A SUFFICIENT CONDITION FOR THE QUASIPOTENTIAL TO BE THE RATE FUNCTION OF THE INVARIANT MEASURE OF COUNTABLE-STATE MEAN-FIELD INTERACTING PARTICLE SYSTEMS

SARATH YASODHARAN,* *Brown University*
RAJESH SUNDARESAN,** *Indian Institute of Science*

Abstract

This paper considers the family of invariant measures of Markovian mean-field interacting particle systems on a countably infinite state space and studies its large deviation asymptotics. The Freidlin–Wentzell quasipotential is the usual candidate rate function for the sequence of invariant measures indexed by the number of particles. The paper provides two counterexamples where the quasipotential is not the rate function. The quasipotential arises from finite-horizon considerations. However, there are certain barriers that cannot be surmounted easily in any finite time horizon, but these barriers can be crossed in the stationary regime. Consequently, the quasipotential is infinite at some points where the rate function is finite. After highlighting this phenomenon, the paper studies some sufficient conditions on a class of interacting particle systems under which one can continue to assert that the Freidlin–Wentzell quasipotential is indeed the rate function.

Keywords: Mean-field interaction; invariant measure; large deviations; static large deviation; Freidlin–Wentzell quasipotential; relative entropy

2020 Mathematics Subject Classification: Primary 60F10

Secondary 60K35; 82C22; 60J74; 90B15

1. Introduction

For a broad class of Markov processes, such as small-noise diffusions, finite-state mean-field models, simple exclusion processes, etc., it is well known that the Freidlin–Wentzell quasipotential is the rate function that governs the large deviation principle (LDP) for the family of invariant measures [7, 17, 18, 33]. The quasipotential is the minimum cost (arising from the rate function for a process-level LDP) associated with trajectories of arbitrary but finite duration, with fixed initial and terminal conditions. We begin this paper with two counterexamples of independently evolving countable-state particle systems for which the quasipotential is not the rate function for the family of invariant measures. The family of invariant measures for each of these counterexamples satisfies the LDP with a suitable relative entropy as its rate function, and we show that the quasipotential is not the same as this relative entropy. Specifically,

---

we show that there are points in the state space where the rate function is finite, but the quasipotential is infinite. These points cannot be reached easily via trajectories of arbitrary but finite time duration. The barriers to reaching these points are surmounted in the stationary regime. There are, however, some sufficient conditions, at least on a family of such countable-state interacting particle systems, where the Freidlin–Wentzell quasipotential is indeed the correct rate function; this will be the main result of this paper. Intuitively, the sufficient conditions cut down the speed of outward excursions and ensure that the insurmountable barriers for the finite-horizon trajectories continue to be insurmountable in the stationary regime.

Before we describe the counterexamples and the main result, let us introduce some notation and describe the model of a countable-state mean-field interacting particle system. Let $\mathcal{Z}$ denote the set of non-negative integers, and let $(\mathcal{Z}, \mathcal{E})$ denote a directed graph on $\mathcal{Z}$. Let $\mathcal{M}_1(\mathcal{Z})$ denote the space of probability measures on $\mathcal{Z}$ equipped with the total variation metric (which we denote by d). For each $N \geq 1$, let $\mathcal{M}_1^N(\mathcal{Z}) \subset \mathcal{M}_1(\mathcal{Z})$ denote the set of probability measures on $\mathcal{Z}$ that can arise as empirical measures of N-particle configurations on $\mathcal{Z}^N$. For each $N \geq 1$, we consider a Markov process with the infinitesimal generator acting on functions f on $\mathcal{M}_1^N(\mathcal{Z})$ as follows:

$$\mathcal{L}^N f(\xi) := \sum_{(z,z') \in \mathcal{E}} N \xi(z) \lambda_{z,z'}(\xi) \left[f\left(\xi + \frac{\delta_{z'}}{N} - \frac{\delta_z}{N} \right) - f(\xi) \right], \quad \xi \in \mathcal{M}_1^N(\mathcal{Z}); \qquad (1.1)$$

here $\lambda_{z,z'} : \mathcal{M}_1(\mathcal{Z}) \to \mathbb{R}_+$, $(z, z') \in \mathcal{E}$, are given functions that describe the transition rates, and δ denotes the Dirac measure. Such processes arise as the empirical measures of weakly interacting Markovian mean-field particle systems where the evolution of the state of a particle depends on the states of the other particles only through the empirical measure of the states of all the particles. Under suitable assumptions on the model, the martingale problem for $\mathcal{L}^N$ is well posed and the associated Markov process possesses a unique invariant probability measure $\wp^N$. This paper highlights certain nuances associated with the LDP for the sequence $\{\wp^N, N \geq 1\}$ on $\mathcal{M}_1(\mathcal{Z})$.

Fix $T > 0$ and let $\mu_{\nu_N}^N$ denote the Markov process with initial condition $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$ whose infinitesimal generator is $\mathcal{L}^N$. Its sample paths are elements of $D([0, T], \mathcal{M}_1^N(\mathcal{Z}))$, the space of $\mathcal{M}_1^N(\mathcal{Z})$-valued functions on $[0, T]$ that are right-continuous with left limits equipped with the Skorokhod topology. Such processes have been well studied in the past. Under mild conditions on the transition rates, when $\nu_N \to \nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$, it is well known that the family $\{\mu_{\nu_N}^N, N \geq 1\}$ converges in probability, in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$, as $N \to \infty$ to the *mean-field limit* (see McKean [25] in the context of interacting diffusions and Bordenave *et al.* [6] in the context of countable-state mean-field models):

$$\dot{\mu}(t) = \Lambda_{\mu(t)}^* \mu(t), \quad \mu(0) = \nu, \quad t \in [0, T]. \qquad (1.2)$$

Here $\dot{\mu}(t)$ denotes the derivative of μ at time t; Λ_ξ, $\xi \in \mathcal{M}_1(\mathcal{Z})$, denotes the rate matrix when the empirical measure is ξ (i.e., $\Lambda_\xi(z, z') = \lambda_{z,z'}(\xi)$ when $(z, z') \in \mathcal{E}$, $\Lambda_\xi(z, z') = 0$ when $(z, z') \notin \mathcal{E}$, and $\Lambda_\xi(z, z) = -\sum_{z' \neq z} \lambda_{z,z'}(\xi)$); and Λ_ξ^* denotes the transpose of Λ_ξ. The above dynamical system on $\mathcal{M}_1(\mathcal{Z})$ is called the McKean–Vlasov equation. This mean-field convergence allows one to view the process $\mu_{\nu_N}^N$ as a small random perturbation of the dynamical system (1.2). The starting point of our study of the asymptotics of $\{\wp^N, N \geq 1\}$ is the process-level LDP for $\{\mu_{\nu_N}^N, \nu_N \in \mathcal{M}_1^N(\mathcal{Z}), N \geq 1\}$, whenever ν_N converges to ν in $\mathcal{M}_1(\mathcal{Z})$. This LDP was established by Léonard [21] when the initial conditions are fixed, and by Borkar and Sundaresan [7] when the initial conditions converge in $\mathcal{M}_1(\mathcal{Z})$. The rate function of this LDP
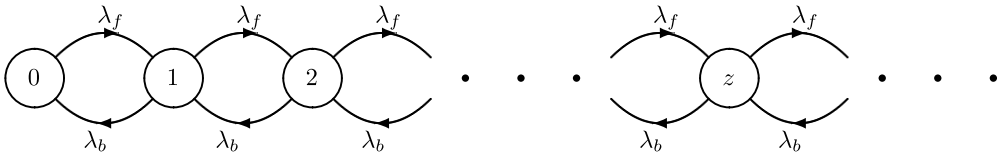
FIGURE 1. Transition rates of an M/M/1 queue.

is governed by 'costs' associated with trajectories on $[0, T]$ with initial condition ν, which we denote by $S_{[0,T]}(\varphi|\nu)$, $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ (see (2.5) for its definition).

