$$

In order to see that (iii) holds for $\widetilde{\Psi}$, it remains to show that $\lim_{n\to\infty}\lim_{j\to\infty}\sum_{m=0}^{\infty} m\widetilde{\mu}_m^{\delta_n,\lambda_{n,j}} = \sum_{m=0}^{\infty} m\widetilde{\mu}_m$. For $N\in\mathbb{N}$ and for any continuous function $f\colon W\to\mathbb{R}$, we estimate

$$\left|\left\langle \sum_{m=0}^{\infty} m(\widetilde{\mu}_m^{\delta_n,\lambda_{n,j}}-\widetilde{\mu}_m), f\right\rangle\right| \leq \sum_{m=0}^{N} m\left|\langle \widetilde{\mu}_m^{\delta_n,\lambda_{n,j}}-\widetilde{\mu}_m, f\rangle\right| + \sum_{m=N+1}^{\infty} \|f\|_\infty m\left|\widetilde{\mu}_m^{\delta_n,\lambda_{n,j}}(W)-\widetilde{\mu}_m(W)\right|,$$

where we write $\langle \nu, f\rangle$ for the integral of the function f against the measure ν. The first term on the r.h.s. clearly tends to 0 as $j\to\infty$, followed by $n\to\infty$, for any fixed N. The second term can further be estimated from above as follows

$$\|f\|_\infty \sum_{m>N} \frac{m(m-1)}{N-1}(\widetilde{\mu}_m^{\delta_n,\lambda_{n,j}}(W)+\widetilde{\mu}_m(W)) \leq \frac{2\|f\|_\infty}{N}C.$$

This clearly tends to 0 as $N\to\infty$. One can analogously show that $\sum_{m=0}^{\infty}\widetilde{\mu}_m^{\delta_n,\lambda_{n,j}}$ tends to $\sum_{m=0}^{\infty}\widetilde{\mu}_m$ as $j\to\infty$ followed by $n\to\infty$, and hence condition (ii) from (1.12) holds. Also we have $\sum_{m=0}^{\infty} m(m-1)\widetilde{\mu}_m(W)\leq C$. Altogether, $\widetilde{\Psi}$ is an admissible trajectory setting.

Now, using the arguments of the proofs of Propositions 3.1 and 3.2 (which also involve the coarsened limits $\widetilde{\Psi}^{\delta_n}$ for fixed $n\in\mathbb{N}$) for the subsequential limits $j\to\infty$ followed by $n\to\infty$, we conclude that

$$\lim_{n\to\infty}\lim_{j\to\infty}\frac{\# J^{\delta_n,\lambda_{n,j}}(\widetilde{\Psi}^{\delta_n,\lambda_{n,j}})}{\inf_{r\in J^{\delta_n,\lambda_{n,j}}(\widetilde{\Psi}^{\delta_n,\lambda_{n,j}})}\prod_{i\in I^{\lambda_{n,j}}} N(\lambda_{n,j})^{-(r_{-1}^i-1)}} = \mathrm{I}(\widetilde{\Psi})$$

and

$$\lim_{n\to\infty}\lim_{j\to\infty}\inf_{r\in J^{\delta_n,\lambda_{n,j}}(\widetilde{\Psi}^{\delta_n,\lambda_{n,j}})}\frac{1}{\lambda_{n,j}}\mathfrak{S}(r) = \mathrm{S}(\widetilde{\Psi}).$$

Furthermore, Fatou's lemma implies that

$$-\beta\liminf_{n\to\infty}\liminf_{j\to\infty}\inf_{r\in J^{\delta_n,\lambda_{n,j}}(\widetilde{\Psi}^{\delta_n,\lambda_{n,j}})}\frac{1}{\lambda_{n,j}}\mathfrak{M}(r) \leq -\beta M(\widetilde{\Psi}). \tag{3.29}$$

Thus, we conclude that (3.25) (and therefore the upper bound in Theorem 1.2) holds, as soon as Lemma 3.5 is formulated and verified. This we do now.

Lemma 3.5. *For any $\delta\in\mathbb{B}$, almost surely,*

$$\limsup_{\lambda\to\infty}\frac{1}{\lambda}\log\# G(\delta,\lambda) = 0.$$

Proof. For $\lambda>0$, let $G_1(\delta,\lambda)$ denote the set of $(\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}}$ satisfying part (5) from Definition 2.1. It is easily seen that its cardinality increases only polynomially in λ. Now, given $(\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}}\in G_1(\delta,\lambda)$, we will give an upper bound for the number of $(\mu_m^{\delta,\lambda})_{m=0}^{\infty})$ such that the pair of these tuples is in $G(\delta,\lambda)$. This is much more demanding, since there is *a priori* no upper bound for m. We will provide a λ-dependent one.

For any $\lambda>0$, $\Psi^{\delta,\lambda}\in G(\delta,\lambda)$ and $j=1,\dots,\delta^{-d}$ we have that

$$\sum_{m=0}^{\infty} m\mu_m^{\delta,\lambda}(W_j^\delta) = \sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k^{\delta,\lambda}(W_j^\delta) \leq (k_{\max}-1)N(\lambda),$$

in particular $\mu_m^{\delta,\lambda}(W_j^\delta) = 0$ for $m > (k_{\max} - 1)N(\lambda)$. We also have that the numbers $\mu_0^{\delta,\lambda}(W_j^\delta)$, $\dots, \mu_{(k_{\max}-1)N(\lambda)}^{\delta,\lambda}(W_j^\delta)$, are $\frac{1}{\lambda}$ times nonnegative integers.