We assume that ξ^* is the unique globally asymptotically stable equilibrium of (1.2). Define the Freidlin–Wentzell quasipotential

$$V(\xi) := \inf\{S_{[0,T]}(\varphi|\xi^*) : \varphi(0) = \xi^*, \varphi(T) = \xi, T > 0\}, \quad \xi \in \mathcal{M}_1(\mathcal{Z}). \qquad (1.3)$$

From the theory of large deviations of the invariant measure of Markov processes [7, 11, 18, 33], V is a natural candidate for the rate function of the family $\{\wp^N, N \geq 1\}$.

1.1. Two counterexamples

We begin with two counterexamples for which V is not the rate function for the family of invariant measures.

1.1.1. *Non-interacting M/M/1 queues.* Consider the graph $(\mathcal{Z}, \mathcal{E}_Q)$ whose edge set $\mathcal{E}_Q$ consists of forward edges $\{(z, z + 1), z \in \mathcal{Z}\}$ and backward edges $\{(z, z - 1), z \in \mathcal{Z} \setminus \{0\}\}$ (see Figure 1). Let λ_f and λ_b be two positive numbers. Consider the generator L^Q acting on functions f on $\mathcal{Z}$ by

$$L^Q f(z) := \sum_{z' : (z,z') \in \mathcal{E}_Q} \lambda_{z,z'}(f(z') - f(z)), \quad z \in \mathcal{Z},$$

where $\lambda_{z,z+1} = \lambda_f$ for each $z \in \mathcal{Z}$ and $\lambda_{z,z-1} = \lambda_b$ for each $z \in \mathcal{Z} \setminus \{0\}$. When $\lambda_f < \lambda_b$, the invariant probability measure associated with this Markov process is

$$\xi_Q^*(z) := \left(1 - \frac{\lambda_f}{\lambda_b}\right)\left(\frac{\lambda_f}{\lambda_b}\right)^z, \quad z \in \mathcal{Z}.$$

For each $N \geq 1$, we consider N particles, each of which evolves independently as a Markov process on $\mathcal{Z}$ with the infinitesimal generator L^Q. That is, the particles are independent M/M/1 queues. It is easy to check that the empirical measure of the system of particles is also a Markov process on the state space $\mathcal{M}_1^N(\mathcal{Z})$ and it possesses a unique invariant probability measure, which we denote by $\wp_Q^N$.

On one hand, it is straightforward to see that the family $\{\wp_Q^N, N \geq 1\}$ satisfies the LDP on $\mathcal{M}_1(\mathcal{Z})$. Indeed, under stationarity, the state of each particle is distributed as ξ_Q^*. As a consequence, $\wp_Q^N$ is the law of the random variable $\frac{1}{N} \sum_{n=1}^N \delta_{\zeta_n}$ on $\mathcal{M}_1(\mathcal{Z})$, where $\zeta_1, \ldots, \zeta_N$ are independent and identically distributed (i.i.d.) as ξ_Q^*. Therefore, by Sanov's theorem [13, Theorem 6.2.10], $\{\wp_Q^N, N \geq 1\}$ satisfies the LDP with the rate function $I(\cdot\|\xi_Q^*)$, where

$I : \mathcal{M}_1(\mathcal{Z}) \times \mathcal{M}_1(\mathcal{Z}) \to [0, \infty]$ is the relative entropy defined by

$$I(\zeta \| \nu) := \begin{cases} \sum_{z \in \mathcal{Z}} \zeta(z) \log \left(\dfrac{\zeta(z)}{\nu(z)} \right) & \text{if } \zeta \ll \nu, \\ \infty & \text{otherwise,} \end{cases} \tag{1.4}$$

with the convention that $0 \log 0 = 0$. On the other hand, it is natural to conjecture that the rate function for the family $\{\wp_Q^N, N \geq 1\}$ is given by the quasipotential (1.3) with ξ^* replaced by ξ_Q^*. However, as discussed in the next paragraph, the quasipotential is not the same as $I(\cdot \| \xi_Q^*)$. Hence, from the uniqueness of the large deviations rate function [13, Lemma 4.1.4], the quasipotential does not govern the rate function for the family $\{\wp_Q^N, N \geq 1\}$.

We now provide some intuition on why the quasipotential is not the rate function in the example under consideration. For a formal proof, see Section 8. We first introduce some notation. Let $\mathbb{R}^\infty$ denote the infinite product of $\mathbb{R}$ equipped with the product topology. We view $\mathcal{M}_1(\mathcal{Z})$ as the subset $\{x \in \mathbb{R}^\infty : x_i \geq 0 \, \forall i, \sum_{i \geq 0} x_i = 1\}$ of $\mathbb{R}^\infty$ with the subspace topology (e.g., see [15, Chapter 3, Section 2]). If $\xi, f \in \mathbb{R}^\infty$, we define

$$\langle \xi, f \rangle := \lim_{m \to \infty} \sum_{z=0}^{m} \xi(z) f(z), \tag{1.5}$$

whenever the limit exists. Also, define $\vartheta : \mathcal{Z} \to \mathbb{R}_+$ by

$$\vartheta(z) := z \log z, \quad z \in \mathcal{Z}, \tag{1.6}$$

with the convention that $0 \log 0 = 0$, and define $\iota(z) := z$, $z \in \mathcal{Z}$. Using the fact that ξ_Q^* has geometric decay, it can be checked that $I(\xi \| \xi_Q^*)$ is finite if and only if the first moment of ξ (i.e., $\langle \xi, \iota \rangle$) is finite. However, it turns out that $V(\xi)$ (i.e., the quantity in (1.3) with ξ^* replaced by ξ_Q^*) is finite if and only if the ϑ-moment of ξ (i.e., $\langle \xi, \vartheta \rangle$) is finite. In particular, if we consider a $\xi \in \mathcal{M}_1(\mathcal{Z})$ whose first moment is finite but whose ϑ-moment is infinite, then $V(\xi) \neq I(\xi \| \xi_Q^*)$. Let $\varepsilon > 0$, $\xi \in \mathcal{M}_1(\mathcal{Z})$ be such that $\langle \xi, \iota \rangle < \infty$ but $\langle \xi, \vartheta \rangle = \infty$, and consider the ε-neighbourhood of ξ in $\mathcal{M}_1(\mathcal{Z})$. By Sanov's theorem, the probability of this neighbourhood under $\wp_Q^N$ is of the form $\exp\{-N(I(\xi \| \xi_Q^*) + o(1))\}$. For a fixed $T > 0$, let us now try to estimate the probability of $\mu_{\nu_N}^N(T)$ being in this neighbourhood when ν_N is in a small neighbourhood of ξ_Q^*. If the process μ^N is initiated at a ν_N near ξ_Q^*, then the probability that the random variable $\mu_{\nu_N}^N(T)$ is in the ε-neighbourhood of ξ is at most

$$\exp \left\{ -N \left(\inf_{\{\xi' : d(\xi, \xi') \leq \varepsilon\}} V(\xi') + o(1) \right) \right\}.$$

Since V is lower semicontinuous (we prove this in Lemma 5.4), we must have

$$\inf_{\{\xi' : d(\xi, \xi') \leq \varepsilon\}} V(\xi') \to \infty \text{ as } \varepsilon \to 0.$$

Hence we can choose an ε small enough so that $\inf_{\{\xi' : d(\xi, \xi') \leq \varepsilon\}} V(\xi') > 2I(\xi \| \xi_Q^*)$. For this ε, the probability that $\mu_{\nu_N}^N(T)$ lies is the ε-neighbourhood of ξ is bounded above by $\exp\{-N \times (2I(\xi \| \xi_Q^*) + o(1))\}$, which is smaller than $\exp\{-N(I(\xi \| \xi_Q^*) + o(1))\}$, even in the exponential
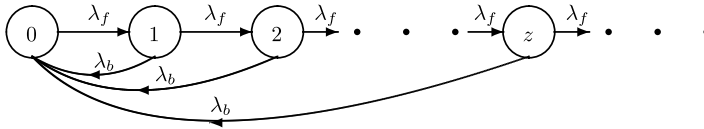
FIGURE 2. Transition rates of a wireless node.

scale, for large enough N. That is, for any arbitrary but fixed T, we can find a small neighbourhood of ξ such that the probability that $\mu^N_{\nu_N}(T)$ lies in that neighbourhood is smaller than what we expect to see in the stationary regime. In other words, there are some barriers in $\mathcal{M}_1(\mathcal{Z})$ that cannot be surmounted in any finite time, but that can be crossed in the stationary regime. These barriers indicate that, to obtain the correct stationary-regime probability of a small neighbourhood of ξ using the dynamics of $\mu^N_{\nu_N}$, one should wait longer than any fixed time horizon. That is, one should consider the random variable $\mu^N_{\nu_N}(T(N))$, where $T(N)$ is a suitable function of N, and estimate the probability that $\mu^N_{\nu_N}(T(N))$ belongs to a small neighbourhood of ξ. However, it is not straightforward to obtain such estimates from the process-level large deviation estimates of $\mu^N_{\nu_N}$, since the latter are usually available for a fixed time duration.