Let $\varepsilon > 0$ be fixed. We claim that for all sufficiently large $\lambda > 0$, there are not more than $\varepsilon N(\lambda) \sim \varepsilon\lambda\mu(W)$ nonzero ones out of these quantities. Indeed, if there were at least $\lceil \varepsilon N(\lambda) \rceil$ nonzero ones, denoted $\mu_{m_0}^{\delta,\lambda}(W_j^\delta), \dots, \mu_{m_{\lceil \varepsilon N(\lambda)\rceil-1}}^{\delta,\lambda}(W_j^\delta)$ with $0 \leq m_0 < m_1 < \dots < m_{\lceil \varepsilon N(\lambda)\rceil -1} \leq (k_{\max} - 1)N(\lambda)$, then we could estimate

$$(k_{\max}-1)N(\lambda) \geq \sum_{m=0}^{(k_{\max}-1)N(\lambda)} \lambda m \mu_m^{\delta,\lambda}(W_j^\delta) \geq \sum_{i=0}^{\lceil \varepsilon N(\lambda)\rceil -1} \lambda m_i \mu_{m_i}^{\delta,\lambda}(W_j^\delta) \mathbb{1}\{\mu_{m_i}^{\delta,\lambda}(W_j^\delta) > 0\}$$

$$= \sum_{i=0}^{\lceil \varepsilon N(\lambda)\rceil -1} \lambda m_i \mu_{m_i}^{\delta,\lambda}(W_j^\delta) \mathbb{1}\{\mu_{m_i}^{\delta,\lambda}(W_j^\delta) \geq \frac{1}{\lambda}\} \geq \sum_{i=0}^{\lceil \varepsilon N(\lambda)\rceil -1} m_i \geq \sum_{m=0}^{\lceil \varepsilon N(\lambda)\rceil -1} m \sim \frac{1}{2}(\varepsilon N(\lambda))(\varepsilon N(\lambda) - 1),$$

which is a contradiction for all $\lambda > 0$ sufficiently large.

Now, $\# G(\delta, \lambda)$ can be estimated as follows. Let us first fix $(\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}} \in G_1(\delta, \lambda)$, i.e., satisfying part (5) from Definition 2.1, and let us count the number of $(\mu_m^{\delta,\lambda})_{m=0}^{(k_{\max}-1)N(\lambda)}$ such that $((\nu_k^{\delta,\lambda})_{k=1}^{k_{\max}}, (\mu_m^{\delta,\lambda})_{m=0}^{(k_{\max}-1)N(\lambda)}))$ lies in $G(\delta, \lambda)$. Out of the $k_{\max}\delta^{-d}N(\lambda)$ quantities $\mu_0^{\delta,\lambda}(W_j^\delta), \dots, \mu_{(k_{\max}-1)N(\lambda)}^{\delta,\lambda}(W_j^\delta)$, $j = 1, \dots, \delta^{-d}$, at most $\lceil \varepsilon N(\lambda) \rceil \delta^{-d}$ are nonzero. The number of ways to choose them equals $\binom{k_{\max}N(\lambda)\delta^{-d}}{\lceil \varepsilon N(\lambda)\rceil \delta^{-d}}$. Having chosen $\lceil \varepsilon N(\lambda) \rceil \delta^{-d}$ potentially nonzero ones so that the remaining $k_{\max}\delta^{-d}N(\lambda) - \lceil \varepsilon N(\lambda) \rceil \delta^{-d}$ ones are equal to zero, according to part (9) of Definition 2.1 we note that the potentially nonzero ones sum up to $N(\lambda)$, and each one has a value in $\frac{1}{\lambda}\mathbb{N}_0$. For this, there are at most $\binom{N(\lambda)+\lceil \varepsilon N(\lambda)\rceil\delta^{-d}-1}{\lceil \varepsilon N(\lambda)\rceil \delta^{-d}-1}$ combinations, for any choice of the set of the potentially nonzero ones. Therefore, using Stirling's formula as in (3.4), for any sufficiently large λ, we have the following estimate

$$\begin{aligned}
\# G(\delta, \lambda) &\leq \# G_1(\delta, \lambda) \binom{k_{\max}N(\lambda)\delta^{-d}}{\lceil \varepsilon N(\lambda)\rceil\delta^{-d}} \binom{N(\lambda) + \lceil \varepsilon N(\lambda)\rceil \delta^{-d} - 1}{\lceil \varepsilon N(\lambda)\rceil \delta^{-d} - 1} \\
&= \mathrm{e}^{o(\lambda)} \exp\Big(- \lambda\mu(W)\Big((k_{\max} - \varepsilon)\delta^{-d} \log \frac{(k_{\max} - \varepsilon)\delta^{-d}}{k_{\max}\delta^{-d}} + \varepsilon\delta^{-d} \log \frac{\varepsilon\delta^{-d}}{k_{\max}\delta^{-d}}\Big)\Big) \\
&\quad \times \exp\Big(- \lambda\mu(W)\Big(\varepsilon\delta^{-d} \log \frac{\varepsilon\delta^{-d}}{1 + \varepsilon\delta^{-d}} + \log \frac{1}{1 + \varepsilon\delta^{-d}}\Big)\Big).
\end{aligned}$$

Making $\varepsilon \downarrow 0$, we conclude that $\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \# G(\delta, \lambda) = 0$. □

3.5 The large deviation principle: proof of Theorem 1.4(i)

In this section, we prove Theorem 1.4(i). The combinatorial essence of this theorem has already been proven in Proposition 3.1, including the relations with δ-coarsenings. What remains to be done is to relate this to the coordinatewise weak convergence on $\mathcal{A}$. We will be able to use some of the arguments of Section 3.4.

The lower semicontinuity of $\mathrm{I} + \mu(W) \log k_{\max}$ was already discussed in Section 1.3, the nonnegativity in Section 1.5. These together mean that $\mathrm{I} + \mu(W) \log k_{\max}$ is a rate function.

We proceed with the proof of the lower bound. Let $G \subseteq \mathcal{A}$ be open. If $\inf_G I = \infty$, then there is nothing to show, therefore let us assume that there exists $\Psi \in G$ with $\mathrm{I}(\Psi) < \infty$. According to Proposition 3.4, there is a standard setting $\underline{\Psi}$ containing Ψ. Since G is open, there exists $\varrho > 0$ such that $B_\varrho(\Psi) \subseteq G$. Let us choose $\delta_0 \in \mathbb{B}$ and, for any $\mathbb{B} \ni \delta \leq \delta_0$, some $\lambda_0 = \lambda_0(\delta) > 0$ such that $\Psi^\delta, \Psi^{\delta,\lambda} \in B_\varrho(\Psi)$ for any $\lambda > \lambda_0$. Now we can estimate, for these δ and λ,

$$\begin{aligned} \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in G) &\geq \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in B_\varrho(\Psi)) \geq \mathrm{P}^{0,0}_{\lambda,X^\lambda}\big((\Psi_\lambda(S))^\delta = \Psi^{\delta,\lambda}\big) \\ &= \frac{1}{Z^{0,0}_\lambda(X^\lambda)} \sum_{r \in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \frac{1}{\prod_{i \in I^\lambda} N(\lambda)^{r^i_{-1}-1}} \geq \frac{\# J^{\delta,\lambda}(\Psi^{\delta,\lambda})}{k_{\max}^{N(\lambda)} \sup_{r \in J^{\delta,\lambda}(\Psi^{\delta,\lambda})} \prod_{i \in I^\lambda} N(\lambda)^{r^i_{-1}-1}}. \end{aligned}$$

Now, using Proposition 3.1 and the fact that $N(\lambda)/\lambda \to \mu(W)$, we obtain

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda} \log \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in G) \geq -\mu(W) \log k_{\max} - \mathrm{I}(\Psi).$$

Note that $\underline{\Psi}$ is not necessarily controlled because $M(\Psi) < \infty$ is not guaranteed. However, since for all $\delta \in \mathbb{B}$, $s = 1, \dots, \delta^{-d}$, $\lambda > 0$, $\mu^{\delta,\lambda}_m(W^\delta_s)/\mu^\delta_m(W^\delta_s)$ does not depend on m, we easily see that Proposition 3.1 holds for this $\underline{\Psi}$ as well. Now, take the supremum over $\Psi \in G \cap \{I < \infty\}$ to conclude that the lower bound holds.

We continue with the upper bound. Let $F \subseteq \mathcal{A}$ be closed. Let us choose an increasing sequence $(\lambda_n)_{n\in\mathbb{N}}$ of positive numbers along which the limit superior in (1.25) is realized. For $\lambda > 0$, let us put

$$O(\lambda) = \big\{\Psi \in \mathcal{A} \colon \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) = \Psi) > 0\big\}.$$

If for all but finitely many $n \in \mathbb{N}$ we have $F \cap O(\lambda_n) = \varnothing$, then

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda} \log \mathrm{P}^{0,0}_{\lambda,X^\lambda}(\Psi_\lambda(S) \in F) = -\infty. \tag{3.30}$$

Therefore, without loss of generality, we can assume that $O(\lambda_n) \cap F$ is non-empty for all $n \in \mathbb{N}$. For $\delta \in \mathbb{B}$ and $A \subset \mathcal{A}$, let us write $A^\delta = \{\Psi^\delta \colon \Psi \in A\}$, where Ψ^δ is the coordinatewise δ-coarsened version of Ψ. Then we have

$$\begin{aligned} \mathrm{P}^{0,0}_{\lambda_n,X^{\lambda_n}}\big(\Psi_{\lambda_n}(S) \in F\big) &= \mathrm{P}^{0,0}_{\lambda_n,X^{\lambda_n}}\big(\Psi_{\lambda_n}(S) \in F \cap O(\lambda_n)\big) = \mathrm{P}^{0,0}_{\lambda_n,X^{\lambda_n}}\big((\Psi_{\lambda_n}(S))^\delta \in (F \cap O(\lambda_n))^\delta\big) \\ &\leq \#(F \cap O(\lambda_n))^\delta \sup_{\Psi \in F \cap O(\lambda_n)} \frac{\# J^{\delta,\lambda_n}(\Psi^\delta)}{k_{\max}^{N(\lambda_n)} \inf_{r \in J^{\delta,\lambda_n}(\Psi^\delta)} \prod_{i \in I^{\lambda_n}} N(\lambda_n)^{r^i_{-1}-1}}. \end{aligned} \tag{3.31}$$

It is clear that $(F \cap O(\lambda_n))^\delta \subseteq G(\delta,\lambda_n) = (O(\lambda_n))^\delta$ for all $n \in \mathbb{N}$ and $\delta \in \mathbb{B}$, where $G(\delta,\lambda_n)$ was defined in Section 3.4. Hence, by Lemma 3.5,

$$\limsup_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{\lambda_n} \log \#(F \cap O(\lambda_n))^\delta = 0.$$

It remains to show that

$$\limsup_{\delta\downarrow 0} \limsup_{n\to\infty} \frac{1}{\lambda_n} \log \Big[\sup_{\Psi \in F \cap O(\lambda_n)} \frac{\# J^{\delta,\lambda_n}(\Psi^\delta)}{\inf_{r \in J^{\delta,\lambda_n}(\Psi^\delta)} \prod_{i \in I^{\lambda_n}} N(\lambda_n)^{r^i_{-1}-1}} \Big] \leq -\inf_{\Psi \in F} \mathrm{I}(\Psi). \tag{3.32}$$

One can do this analogously to the proof of the upper bound of Theorem 1.2 starting from (3.27). Indeed, using Prohorov's theorem together with a diagonal sequence argument, we find $\Psi^* \in \mathcal{A}$ that the maximizer in (3.31) converges to along a subsequence of δ's and λ_n's. The limit lies in F because F is closed. Using the lower semicontinuity of I together with Fatou's lemma, we conclude that the left-hand side of (3.32) is not larger than $-\mathrm{I}(\Psi^*)$, which itself is not larger than $-\inf_F I$. This finishes the proof of the upper bound in Theorem 1.4 (i).