There are natural barriers in the context of finite-state mean-field models when the limiting dynamical system has multiple (but finitely many) stable equilibria [36]. In such situations, passages from a neighbourhood of one equilibrium to a neighbourhood of another take place over time durations of the form $\exp\{N \times O(1)\}$ where N is the number of particles (here, $O(1)$ refers to a bounded sequence). Interestingly, these barriers can be surmounted using trajectories of finite time durations; i.e., for any fixed T, the probability that the empirical measure process reaches a neighbourhood of an equilibrium at time T when it is initiated in a small neighbourhood of another equilibrium is of the form $\exp\{-N \times O(1)\}$. In contrast, in the case of the above counterexample, the barriers cannot be surmounted in finite time durations; for any fixed T, the probability that $\mu^N(T)$ reaches a small neighbourhood of a point in $\mathcal{M}_1(\mathcal{Z})$ with finite first moment but infinite ϑ-moment when it is initiated from a neighbourhood of ξ^*_Q is of the form $\exp\{-N \times \omega(1)\}$ (here, $\omega(1)$ refers to a sequence that goes to ∞). Hence we anticipate that the barriers that we encounter in the above counterexample are somehow more difficult to surmount than those that arise in the case of finite-state mean-field models with multiple stable equilibria.

1.1.2. *Non-interacting nodes in a wireless network.* We provide another counterexample where the issue is similar. Consider the graph $(\mathcal{Z}, \mathcal{E}_W)$ whose edge set $\mathcal{E}_W$ consists of forward edges $\{(z, z+1), z \in \mathcal{Z}\}$ and backward edges $\{(z, 0), z \in \mathcal{Z} \setminus \{0\}\}$ (see Figure 2). Let λ_f and λ_b be positive numbers. Consider the generator L^W acting on functions f on $\mathcal{Z}$ by

$$L^W f(z) := \sum_{z' : (z,z') \in \mathcal{E}_W} \lambda_{z,z'}(f(z') - f(z)), \quad z \in \mathcal{Z},$$

where $\lambda_{z,z+1} = \lambda_f$ for each $z \in \mathcal{Z}$ and $\lambda_{z,0} = \lambda_b$ for each $z \in \mathcal{Z} \setminus \{0\}$. The invariant probability measure associated with this Markov process is

$$\xi^*_W(z) := \frac{\lambda_b}{\lambda_f + \lambda_b} \left(\frac{\lambda_f}{\lambda_f + \lambda_b} \right)^z, \quad z \in \mathcal{Z}.$$

Similarly to the previous example, for each $N \geq 1$, we consider N particles, each of which evolves independently as a Markov process on $\mathcal{Z}$ with the infinitesimal generator L^W. It is

easy to check that the empirical measure of the system of particles possesses a unique invariant probability measure, which we denote by $\wp_W^N$. Under stationarity, the state of each particle is distributed as ξ_W^*. As a consequence, $\wp_W^N$ is the law of the random variable $\frac{1}{N}\sum_{n=1}^N \delta_{\zeta_n}$ on $\mathcal{M}_1(\mathcal{Z})$, where $\zeta_1, \ldots, \zeta_N$ are i.i.d. ξ_W^*. Hence, by Sanov's theorem, the family $\{\wp_W^N, N \geq 1\}$ satisfies the LDP with the rate function $I(\cdot \| \xi_W^*)$. As we show in Section 8, in this example too, the quasipotential (1.3) with ξ^* replaced by ξ_W^* is not the same as $I(\cdot \| \xi_W^*)$. As in the previous example, there are points ξ where $V(\xi) = \infty$ but $I(\xi \| \xi_Q^*) < \infty$, points ξ that have a finite first moment but infinite ϑ-moment. Once again, the quasipotential does not govern the rate function for the family $\{\wp_W^N, N \geq 1\}$.

1.2. Assumptions and main result

We now provide some assumptions on the model of countable-state mean-field interacting particle systems that ensure that the barriers in $\mathcal{M}_1(\mathcal{Z})$ that are insurmountable using trajectories of arbitrary but finite time duration remain insurmountable in the stationary regime as well. Under these assumptions, we prove the main result of this paper, namely, that the sequence of invariant measures $\{\wp^N, N \geq 1\}$ satisfies the LDP with rate function V.

1.2.1. *Assumptions.* Our first set of assumptions is on the mean-field interacting particle system (i.e., on the generator $\mathscr{L}^N$ defined in (1.1)).

(A1) The edge set is given by $\mathcal{E} = \{(z, z+1), z \in \mathcal{Z}\} \cup \{(z, 0), z \in \mathcal{Z} \setminus \{0\}\}$.

(A2) There exist positive constants $\overline{\lambda}$ and $\underline{\lambda}$ such that

$$\frac{\underline{\lambda}}{z+1} \leq \lambda_{z,z+1}(\xi) \leq \frac{\overline{\lambda}}{z+1} \quad \text{and} \quad \underline{\lambda} \leq \lambda_{z,0}(\xi) \leq \overline{\lambda}$$

for all $\xi \in \mathcal{M}_1(\mathcal{Z})$.

(A3) The functions $(z+1)\lambda_{z,z+1}(\,\cdot\,)$, $z \in \mathcal{Z}$, and $\lambda_{z,0}(\,\cdot\,)$, $z \in \mathcal{Z} \setminus \{0\}$, are uniformly Lipschitz continuous on $\mathcal{M}_1(\mathcal{Z})$.

Note that Assumption (A1) considers a specific transition graph (Figure 2) for each particle. This graph arises in the contexts of random backoff algorithms for medium access in wireless local area networks [20] and decentralised control of loads in a smart grid [26]. Assumption (A2) ensures that the forward transition rate at state z decays as $1/z$. This key assumption cuts down the speed of outward excursions and enables us to overcome the issue described in the counterexamples. To highlight this, consider a modified example of Section 1.1.2 where $\lambda_{z,z+1} = \lambda_f/(z+1)$, $z \in \mathcal{Z}$; the rest of the description remains the same. Let $\tilde{\xi}_W \in \mathcal{M}_1(\mathcal{Z})$ denote the invariant probability measure associated with one particle. It can be checked that $\tilde{\xi}_W(z)$ is of the order of $\exp\{-\vartheta(z)\}$, unlike ξ_W^*, which has geometric decay. As a consequence, $I(\xi \| \tilde{\xi}_W)$ is finite if and only if the ϑ-moment of ξ is finite. Hence, by imposing (A2), we have ensured that the barriers in $\mathcal{M}_1(\mathcal{Z})$ that are insurmountable for finite-time-duration trajectories continue to remain insurmountable in the stationary regime; this is the key property that enables us to prove the main result of this paper. Assumption (A3) is a uniform Lipschitz continuity property for the transition rates which is required for the process-level LDP for $\mu_{\nu_N}^N$ to hold and for the McKean–Vlasov equation (1.2) to be well posed.