4 Analysis of the minimizers

This section is devoted to the proof of Proposition 1.3. In particular, in Section 4.1, we show that the infimum in (1.17) is attained and, for any minimizer $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$, for any $k \in [k_{\max}]$, $\mu^{\otimes k}$ is absolutely continuous with respect to ν_k and μ is absolutely continuous with respect to each μ_m. This is a prerequisite for perturbing the minimizer in many admissible directions. In Section 4.2 we finish the proof of Proposition 1.3 by deriving the Euler–Lagrange equations. For the remainder of this section, we fix all parameters W, μ, γ, β and $k_{\max}$. Moreover, we use the following representation of I from (1.16).

$$\mathrm{I}(\Psi) = \sum_{k=1}^{k_{\max}} H_{W^k}(\mu \otimes M^{\otimes(k-1)}) + \sum_{m=0}^{\infty} H_W(\mu_m \mid \mu) - \mu_m(W) \log \frac{(\mathrm{e}\mu(W))^{-m}}{m!}.$$

4.1 Existence and positivity of the minimizers

We start with the following lemma, which follows almost immediately from the arguments of the proof of the upper bound of Theorem 1.2 in Section 3.4.

Lemma 4.1. *The set of minimizers for the variational formula in* (1.17) *is non-empty, compact and convex.*

Proof. Recall that the three functionals I, S, M are lower semicontinuous and convex. Furthermore, it is clear that we can restrict the infimum in (1.17) to those Ψ that satisfy also $\mathrm{M}(\Psi) \le C$ for any sufficiently large C. But, as we have seen in Section 3.4, this set of Ψ's is compact. From this, all our assertions easily follow. □

Now we prove that, for each minimizer Ψ, $\mu^{\otimes k}$ is absolutely continuous with respect to ν_k and μ is absolutely continuous with respect to each μ_m. (Note that the opposite absolute continuities are true by finiteness of the entropies.) We need to show this only for $k_{\max} > 1$, as we explained after Proposition 1.3.

Lemma 4.2. *If $k_{\max} > 1$ and $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^{\infty})$ is a minimizer of* (1.17)*, then $\mu^{\otimes k} \ll \nu_k$ for any $k \in [k_{\max}]$, and $\mu \ll \mu_m$ for any $m \in \mathbb{N}_0$.*

Proof. The essence of the proof is the following. The congestion term $\mathrm{M}(\cdot)$ and the SIR term $\mathrm{S}(\cdot)$ are linear in each μ_m respectively ν_k, as well as the third term in $\mathrm{I}(\cdot)$ in (1.16) in each μ_m. On the other hand, the function $x \mapsto x \log x$ has the slope $-\infty$ at $x \downarrow 0$. We show the following assertions about the minimizer Ψ step by step as follows. Recall that $M = \sum_{m \in \mathbb{N}_0} m\mu_m = \sum_{k \in [k_{\max}]} \sum_{l=1}^{k-1} \pi_l \nu_k$. We write $\ge$ and $>$, respectively, between measures in $\mathcal{M}(W^k)$ if their difference lies in $\mathcal{M}(W^k)$, respectively in $\mathcal{M}(W^k) \setminus \{0\}$.

Fix a measurable set $A \subset W$ such that $\mu(A) > 0$. Then we have:

1 $M(A) > 0$.

2 for any $m_1 < m_0 < m_2$ such that $\mu_{m_1}(A) > 0$ and $\mu_{m_2}(A) > 0$, also $\mu_{m_0}(A) > 0$.

3 $\mu_0(A) > 0$.

4 $\mu_m(A) > 0$ for any $m \ge k_{\max}$.

5 $\nu_k(A^k) > 0$ for any $k \in [k_{\max}]$.

Indeed, these steps are verified respectively as follows. In each of the steps, for $\varepsilon \in (0,1)$, we construct an admissible trajectory setting $\Psi^\varepsilon = ((\nu_k^\varepsilon)_{k=1}^{k_{\max}}, (\mu_m^\varepsilon)_{m=0}^\infty)$ such that $\mathrm{I}(\Psi^\varepsilon) + \gamma \mathrm{S}(\Psi^\varepsilon) + \beta \mathrm{M}(\Psi^\varepsilon) < \mathrm{I}(\Psi) + \gamma \mathrm{S}(\Psi) + \beta \mathrm{M}(\Psi)$ for sufficiently small $\varepsilon > 0$, and therefore Ψ is not a minimizer of (1.17).

1 If $M(A) = 0$, then in particular $\mu_0(A) = \nu_1(A) = \mu(A)$ and $\mu_m(A) = 0$ for all $m > 0$. Also, $\pi_1\nu_2(A) = \nu_2(W \times A) = 0$, according to the definition of M.

Let us define Ψ^ε as follows: $\nu_2^\varepsilon = (1-\varepsilon)\nu_2 + \varepsilon(\mu^{\otimes 2})/\mu(W)$, $\nu_k^\varepsilon = (1-\varepsilon)\nu_k$ for $k \neq 2$, $\mu_1^\varepsilon = (1-\varepsilon)\mu_1 + \varepsilon\mu$ and $\mu_m^\varepsilon = (1-\varepsilon)\mu_m$ for $m \neq 1$. Then we compute and estimate the three terms of the entropy $\mathrm{I}(\Psi)$ as follows.

$$\begin{aligned}
&\sum_{k=1}^{k_{\max}} H_{W^k}\big(\nu_k^\varepsilon \mid \mu \otimes (M^\varepsilon)^{\otimes(k-1)}\big) \\
&\leq \sum_{k=1}^{k_{\max}} H_{W\times(W\setminus A)^{k-1}}((1-\varepsilon)\nu_k \mid \mu \otimes (M^\varepsilon)^{\otimes(k-1)}) + H_{W\times A}\Big(\frac{\varepsilon\mu^{\otimes 2}}{\mu(W)} \mid \varepsilon\mu^{\otimes 2}\Big) + \mathcal{O}(\varepsilon) \\
&\leq \sum_{k=1}^{k_{\max}} H_{W^k}(\mu \otimes M^{\otimes(k-1)}) + \mathcal{O}(\varepsilon),
\end{aligned}$$

furthermore

$$\begin{aligned}
&\sum_{m=0}^{\infty} H_W(\mu_m^\varepsilon \mid \mu) - \mu_m^\varepsilon(W) \log \frac{(\mathrm{e}\mu(W))^{-m}}{m!} \\
&\leq H_W((1-\varepsilon)\mu_m \mid \mu) - \mu_m(W) \log \frac{(\mathrm{e}\mu(W))^{-m}}{m!} + \mu(A)\varepsilon\log\varepsilon + \mathcal{O}(\varepsilon).
\end{aligned} \tag{4.1}$$

For the second term we used the convexity of the relative entropy in the form

$$H_W((1-\varepsilon)\nu_1 + \varepsilon\mu \mid \mu) \leq (1-\varepsilon) H_W(\nu_1 \mid \mu) \leq H_W(\nu_1 \mid \mu) + \mathcal{O}(\varepsilon). \tag{4.2}$$

This in turn follows from [HJKP15, Lemmas 3.10, 3.11], which implies that, for any $k \in \mathbb{N}$, $\xi, \eta \in \mathcal{M}(W^k)$ with $\eta \neq 0$ and $\xi \ll \eta$,

$$\Big| H_{W^k}(\xi \mid \eta) - H_{W^k}((1-\varepsilon)\xi \mid \eta) \Big| \underset{\varepsilon\downarrow 0}{\asymp} \varepsilon.$$

It follows that, as $\varepsilon \downarrow 0$,

$$\mathrm{I}(\Psi^\varepsilon) + \gamma S(\Psi^\varepsilon) + \beta \mathrm{M}(\Psi^\varepsilon) - \big[\mathrm{I}(\Psi) + \gamma S(\Psi) + \beta \mathrm{M}(\Psi)\big] \leq \mathcal{O}(\varepsilon) + \mu(A)\varepsilon\log\varepsilon, \tag{4.3}$$

which is negative for all sufficiently small $\varepsilon > 0$. Thus, Ψ is not a minimizer.

2 If $M(A) > 0$ but $\mu_{m_1}(A) > 0$, $\mu_{m_2}(A) > 0$ and $\mu_{m_0}(A) = 0$ for some $m_1 < m_0 < m_2$, then let $\nu_k^\varepsilon = \nu_k$ for all $k \in [k_{\max}]$ and let $\mu_{m_0}^\varepsilon = (1-\varepsilon)\mu_{m_0} + \varepsilon(\alpha_1\mu_{m_1} + \alpha_2\mu_{m_2})$, $\mu_{m_1}^\varepsilon = (1-\alpha_1\varepsilon)\mu_{m_1}$, $\mu_{m_2}^\varepsilon = (1-\varepsilon\alpha_2)\mu_{m_2}$, where $\alpha_1, \alpha_2 \in (0,1)$ are such that $\alpha_1 + \alpha_2 = 1$ and $m_1\alpha_1 + m_2\alpha_2 = m_0$. Then, Ψ^ε is an admissible trajectory setting with $M^\varepsilon = M$. It follows similarly to the previous

computation that $\mathrm{I}(\Psi^\varepsilon)+\gamma\mathrm{S}(\Psi^\varepsilon)+\beta\mathrm{M}(\Psi^\varepsilon)<\mathrm{I}(\Psi)+\gamma\mathrm{S}(\Psi)+\beta\mathrm{M}(\Psi)$ for all sufficiently small $\varepsilon>0$. However, instead of (4.1), we have

$$\begin{aligned}&\sum_{m=0}^\infty H_W(\mu_m^\varepsilon\mid\mu)-\mu_m^\varepsilon(W)\log\frac{(\mathrm{e}\mu(W))^{-m}}{m!}\\&\le\sum_{m=0}^\infty H_W(\mu_m\mid\mu)-\mu_m(W)\log\frac{(\mathrm{e}\mu(W))^{-m}}{m!}+(\alpha_1\mu_{m_1}(A)+\alpha_2\mu_{m_2}(A))\varepsilon\log\varepsilon+\mathcal{O}(\varepsilon),\end{aligned}$$

as $\varepsilon\downarrow 0$.