Our second set of assumptions is on the McKean–Vlasov equation (1.2). Let μ_ν, $\nu \in \mathcal{M}_1(\mathcal{Z})$, denote the solution to the limiting dynamics (1.2) with initial condition $\nu \in \mathcal{M}_1(\mathcal{Z})$. Recall the function ϑ. Define $\mathscr{K}_M := \{\xi \in \mathcal{M}_1(\mathcal{Z}) : \langle \xi, \vartheta \rangle \leq M\}$, $M > 0$.

(B1)   There exists a unique globally asymptotically stable equilibrium ξ^* for the McKean–Vlasov equation (1.2).

(B2)   We have $\langle \xi^*, \vartheta \rangle < \infty$ and $\lim_{t \to \infty} \sup_{\nu \in \mathscr{K}_M} \langle \mu_\nu(t), \vartheta \rangle = \langle \xi^*, \vartheta \rangle$ for each $M > 0$.

The first assumption above asserts that all the trajectories of (1.2) converge to ξ^* as time becomes large. The proof of the LDP upper and lower bounds for the family $\{\wp^N, N \geq 1\}$ involves construction of trajectories that start at suitable compact sets, reach the stable equilibrium ξ^* using arbitrarily small cost, and then terminate at a desired point in $\mathcal{M}_1(\mathcal{Z})$ starting from ξ^*. All of these are enabled by Assumption (B1) (see more remarks about this assumption in Section 1.3). The second assumption asserts that the ϑ-moment of the solution to the limiting dynamics converges uniformly over initial conditions lying in sets of bounded ϑ-moment. In the case of a non-interacting system that satisfies (A1) but with constant forward transition rates (for example, see L^W in Section 1.1.2), the analogue of this assumption can easily be verified: the first moment of the solution to the limiting dynamics converges uniformly over initial conditions lying in sets of bounded first moment. In fact, one can explicitly write down the first moment of the solution to the limiting dynamics in this case and verify this assumption easily. Assumption (B2) is the analogous statement for our mean-field system that satisfies the $1/z$-decay of the forward transition rates in Assumption (A2).

1.2.2. *Main result.* We now state the main result of this paper, namely the LDP for the family of invariant measures $\{\wp^N, N \geq 1\}$ under Assumptions (A1)–(A3) and (B1)–(B2).

We first assert the existence and uniqueness of the invariant measure $\wp^N$ for $\mathscr{L}^N$ for each $N \geq 1$, and the exponential tightness of the family $\{\wp^N, N \geq 1\}$.

Proposition 1.1. *Assume* (A1) *and* (A2). *For each $N \geq 1$, $\mathscr{L}^N$ admits a unique invariant probability measure $\wp^N$. Furthermore, the family $\{\wp^N, N \geq 1\}$ is exponentially tight in $\mathcal{M}_1(\mathcal{Z})$.*

Recall the quasipotential V defined in (1.3). We now state the main result of this paper.

Theorem 1.1. *Assume* (A1), (A2), (A3), (B1), *and* (B2). *Then the family of probability measures $\{\wp^N, N \geq 1\}$ satisfies the LDP on $\mathcal{M}_1(\mathcal{Z})$ with rate function V.*

The proof of this result is carried out in Sections 4–7. We begin with the process-level uniform LDP for $\mu_{\nu_N}^N$ over compact subsets of $\mathcal{M}_1(\mathcal{Z})$; this uniform LDP gives us the large deviation estimates for the process $\mu_{\nu_N}^N$ uniformly over the initial conditions ν_N lying in a given compact set (see Definition 2.2 and Theorem 2.1). We prove the LDP for the family $\{\wp^N, N \geq 1\}$ by transferring this process-level uniform LDP for $\mu_{\nu_N}^N$ over compact subsets of $\mathcal{M}_1(\mathcal{Z})$ to the stationary regime. The proof of the LDP lower bound (in Section 4) considers specific trajectories and lower-bounds the probability of small neighbourhoods of points in $\mathcal{M}_1(\mathcal{Z})$ under $\wp^N$ using the probability that the process $\mu_{\nu_N}^N$ remains close to these trajectories. For the proof of the upper bound, we require certain regularity properties of the quasipotential. These properties are established in Section 5. We first show a controllability property (this terminology is from Cerrai and Röckner [11]) for V: $V(\xi)$ is finite if and only if $\langle \xi, \vartheta \rangle < \infty$. Using the lower bound proved in Section 4, we then show that the level sets of V are compact subsets of $\mathcal{M}_1(\mathcal{Z})$. Since $\mathcal{M}_1(\mathcal{Z})$ is not locally compact and V has compact lower level sets, we do not expect V to be continuous on $\mathcal{M}_1(\mathcal{Z})$. Indeed, if $\xi \in \mathcal{M}_1(\mathcal{Z})$ is such that V is continuous at ξ and $V(\xi) < \infty$, given $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(\xi', \xi) < \delta$ implies that $|V(\xi') - V(\xi)| < \varepsilon$. In particular, $\{\xi' \in \mathcal{M}_1(\mathcal{Z}) : V(\xi') \leq V(\xi) + \varepsilon\} \supset B(\xi, \delta)$. Since $\{\xi' \in \mathcal{M}_1(\mathcal{Z}) : V(\xi') \leq V(\xi) + \varepsilon\}$ is compact in $\mathcal{M}_1(\mathcal{Z})$, this shows that ξ has a relatively compact neighbourhood in $\mathcal{M}_1(\mathcal{Z})$, which is a contradiction. This shows that, for any $\xi \in \mathcal{M}_1(\mathcal{Z})$ such

that $V(\xi) < \infty$, V is discontinuous at ξ. However, we prove the following *small-cost connection* property: whenever $\xi_n \to \xi^*$ in $\mathcal{M}_1(\mathcal{Z})$ and $\langle \xi_n, \vartheta \rangle \to \langle \xi^*, \vartheta \rangle$ as $n \to \infty$, we have $\lim_{n \to \infty} V(\xi_n) = V(\xi^*) = 0$. These properties of the quasipotential are then used to transfer the process-level uniform LDP upper bound for $\mu_{\nu_N}^N$ (uniform over compact subsets of $\mathcal{M}_1(\mathcal{Z})$) to the LDP upper bound for the family of invariant measures. The proof of the upper bound is carried out in Section 6. Finally, we complete the proof of the theorem in Section 7.

While the proofs of our lower and upper bounds follow the general methodology of Sowers [33], there are significant model-specific difficulties that arise in our context. The main novelty in the proof of Theorem 1.1 is to establish the small-cost connection property of the quasipotential V under Assumptions (A1)–(A3) and (B1)–(B2). That is, we can find trajectories of small cost that start at ξ^* and end at points in $\mathcal{M}_1(\mathcal{Z})$ whose ϑ-moment is not very far from that of ξ^*. In the work of Sowers [33], this is carried out by considering the 'straight-line' trajectory that connects the attractor to the nearby point under consideration. Such a trajectory may not have small cost in our case since the mass transfer is restricted to the edges in $\mathcal{E}$. We overcome this difficulty by considering a mass transfer of piecewise constant velocity via the edges in $\mathcal{E}$. We then carefully estimate the cost of this trajectory and prove the necessary small-cost connection property. We also simplify the proof of the compactness of the lower level sets of V; while Sowers [34, Proposition 7] studies the minimisation of the costs of trajectories over the infinite horizon, we arrive at it by using the LDP lower bound and the exponential tightness of the family $\{\wp^N, N \geq 1\}$. We also remark that the methodology of Sowers [33] has been used by Cerrai and Röckner [11] in the context of stochastic reaction diffusion equations and by Cerrai and Paskal [9] in the context of two-dimensional stochastic Navier–Stokes equations.

1.3. Discussion and future directions

The main result and the counterexamples suggest that in order for the family of invariant measures of a Markov process to satisfy the LDP with rate function governed by the Freidlin–Wentzell quasipotential, one must have some *good* properties on the model under consideration. In the case of our main result, this goodness is achieved by the $1/z$-decay of the forward transition rates from Assumption (A2). We use this assumption to show the exponential tightness of the invariant measure over compact subsets with bounded ϑ-moments. It also enables us to show the regularity properties of the quasipotential which are required to transfer the process-level large deviation result to the stationary regime. However, a general treatment of the LDP for the family of invariant measures of a Markov process (that encompasses the cases of [7, 9, 11, 17, 33]), especially when the ambient state space is not locally compact, is missing from the literature.

One of the assumptions that plays a significant role in the proof of our main result is the existence of a unique globally asymptotically stable equilibrium for the limiting dynamics (Assumption (B1)). In the works of Sowers [33], Cerrai and Röckner [11], and Cerrai and Paskal [9], the model assumptions ensure that (B1) holds. In general, the limiting dynamical system (1.2) could possess multiple ω-limit sets. In that case, the approach of our proofs breaks down. A well-known approach to studying large deviations of the invariant measures in such cases is to focus on small neighbourhoods of these ω-limit sets and then analyse the discrete-time Markov chain that evolves on these neighbourhoods. The LDP then follows from the estimates of the invariant measure of this discrete-time chain (see Freidlin and Wentzell [18, Chapter 6, Section 4]). However, this approach requires the *uniform LDP over open subsets* of $\mathcal{M}_1(\mathcal{Z})$, which is not yet available for our mean-field model. If this can be established, along with the regularity properties of the quasipotential established in Section 5, one can use the

above idea not only to extend our main result to the case when the limiting dynamical system possesses multiple ω-limit sets, but also to study exit problems and metastability phenomena in our mean-field model.

Another definition of the quasipotential appears in the literature. It is given by the minimisation of costs of the form $S_{(-\infty,0]}(\varphi)$ over infinite-horizon trajectories φ on $(-\infty, 0]$ such that the terminal condition $\varphi(0)$ is fixed and $\varphi(t) \to \xi^*$ as $t \to -\infty$ (see Sowers [33], Cerrai and Röckner [11]). While it is clear that the above definition of the quasipotential is a lower bound for V in (1.3), unlike in Sowers [33] and Cerrai and Röckner [11], we are not able to show that the two definitions are the same. A proof of this equality, or otherwise, would add more insight on the general case.

We remark that Assumption (A3) does not play a role in the proof of our main result. It is used to invoke the process-level LDP for $\mu_{\nu_N}^N$ (see Theorem 2.1) and the well-posedness of the limiting dynamical system (1.2). If these two properties are established through some other means, then the proof of Theorem 1.1 will hold verbatim without the need for Assumption (A3).

Finally, we mention that a time-independent variational formula for the quasipotential is available for some non-reversible models in statistical mechanics; see Bertini *et al.* [2, 3]. It is not clear whether the quasipotential V in (1.3) admits a time-independent variational form. This would be an interesting direction to explore.

1.4. Related literature

Process-level large deviations of small-noise diffusion processes have been well studied in the past. For finite-dimensional large deviation problems, see Freidlin and Wentzell [18, Chapter 5], Liptser [23], Veretennikov [35], Puhalskii [29], and the references therein. For infinite-dimensional problems where the state space is not locally compact, see Sowers [34] and Cerrai and Röckner [10]. More recently, uniform large deviation principles (uniform LDPs) for Banach-space-valued stochastic differential equations over the class of bounded and open subsets of the Banach space have been studied by Salins *et al.* [31]. These have been used to study the exit times and metastability in such processes; see Salins and Spiliopoulos [32]. While the above works focus on diffusion processes, our work focuses on the stationary-regime large deviations of countable-state mean-field models with jumps. In the spirit of the small-noise problems listed above, our process $\mu_{\nu_N}^N$ can be viewed as a small random perturbation of the dynamical system (1.2) on $\mathcal{M}_1(\mathcal{Z})$.

In the context of interacting particle systems, Dawson and Gärtner [12] established the process-level LDP for weakly interacting diffusion processes, and Léonard [21] and Borkar and Sundaresan [7] extended this to mean-field interacting particle systems with jumps. In this work, we focus on the stationary-regime large deviations of mean-field models with jumps when the state of each particle comes from a countable set. For small-noise diffusion process on Euclidean spaces and finite-state mean-field models, since the state space (on which the empirical measure process evolves) is locally compact, the process-level large deviation results have been extended in a straightforward manner to the uniform LDP over the class of open subsets of the space. Such uniform large deviation estimates have been used to prove the large deviations of the invariant measure and the exit time estimates; see Freidlin and Wentzell [18, Chapter 6] in the context of diffusion processes, and Borkar and Sundaresan [7, 36] in the context of finite-state mean-field models. One of the key ingredients in these proofs is the continuity of the quasipotential. However, in our case, the state space $\mathcal{M}_1(\mathcal{Z})$ is infinite-dimensional and not locally compact. Therefore, since the quasipotential (1.3) is expected to

have compact lower level sets, we do not expect it to be continuous on $\mathcal{M}_1(\mathcal{Z})$, unlike in the finite-dimensional problems mentioned above. Hence the ideas presented in [7] are not directly applicable to our context of the LDP for the family of invariant measures.

Large deviations of the family of invariant measures for small-noise diffusion processes on non-locally-compact spaces have also been studied in the past; see Sowers [33] and Cerrai and Röckner [11]. They have a unique attractor for the limiting dynamics, and the proof essentially involves conversion of the uniform LDP over the finite time horizon to the stationary regime. Martirosyan [24] studied a situation where the limiting dynamical system possesses multiple attractors. For the study of large deviations of the family of invariant measures for simple exclusion processes, see Bodineau and Giacomin [5] and Bertini *et al.* [3]. More recently, Farfán *et al.* [17] extended this to a simple exclusion process whose limiting hydrodynamic equation has multiple attractors. Their proof proceeds similarly to the case of finite-dimensional diffusions in Freidlin and Wentzell [18, Chapter 6, Section 4], by first approximating the process near the attractors and then using the Khasminskii reconstruction formula [19, Chapter 4, Section 4]. In particular, it requires the uniform LDP to hold over open subsets of the state space. Since their state space, although infinite-dimensional, is compact, the proof of the uniform LDP over open subsets easily follows from the process-level LDP. Also, the compactness of the state space simplifies the proof of the small-cost connection property from the attractors to nearby points, a property needed in the Khasminskii reconstruction. Although we restrict our attention to the case of a unique globally asymptotically stable equilibrium as in [11, 33], the main novelty of our work is that we establish certain regularity properties of the quasipotential for countable-state mean-field models with jumps which have not previously been established. We then use these properties to prove the LDP for the family of invariant measures. Furthermore, we exhibit two counterexamples in which the stationary-regime LDP's rate functions are not governed by the usual quasipotential. To the best of our knowledge, such examples, in which the LDP for the family of invariant measures holds but its rate functions are not governed by the usual Freidlin–Wentzell quasipotential, are new. These examples are constructed in such a way that the particle systems do not possess the *small-cost connection* property from the attractor to nearby points with finite first moment but infinite ϑ-moment.

Puhalskii [28] studied large deviations of the family of invariant measures for a queueing network in a finite-dimensional setting. In addition, Puhalskii [30] studied large deviations of the family of invariant measures for a stochastic process under some general conditions. One of the conditions in [30] is the small-cost connection property between any two nearby points in the state space, which we do not expect to hold in our countable-state mean-field model since our state space is infinite-dimensional.

1.5. Organisation

This paper is organised as follows. In Section 2, we provide preliminary results on the large deviations over finite time horizons. The proof of the main result is carried out in Sections 3–7. In Section 3, we prove the existence, uniqueness, and exponential tightness of the family of invariant measures. In Section 4, we prove the LDP lower bound for the family of invariant measures. In Section 5, we establish some regularity properties of the quasipotential V defined in (1.3). In Section 6, we prove the LDP upper bound for the family of invariant measures. In Section 7, we complete the proof of the main result. Finally in Section 8, we prove that the quasipotential differs from the relative entropy (with respect to the globally asymptotically stable equilibrium) for the two counterexamples discussed in Section 1.1.