3 If $M(A)>0$ but $\mu_0(A)=0$, let $\nu_k^\varepsilon=(1-\varepsilon)\nu_k$ for all $1<k\le k_{\max}$, $\mu_m^\varepsilon=(1-\varepsilon)\mu_m$ for all $m>0$, $\mu_0^\varepsilon=\varepsilon\mu+(1-\varepsilon)\mu_0$ and $\nu_1^\varepsilon=(1-\varepsilon)\nu_1+\varepsilon\mu$. It is again sufficient to consider the entropy terms in I. The summands on $k>1$ can be estimated as follows.

$$\begin{aligned}\sum_{k=2}^{k_{\max}}H_{W^k}(\nu_k^\varepsilon\mid\mu\otimes(M^\varepsilon)^{(k-1)})&=\sum_{k=2}^{k_{\max}}H_{W^k}((1-\varepsilon)\nu_k\mid(1-\varepsilon)^{k-1}\mu\otimes M^{k-1})\\&\le\sum_{k=2}^{k_{\max}}H_{W^k}(\nu_k\mid\mu\otimes M^{(k-1)})+\mathcal{O}(\varepsilon).\end{aligned}$$

The summand for $k=1$ can be estimated with the help of (4.2) For the summand for $m=0$, we have

$$\begin{aligned}H_W(\mu_0^\varepsilon\mid\mu)&=H_{W\setminus A}((1-\varepsilon)\mu_0+\varepsilon\mu\mid\mu)+\mu(A)\varepsilon\log\varepsilon\\&\le H_{W\setminus A}((1-\varepsilon)\mu_0\mid\mu)+\mu(A)\varepsilon\log\varepsilon+\mathcal{O}(\varepsilon)=H_W(\mu_0\mid\mu)+\mu(A)\varepsilon\log\varepsilon+\mathcal{O}(\varepsilon).\end{aligned}$$

while the remaining sum is handled as follows.

$$\begin{aligned}&\sum_{m=1}^\infty H_W(\mu_m^\varepsilon\mid\mu)-\mu_m^\varepsilon(W)\log\frac{(\mathrm{e}\mu(W))^{-m}}{m!}\\&=\sum_{m=1}^\infty H_W((1-\varepsilon)\mu_m\mid\mu)-\mu_m(W)\log\frac{(\mathrm{e}\mu(W))^{-m}}{m!}+\mathcal{O}(\varepsilon).\end{aligned}$$

Thus, (4.3) holds also here, which implies the claim.

4 If $M(A)>0$ but $\mu_{m_0}(A)=0$ for some $m_0\ge k_{\max}$, let $\mu_{m_0}^\varepsilon=(1-\varepsilon)\mu_{m_0}+\varepsilon M/m_0$, $\mu_m^\varepsilon=(1-\varepsilon)\mu_m$ for $m\notin\{0,m_0\}$ and, moreover $\nu_k^\varepsilon=\nu_k$ for all $k\in[k_{\max}]$.

$$\sum_{m=1}^\infty m\mu_m^\varepsilon=(1-\varepsilon)\sum_{m=1}^\infty m\mu_m+\frac{\varepsilon m_0}{m_0}\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k=\sum_{k=1}^{k_{\max}}\sum_{l=1}^{k-1}\pi_l\nu_k,$$

as required.

On the other hand, we have

$$\mu-\sum_{m=1}^\infty\mu_m^\varepsilon\ge\mu-(1-\varepsilon)\sum_{m=1}^\infty\mu_m-\frac{\varepsilon(k_{\max}-1)}{m_0}\mu\ge(1-\varepsilon)\mu-(1-\varepsilon)\sum_{m=1}^\infty\mu_m=(1-\varepsilon)\mu_0.$$

Therefore, if we put $\mu_0^\varepsilon = \mu - \sum_{m=1}^\infty \mu_m^\varepsilon$, then $\mu_0^\varepsilon \geq (1-\varepsilon)\mu_0$ and Ψ^ε is an admissible trajectory setting. Now we can proceed analogously to (3) to conclude that $\mathrm{I}(\Psi^\varepsilon) + \gamma\mathrm{S}(\Psi^\varepsilon) + \beta\mathrm{M}(\Psi^\varepsilon) < \mathrm{I}(\Psi) + \gamma\mathrm{S}(\Psi) + \beta\mathrm{M}(\Psi)$ for sufficiently small $\varepsilon > 0$.

The proof of (5) is very similar to the ones of (2), (3) and (4), therefore we leave it to the reader.

□

4.2 Deriving the Euler–Lagrange equations

In this section, we finish the proof of Proposition 1.3. According to the results of Section 4.1, now we see that (1.17) exhibits at least one minimizer, and all minimizers have almost everywhere positive Lebesgue density on the corresponding powers of $\operatorname{supp}\mu$. Knowing this, we now carry out the perturbation analysis for the minimizer(s) of the optimization problem in (1.17) and derive the shape of the minimizers in most explicit terms.

We use the method of Lagrange multipliers in the framework of a perturbation argument. Let $\Psi = ((\nu_k)_{k=1}^{k_{\max}}, (\mu_m)_{m=0}^\infty)$ minimize (1.17). Fix any collection of signed measures $\Phi = ((\tau_k)_{k=1}^{k_{\max}}, (\sigma_m)_{m=0}^\infty)$ such that only finitely many σ_m's are different from zero, each τ_k and each σ_m has a simple function as a Lebesgue density and they satisfy the following constraints:

$$(i)\quad \sum_{k=1}^{k_{\max}} \pi_0 \tau_k = 0, \qquad (ii)\quad \sum_{m=0}^\infty \sigma_m = 0, \qquad (iii)\quad \sum_{m=0}^\infty m\sigma_m = \sum_{k=1}^{k_{\max}} \sum_{l=1}^{k-1} \pi_l \tau_k. \tag{4.4}$$

Then it follows from Lemma 4.2 that, for any $\varepsilon \in \mathbb{R}$ with sufficiently small $|\varepsilon|$, $\Psi + \varepsilon\Phi = ((\nu_k + \varepsilon\tau_k)_{k=1}^{k_{\max}}, (\mu_m + \varepsilon\sigma_m)_{m=0}^\infty)$ is a collection of (non-negative!) measures that satisfies (1.12) and is therefore admissible in the variational formula in (1.17). That (1.12) is satisfied follows easily from (4.4). Furthermore, the non-negativity follows from the fact that each τ_k and each σ_m is a finite linear combination of measures of the form $\mathbb{1}_A \,\mathrm{dLeb}$ with $A \subset W$. Since only finitely many such summands are involved, there is a constant $C > 0$ such that $|\tau_k| \leq C\nu_k$ and $|\sigma_m| \leq C\mu_m$ for any $k \in [k_{\max}]$ and $m \in \mathbb{N}_0$, and therefore it suffices to take $|\varepsilon| < 1/C$.

From minimality, we deduce that

$$0 = \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \Big(\mathrm{I}(\Psi+\varepsilon\Phi) + \gamma\mathrm{S}(\Psi+\varepsilon\Phi) + \beta\mathrm{M}(\Psi+\varepsilon\Phi) \Big). \tag{4.5}$$

We calculate the latter two terms as

$$\frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \Big(\gamma\mathrm{S}(\Psi+\varepsilon\Phi) + \beta\mathrm{M}(\Psi+\varepsilon\Phi) \Big) = \gamma \sum_{k\in[k_{\max}]} \langle \tau_k, f_k \rangle + \beta \sum_{m\in\mathbb{N}_0} m(m-1)\sigma_m(W),$$

where, as before, we used the notation $\langle \mu, f\rangle$ for the integral of a function f with respect to a measure μ. Abbreviating $M = \sum_{k\in[k_{\max}]} \sum_{l=1}^{k-1} \pi_l \nu_k$ and $M_\tau = \sum_{k\in[k_{\max}]} \sum_{l=1}^{k-1} \pi_l \tau_k$, we see that

$$\begin{aligned} \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \mathrm{I}(\Psi+\varepsilon\Phi) = & \sum_{k\in[k_{\max}]} \Big\langle \tau_k, 1 + \log\frac{\mathrm{d}\nu_k}{\mathrm{d}\mu^{\otimes k}} \Big\rangle + \sum_{m\in\mathbb{N}_0} \Big\langle \sigma_m, 1 + \log\frac{\mathrm{d}\mu_m}{\mathrm{d}\mu} \Big\rangle - \sigma_m(W) \log\frac{(\mathrm{e}\mu(W))^{-m}}{m!} \\ & - \Big\langle M_\tau, 1 + \log\frac{\mathrm{d}M}{\mathrm{d}\mu} \Big\rangle. \end{aligned} \tag{4.6}$$

Summarizing, we obtain from (4.5) that

$$0 = \Big\langle \Phi, \big((h_k)_{k\in[k_{\max}]}, (g_m)_{m\in\mathbb{N}_0}\big)\Big\rangle, \tag{4.7}$$

where

$$h_k = \gamma f_k + 2 - k + \log\frac{\mathrm{d}\nu_k}{\mathrm{d}(\mu\otimes M^{\otimes(k-1)})} \qquad \text{and} \qquad g_m = \beta m(m-1) + 1 + \log\frac{\mathrm{d}\mu_m}{\mathrm{d}\mu} - \log\frac{(\mathrm{e}\mu(W))^{-m}}{m!}.$$

We conceive Φ as an element of the vector space

$$A = \prod_{k\in[k_{\max}]} \mathcal{M}_\pm(W^k) \times \mathcal{M}_\pm(W)^{\mathbb{N}_0}$$

where $\mathcal{M}_\pm$ is the set of signed measures, and $((h_k)_{k\in[k_{\max}]}, (g_m)_{m\in\mathbb{N}_0})$ as a function on $\prod_{k\in[k_{\max}]} W^k \times W^{\mathbb{N}_0}$. The condition in (4.4) means that Φ is perpendicular to any function in

$$\begin{aligned}\mathcal{F} = \Big\{ &((\varphi_k)_{k\in[k_{\max}]}, (\psi_m)_{m\in\mathbb{N}_0})\colon \varphi_k\colon W^k\to\mathbb{R}, \psi_m\colon W\to\mathbb{R} \text{ bounded and measurable for any } k,m,\\ &\exists \widetilde{A},\widetilde{B},\widetilde{C}\colon W\to\mathbb{R}\colon \varphi_k(x_0,\dots,x_{k-1}) = \widetilde{A}(x_0) + \sum_{l=1}^{k-1}\widetilde{C}(x_l),\\ &\text{and } \psi_m(x) = \widetilde{B}(x) - m\widetilde{C}(x) \text{ for } x,x_0,\dots,x_{k-1}\in W\Big\}.\end{aligned}$$

We have shown that, if Φ is perpendicular to any simple function in $\mathcal{F}$, then it is also perpendicular to $((h_k)_{k\in[k_{\max}]}, (g_m)_{m\in\mathbb{N}_0})$. Since $\mathcal{F}$ is a closed linear subspace of A, it follows that it contains this element. That is, there are three functions $\widetilde{A},\widetilde{B},\widetilde{C}$ on W such that, for any k respectively m,

$$h_k(x_0,\dots,x_{k-1}) = \widetilde{A}(x_0) + \sum_{l=1}^{k-1}\widetilde{C}(x_l) \qquad \text{and} \qquad g_m(x) = \widetilde{B}(x) - m\widetilde{C}(x), \qquad x,x_0,\dots,x_{k-1}\in W.$$

Using an obvious substitution, this is equivalent to the existence of three positive functions A,B,C such that

$$\nu_k(\mathrm{d}x_0,\dots,\mathrm{d}x_{k-1}) = \mu(\mathrm{d}x_0)\,A(x_0)\prod_{l=1}^{k-1}\big(C(x_l)M(\mathrm{d}x_l)\big)\mathrm{e}^{-\gamma f_k(x_0,\dots,x_{k-1})}, \qquad k\in[k_{\max}], \tag{4.8}$$

$$\mu_m(\mathrm{d}x) = \mu(\mathrm{d}x)\,B(x)\frac{C(x)^m}{m!}\mathrm{e}^{-\beta m(m-1)}, \qquad m\in\mathbb{N}_0. \tag{4.9}$$