2. Preliminaries

2.1. Frequently used notation

We first summarise the notation used throughout the paper. Let $\mathcal{Z}$ denote the set of nonnegative integers, and let $(\mathcal{Z}, \mathcal{E})$ denote a directed graph on $\mathcal{Z}$. Let $\mathbb{R}^\infty$ denote the infinite product of $\mathbb{R}$ equipped with the topology of pointwise convergence. Let $C_0(\mathcal{Z})$ denote the space of functions on $\mathcal{Z}$ with compact support. Recall that $\mathcal{M}_1(\mathcal{Z})$ denotes the space of probability measures on $\mathcal{Z}$ equipped with the total variation metric (denoted by d). This metric generates the topology of weak convergence on $\mathcal{M}_1(\mathcal{Z})$. By Scheffé's lemma [15, Chapter 3, Section 2], $\mathcal{M}_1(\mathcal{Z})$ can be identified with the subset $\{x \in \mathbb{R}^\infty : x_i \geq 0 \, \forall i, \, \sum_{i \geq 0} x_i = 1\}$ of $\mathbb{R}^\infty$ with the subspace topology. For each $N \geq 1$, recall that $\mathcal{M}_1^N(\mathcal{Z}) \subset \mathcal{M}_1(\mathcal{Z})$ denotes the space of probability measures on $\mathcal{Z}$ that can arise as empirical measures of N-particle configurations on $\mathcal{Z}^N$. Recall ϑ as defined in (1.6). Given $\alpha \in C_0(\mathcal{Z})$ and $g \in \mathbb{R}^\infty$, let the bracket $\langle \alpha, g \rangle$ denote $\sum_{z \in \mathcal{Z}} \alpha(z) g(z)$. Similarly, given $f, g \in \mathbb{R}^\infty$, let the bracket $\langle f, g \rangle$ denote $\lim_{n \to \infty} \sum_{k=0}^n f(k) g(k)$, whenever the limit exists. For $M > 0$, define

$$\mathcal{K}_M := \{\xi \in \mathcal{M}_1(\mathcal{Z}) : \langle \xi, \vartheta \rangle \leq M\}; \tag{2.1}$$

by Prokhorov's theorem, $\mathcal{K}_M$ is a compact subset of $\mathcal{M}_1(\mathcal{Z})$. Define $\mathcal{K} := \bigcup_{M \geq 1} \mathcal{K}_M$. Let $\xi^* \in \mathcal{M}_1(\mathcal{Z})$ denote the globally asymptotically stable equilibrium for the McKean–Vlasov equation (1.2) (see Assumption (B1)). For each $\Delta > 0$, define

$$K(\Delta) := \{\xi \in \mathcal{K} : d(\xi^*, \xi) \leq \Delta \text{ and } |\langle \xi^*, \vartheta \rangle - \langle \xi, \vartheta \rangle| \leq \Delta\}; \tag{2.2}$$

note that $K(\Delta)$ depends on ξ^* as well (which we omit from the notation, for ease of reading). Define

$$\tau(u) := e^u - u - 1, \qquad u \in \mathbb{R}. \tag{2.3}$$

Note that τ is the log-moment generating function of the centred unit-rate Poisson law, and define its convex dual by

$$\tau^*(u) := \begin{cases} \infty & \text{if } u < -1, \\ 1 & \text{if } u = -1, \\ (u+1)\log(u+1) - u & \text{if } u > -1. \end{cases} \tag{2.4}$$

For a complete and separable metric space $(\mathcal{S}, d_0)$, $A \subset \mathcal{S}$, and $x \in \mathcal{S}$, let $d_0(x, A)$ denote $\inf_{y \in A} d_0(x, y)$. For a set A, let $\sim A$ denote the complement of A. For two numbers a and b, let $a \vee b$ (resp. $a \wedge b$) denote the maximum (resp. minimum) of a and b. Also, let $a^+ = \max\{a, 0\}$. For a metric space $\mathcal{S}$, let $\mathcal{B}(\mathcal{S})$ denote the Borel σ-field on $\mathcal{S}$. Finally, C denotes a constant, and its value may be different at each occurrence.

2.1.1. *Notation related to the dynamics.* Let $D([0, T], \mathcal{S})$ denote the space of $\mathcal{S}$-valued functions on $[0, T]$ that are right-continuous with left limits. It is equipped with the Skorokhod topology, which makes it a complete and separable metric space (see, e.g., Ethier and Kurtz [16, Chapter 3]). Let ρ denote a metric on $D([0, T], \mathcal{S})$ that generates the Skorokhod topology. An element of $D([0, T], \mathcal{S})$ is called a 'trajectory', and we shall refer to the process-level large deviations rate function evaluated on a trajectory as the 'cost' associated with that trajectory. For a trajectory φ, let both φ_t and $\varphi(t)$ denote the evaluation of φ at time t. For $N \geq 1$ and

$\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$, let $\mathbb{P}_{\nu_N}^N$ denote the solution to the $D([0, T], \mathcal{M}_1^N(\mathcal{Z}))$-valued martingale problem for $\mathscr{L}^N$ with initial condition $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$ (whenever the martingale problem for $\mathscr{L}^N$ is well posed). Let $\mu_{\nu_N}^N$ denote the random element of $D([0, T], \mathcal{M}_1^N(\mathcal{Z}))$ whose law is $\mathbb{P}_{\nu_N}^N$. For each $\xi \in \mathcal{M}_1(\mathcal{Z})$, let L_ξ denote the generator acting on functions f on $\mathcal{Z}$ by

$$f \mapsto L_\xi(z) := \sum_{z' : (z, z') \in \mathcal{E}} \lambda_{z, z'}(\xi)(f(z') - f(z)), \qquad z \in \mathcal{Z},$$

i.e., the generator of the single particle evolving on $\mathcal{Z}$ under the static mean-field ξ.

Let $C_0^1([0, T] \times \mathcal{Z})$ denote the space of real-valued functions on $[0, T] \times \mathcal{Z}$ with compact support that are continuously differentiable in the first argument. Given a trajectory $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ such that the mapping $[0, T] \ni t \mapsto \varphi_t \in \mathcal{M}_1(\mathcal{Z})$ is absolutely continuous (see Dawson and Gärtner [12, Section 4.1]), one can define $\dot{\varphi}_t \in \mathbb{R}^\infty$ for almost all $t \in [0, T]$ such that

$$\langle \varphi_t, f_t \rangle = \langle \varphi_0, f_0 \rangle + \int_{[0,t]} \langle \dot{\varphi}_u, f_u \rangle du + \int_{[0,t]} \langle \varphi_u, \partial_u f_u \rangle du$$

holds for each $f \in C_0^1([0, T] \times \mathcal{Z})$ and $t \in [0, T]$.

Finally, let $\mathcal{M}_1(D([0, T], \mathcal{Z}))$ denote the space of probability measures on $D([0, T], \mathcal{Z})$ equipped with the usual weak topology. Also, let $\mathcal{M}_1(\mathcal{M}_1(D([0, T], \mathcal{Z})))$ denote the space of probability measures on $\mathcal{M}_1(D([0, T], \mathcal{Z}))$ equipped with the weak topology.

2.2. Process-level large deviations

We first recall the definition of the LDP for a family of random variables indexed by one parameter.

Definition 2.1. *(Large deviation principle.)* Let $(\mathcal{S}, d_0)$ be a metric space. We say that a family $\{X^N, N \geq 1\}$ of $\mathcal{S}$-valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies the large deviation principle (LDP) with rate function $I : S \to [0, \infty]$ if the following hold:

- (Compactness of level sets.) For any $s \geq 0$, $\Phi(s) := \{x \in \mathcal{S} : I(x) \leq s\}$ is a compact subset of $\mathcal{S}$.

- (LDP lower bound.) For any $\gamma > 0$, $\delta > 0$, and $x \in \mathcal{S}$, there exists $N_0 \geq 1$ such that

$$\mathbb{P}(d_0(X^N, x) < \delta) \geq \exp\{-N(I(x) + \gamma)\}$$

for any $N \geq N_0$.

- (LDP upper bound.) For any $\gamma > 0$, $\delta > 0$, and $s > 0$, there exists $N_0 \geq 1$ such that

$$\mathbb{P}(d_0(X^N, \Phi(s)) \geq \delta) \leq \exp\{-N(s - \gamma)\}$$

for any $N \geq N_0$.

This definition is also used to study the large deviations of a family of probability measures. For each $N \geq 1$, let $\mathbb{P}^N = \mathbb{P} \circ (X^N)^{-1}$, the law of the random variable X_N on $(\mathcal{S}, d_0)$. We say that the family of probability measures $\{\mathbb{P}^N, N \geq 1\}$ satisfies the LDP on $(\mathcal{S}, d_0)$ with rate function I if the sequence of $\mathcal{S}$-valued random variables $\{X^N, N \geq 1\}$ satisfies the LDP with rate function I.

The LDP lower bound in the above definition is equivalent to the following statement [18, Chapter 3, Section 3]:

$$\liminf_{N\to\infty} \frac{1}{N}\log \mathbb{P}(X^N \in G) \geq -\inf_{x\in G} I(x), \text{ for all } G \subset S \text{ open.}$$

Similarly, under the compactness of the level sets of the rate function I, the LDP upper bound above is equivalent to the following statement:

$$\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}(X^N \in F) \leq -\inf_{x\in F} I(x), \text{ for all } F \subset S \text{ closed.}$$

To study the LDP for the family of invariant measures, we require estimates on the probabilities of the process-level large deviations of $\mu_{\nu_N}^N$. In particular, we consider hitting times of $\mu_{\nu_N}^N$ on certain subsets of the state space $\mathcal{M}_1(\mathcal{Z})$ and apply the process-level LDP lower and upper bounds for $\mu_{\nu_N}^N$ starting at these subsets. Therefore, in addition to the scaling parameter N, we must consider the process $\mu_{\nu_N}^N$ indexed by the initial condition $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$. To study the process-level large deviations of such stochastic processes indexed by two parameters, we use the following definition of the uniform LDP (see Freidlin and Wentzell [18, Chapter 3, Section 3]).

Definition 2.2. *(Uniform large deviation principle.)* We say that the family $\{\mu_{\nu_N}^N, \nu_N \in \mathcal{M}_1^N(\mathcal{Z}), N \geq 1\}$ of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$-valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies the uniform large deviation principle (uniform LDP) over the class $\mathcal{A}$ of subsets of $\mathcal{M}_1^N(\mathcal{Z})$ with the family of rate functions $\{I_\nu, \nu \in \mathcal{M}_1(\mathcal{Z})\}$, $I_\nu : D([0, T], \mathcal{M}_1(\mathcal{Z})) \to [0, +\infty]$, $\nu \in \mathcal{M}_1(\mathcal{Z})$, if the following hold:

- (Compactness of level sets.) For each $K \subset \mathcal{M}_1(\mathcal{Z})$ compact and $s \geq 0$, $\bigcup_{\nu\in K} \Phi_\nu(s)$ is a compact subset of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$, where

$$\Phi_\nu(s) := \{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi(0) = \nu, I_\nu(\varphi) \leq s\}.$$

- (Uniform LDP lower bound.) For any $\gamma > 0$, $\delta > 0$, $s > 0$, and $A \in \mathcal{A}$, there exists $N_0 \geq 1$ such that

$$\mathbb{P}(\rho(\mu_{\nu_N}^N, \varphi) < \delta) \geq \exp\{-N(I_{\nu_N}(\varphi) + \gamma)\}$$

for all $\nu_N \in A \cap \mathcal{M}_1^N(\mathcal{Z})$, $\varphi \in \Phi_{\nu_N}(s)$, and $N \geq N_0$.

- (Uniform LDP upper bound.) For any $\gamma > 0$, $\delta > 0$, $s_0 > 0$, and $A \in \mathcal{A}$, there exists $N_0 \geq 1$ such that

$$\mathbb{P}(\rho(\mu_{\nu_N}^N, \Phi_{\nu_N}(s)) \geq \delta) \leq \exp\{-N(s - \gamma)\},$$

for all $\nu_N \in A \cap \mathcal{M}_1^N(\mathcal{Z})$, $s \leq s_0$, and $N \geq N_0$.

Note that the initial conditions in the upper and lower bounds lie in $A \cap \mathcal{M}_1^N(\mathcal{Z})$, unlike in the definition in [18, Chapter 3, Section 3].

We now make some definitions. Recall τ as defined in (2.3). For each $\nu \in \mathcal{M}_1(\mathcal{Z})$ and $T > 0$, define the functional $S_{[0,T]}(\cdot|\nu) : D([0, T], \mathcal{M}_1(\mathcal{Z})) \to [0, \infty]$ by

$$S_{[0,T]}(\varphi|\nu) := \int_{[0,T]} \sup_{\alpha\in C_0(\mathcal{Z})} \left\{ \langle \alpha, \dot{\varphi}_t - \Lambda_{\varphi_t}^* \varphi_t \rangle - \sum_{(z,z')\in\mathcal{E}} \tau(\alpha(z') - \alpha(z))\lambda_{z,z'}(\varphi_t)\varphi_t(z) \right\} dt,$$

(2.5)

whenever $\varphi(0) = \nu$ and the mapping $[0, T] \ni t \mapsto \varphi(t) \in \mathcal{M}_1(\mathcal{Z})$ is absolutely continuous; $S_{[0,T]}(\varphi|\nu) = \infty$ otherwise. Define the lower level sets of the functional $S_{[0,T]}(\cdot|\nu)$ by

$$\Phi_\nu^{[0,T]}(s) := \{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi(0) = \nu, S_{[0,T]}(\varphi|\nu) \le s\}, \quad s > 0, \ \nu \in \mathcal{M}_1(\mathcal{Z}).$$

The next lemma asserts that these level sets are compact in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ when the initial conditions belong to a compact subset of $\mathcal{M}_1(\mathcal{Z})$. The proof is deferred to Appendix A.

Lemma 2.1. *For each $T > 0$, $s > 0$, and $K \subset \mathcal{M}_1(\mathcal{Z})$ compact,*

$$\{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi(0) \in K, S_{[0,T]}(\varphi|\varphi(0)) \le s\}$$

is a compact subset of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$.

The starting point of our study of the invariant-measure asymptotics is the following uniform LDP for the family $\{\mu_{\nu_N}^N, \nu_N \in \mathcal{M}_1^N(\mathcal{Z}), N \ge 1\}$ over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|\nu), \nu \in \mathcal{M}_1(\mathcal{Z})\}$. Its proof uses the process-level LDP for $\mu_{\nu_N}^N$ studied in Léonard [21] for a fixed initial condition and its extension (when $\mathcal{Z}$ is a finite set) to the case when initial conditions converge to a point in $\mathcal{M}_1(\mathcal{Z})$ in Borkar and Sundaresan [7]. The proof can be found in Appendix A.

Theorem 2.1. *Fix $T > 0$ and assume (A1), (A2), and (A3). Then the family of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$-valued random variables $\{\mu_{\nu_N}^N, \nu_N \in \mathcal{M}_1^N(\mathcal{Z}), N \ge 1\}$ satisfies the uniform LDP over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|\nu), \nu \in \mathcal{M}_1(\mathcal{Z})\}$.*

The rate function $S_{[0,T]}(\cdot|\nu)$ admits a non-variational representation in terms of a minimal cost 'control' that modulates the transition rates across various edges in $\mathcal{E}$ so that the desired trajectory is obtained. Recall τ^* as defined in (2.4).