From (i) and (ii) in (1.12), we can identify A and B as

$$\frac{1}{A(x_0)} = \sum_{k\in[k_{\max}]}\int_{W^{k-1}}\prod_{l=1}^{k-1}\big(C(x_l)M(\mathrm{d}x_l)\big)\mathrm{e}^{-\gamma f_k(x_0,\dots,x_{k-1})}, \tag{4.10}$$

$$\frac{1}{B(x)} = \sum_{m\in\mathbb{N}_0}\frac{C(x)^m}{m!}\mathrm{e}^{-\beta m(m-1)}. \tag{4.11}$$

Furthermore, condition (iii) says that

$$\frac{1}{C(x)} = \frac{1}{C(x)}\frac{\mu(\mathrm{d}x)}{M(\mathrm{d}x)}\varphi(C(x)) = \Gamma(C\,\mathrm{d}M, x), \qquad x\in W, \tag{4.12}$$

where $\varphi(\alpha) = \sum_{m\in\mathbb{N}_0} m\frac{\alpha^m}{m!}\mathrm{e}^{-\beta m(m-1)} / \sum_{m\in\mathbb{N}_0} \frac{\alpha^m}{m!}\mathrm{e}^{-\beta m(m-1)}$ for $\alpha \in [0,\infty)$ and

$$\Gamma(\mathrm{d}\widetilde{M}, x) = \int_W \mu(\mathrm{d}x_0)\, \frac{\sum_{k\in[k_{\max}]} \int_{W^{k-2}} \prod_{l=1}^{k-2} \widetilde{M}(\mathrm{d}x_l)\, F_k(x_0, x_1, \dots, x_{k-2}, x)}{\sum_{k\in[k_{\max}]} \int_{W^{k-1}} \prod_{l=1}^{k-1} \widetilde{M}(\mathrm{d}x_l)\, \mathrm{e}^{-\gamma f_k(x_0,\dots,x_{k-1})}}, \tag{4.13}$$

where

$$F_k(x_0, x_1, \dots, x_{k-2}, x) = \sum_{l=1}^{k-1} \mathrm{e}^{-\gamma f_k(x_0, y^l)}, \tag{4.14}$$

y^l is the vector of length $k-1$, consisting of $x_1, \dots, x_{k-2}$; augmented by x at the l-th place, and $\widetilde{M}(\mathrm{d}x) = C(x)M(\mathrm{d}x)$. This ends our derivation of the Euler–Lagrange equations for any minimizer Ψ of (1.17).

This description of C and M is rather implicit and involved, therefore we cannot offer any simple criterion for the uniqueness of the minimizers of (1.17). Also, the question of continuity of the tilting functions A, B and C is open.

Since $\mathrm{I} + \gamma\mathrm{S} + \beta\mathrm{M}$ is convex, it follows that any admissible trajectory setting Ψ satisfying (4.8)–(4.14) is a minimizer of (1.17).

Acknowledgements

The authors thank B. Jahnel, C. Hirsch and M. Renger for interesting discussions and comments.

References

[AK08] S. Adams and W. König, Large deviations for many Brownian bridges with symmetrised initial-terminal condition. *Probab. Theory Relat. Fields* **142:1**, 79–124 (2008).

[BB09] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I: Theory.* Now Publishers Inc. (2009).

[Ch09] X. Chen, *Random Walk Intersections: Large Deviations and Related Topics*. Mathematical Surveys and Monographs, AMS, Vol. 157, Providence, RI. (2009).

[DZ98] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edition, Springer, Berlin (1998).

[GT08] A. J. Ganesh and G. L. Torrisi, Large Deviations of the Interference in a Wireless Communication Model, *IEEE Trans. Inform. Theory* **54:8**, (2008).

[GZ93] H-O. Georgii and H. Zessin, Large deviations and the maximum entropy principle for marked point random fields, *Probab. Theory Relat. Fields* **96:2**, 177–204 (1993).

[HJP16] C. Hirsch, B. Jahnel and R. Patterson, Space-time large deviations in capacity-constrained relay networks, *arXiv:1609.06856* (2016).

[HJKP15] C. HIRSCH, B. JAHNEL, P. KEELER and R. PATTERSON, Large deviations in relay-augmented wireless networks, *arXiv:1510.04146* (2015).

[KB14] H. P. KEELER and B. BLASZCZYSZYN, SINR in wireless networks and the two-parameter Poisson-Dirichlet process, *IEEE Wireless Communications Letters, IEEE*, **3:5**, 525–528 (2014).

[K93] J. F. C. KINGMAN, *Poisson Processes*, Oxford University Press, New York (1993).

[KM02] W. KÖNIG and P. MÖRTERS, Brownian intersection local times: upper tail asymptotics and thick points. *Ann. Probab.* **30**, 1605–1656 (2002).

[KM13] W. KÖNIG and C. MUKHERJEE , Large deviations for Brownian intersection measures, *Commun. Pure Appl. Math.* **66:2**, 263-306 (2013).

[SPW07] H. SONG, M. PENG and W. WANG, Node Selection in Relay-based Cellular Networks, *IEEE 2007 International Symposium on Microwave, Antenna, Propagation, and EMC Technologies For Wireless Communications*.

[T08] J. TRASHORRAS, Large Deviations for Symmetrised Empirical Measures. *J. Theoret. Probab.* **21**, 397–412 (2008).