Theorem 2.2. *(Non-variational representation; cf. Léonard [22].) Let $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ be such that $S_{[0,T]}(\varphi|\varphi(0)) < \infty$. Then there exists a measurable function $h_\varphi : [0, T] \times \mathcal{E} \to \mathbb{R}$ such that*

$$\langle \varphi_t, f_t \rangle = \langle \varphi_0, f_0 \rangle + \int_{[0,t]} \langle \varphi_u, \partial_u f_u \rangle du$$
$$+ \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} (f_u(z') - f_u(z))(1 + h_\varphi(u, z, z'))\lambda_{z,z'}(\varphi_u)\varphi_u(z)du \qquad (2.6)$$

holds for all $t \in [0, T]$ and all $f \in C_0^1([0, T] \times \mathcal{Z})$, and $S_{[0,T]}(\varphi|\varphi(0))$ admits the non-variational representation

$$S_{[0,T]}(\varphi|\varphi(0)) = \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} \tau^*(h_\varphi(t, z, z'))\lambda_{z,z'}(\varphi_t)\varphi_t(z)dt. \qquad (2.7)$$

Remark 2.1. It can be shown that the rate function $S_{[0,T]}$ defined in (2.5) can also be expressed as

$$
\begin{aligned}
S_{[0,T]}(\varphi|\nu) = \sup_{f \in C_0^1([0,T]\times\mathcal{Z})} \Bigg\{ & \langle\varphi_T, f_T\rangle - \langle\varphi_0, f_0\rangle - \int_{[0,T]} \langle\varphi_u, \partial_u f_u\rangle du \\
& - \int_{[0,T]} \langle\varphi_u, L_{\varphi_u} f_u\rangle du - \int_{[0,T]} \sum_{(z,z')\in\mathcal{E}} \tau(f_u(z') - f_u(z))\lambda_{z,z'}(\varphi_u)\varphi_u(z) du \Bigg\},
\end{aligned}
$$
(2.8)

for $\varphi \in D([0,T], \mathcal{M}_1(\mathcal{Z}))$; see Léonard [22]. This form of the rate function will be used in the proofs for the counterexamples in Section 8.

3. Invariant measure: existence, uniqueness, and exponential tightness

In this section we prove Proposition 1.1, the existence and uniqueness of the invariant measure $\wp^N$ for $\mathcal{L}^N$ for each $N \geq 1$, and the exponential tightness of the family of invariant measures $\{\wp^N, N \geq 1\}$. The proof relies on the standard Krylov–Bogolyubov argument and a coupling between the interacting particle system under consideration and a non-interacting system with maximal forward transition rates and minimal backward transition rates.

We first introduce some notation for the non-interacting particle system. Let $\bar{L}$ denote the generator acting on functions f on $\mathcal{Z}$ by

$$
\bar{L}f(z) = \sum_{z' : (z,z')\in\mathcal{E}} \lambda_{z,z'}(f(z') - f(z)), \quad z \in \mathcal{Z},
$$
(3.1)

where $\lambda_{z,z+1} = \overline{\lambda}/(z+1)$ and $\lambda_{z,0} = \underline{\lambda}$. For each $z \in \mathcal{Z}$, let $\bar{P}_z$ denote the solution to the $D([0,T], \mathcal{Z})$-valued martingale problem for $\bar{L}$ with initial condition z. Integration with respect to $\bar{P}_z$ is denoted by $\bar{E}_z$. Let $\pi \in \mathcal{M}_1(\mathcal{Z})$ denote the unique invariant probability measure for $\bar{L}$. Let $\bar{P}_\pi$ denote the solution to the martingale problem for $\bar{L}$ with initial law π. Integration with respect to $\bar{P}_\pi$ is denoted by $\bar{E}_\pi$. By solving the detailed balance equations for $\bar{L}$, we see that

$$
\pi(z) \leq \pi(0) \left(\frac{\overline{\lambda}}{\underline{\lambda}}\right)^z \prod_{k=1}^z \frac{1}{k}, \quad z \geq 1.
$$

In particular, $\pi(z)$ has superexponential decay in z, and $\bar{E}_\pi(\exp\{\beta\vartheta(X)\}) < \infty$ for small enough $\beta > 0$, where ϑ is defined in (1.6). Finally, for each $N \geq 1$, let $\bar{\mathbb{P}}_{\nu_N}^N$ denote the solution to the $D([0,T], \mathcal{M}_1^N(\mathcal{Z}))$-valued martingale problem for $\mathcal{L}^N$ with initial condition $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$, $\lambda_{z,z+1}(\zeta)$ replaced by $\overline{\lambda}/(z+1)$ and $\lambda_{z,0}(\zeta)$ replaced by $\underline{\lambda}$ in (1.1), respectively, for each $\zeta \in \mathcal{M}_1(\mathcal{Z})$. Integration with respect to $\bar{\mathbb{P}}_{\nu_N}^N$ is denoted by $\bar{\mathbb{E}}_{\nu_N}^N$. Also, recall $\mathbb{P}_{\nu_N}^N$, $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$, from Section 2.1.1. We are now ready to prove Proposition 1.1.

Proof of Proposition 1.1. Fix $N \geq 1$. We first show the existence and uniqueness of the invariant probability measure for $\mathcal{L}^N$. Consider the family of probability measures $\{\eta_T^N, T \geq 1\}$ on $\mathcal{M}_1(\mathcal{Z})$ defined by

$$
\eta_T^N(A) := \frac{1}{T}\int_0^T \mathbb{P}_{\delta_0}^N(\mu^N(t)\in A)dt, \quad A \in \mathcal{B}(\mathcal{M}_1(\mathcal{Z})), \ T \geq 1.
$$

Let $X_n^N(t)$ denote the state of the *n*th particle at time *t*. Recall the compact sets $\mathcal{K}_M$, $M > 0$, defined in (2.1). We first couple the laws $\mathbb{P}_{\delta_0}^N$ and $\bar{\mathbb{P}}_{\delta_0}^N$. For $\mathbf{z}^N \in \mathcal{Z}^N$, define $\text{emp}(\mathbf{z}^N) := \frac{1}{N} \sum_{n=1}^N \delta_{z_n^N} \in \mathcal{M}_1^N(\mathcal{Z})$. Let $\mathbf{e}_n^N$ denote the N-length vector with a 1 in the *n*th position and 0 everywhere else. Consider the Markov process on $\mathcal{Z}^N \times \mathcal{Z}^N$ with the infinitesimal generator acting on functions f on $\mathcal{Z}^N \times \mathcal{Z}^N$ by

$$
\begin{aligned}
(\mathbf{z}^N, \bar{\mathbf{z}}^N) \mapsto \sum_{n=1}^N &\left[\left(f(\mathbf{z}^N + \mathbf{e}_n^N, \bar{\mathbf{z}}^N + \mathbf{e}_n^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\lambda_{z_n^N, z_n^N+1}(\text{emp}(\mathbf{z}^N)) \wedge \frac{\overline{\lambda}}{\bar{z}_n^N + 1} \right) \right.\\
&+ \left(f(\mathbf{z}^N + \mathbf{e}_n^N, \bar{\mathbf{z}}^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\lambda_{z_n^N, z_n^N+1}(\text{emp}(\mathbf{z}^N)) - \frac{\overline{\lambda}}{\bar{z}_n^N + 1} \right)^+ \\
&+ \left(f(\mathbf{z}^N, \bar{\mathbf{z}}^N + \mathbf{e}_n^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\frac{\overline{\lambda}}{\bar{z}_n^N + 1} - \lambda_{z_n^N, z_n^N+1}(\text{emp}(\mathbf{z}^N)) \right)^+ \\
&+ \left(f(\mathbf{z}^N - z_n^N \mathbf{e}_n^N, \bar{\mathbf{z}}^N - \bar{z}_n^N \mathbf{e}_n^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\lambda_{z_n^N, 0}(\text{emp}(\mathbf{z}^N)) \wedge \underline{\lambda} \right) \mathbf{1}_{\{z_n^N > 0, \bar{z}_n^N > 0\}} \\
&+ \left(f(\mathbf{z}^N - z_n^N \mathbf{e}_n^N, \bar{\mathbf{z}}^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\lambda_{z_n^N, 0}(\text{emp}(\mathbf{z}^N)) - \underline{\lambda} \right)^+ \mathbf{1}_{\{z_n^N > 0\}} \\
&\left. + \left(f(\mathbf{z}^N, \bar{\mathbf{z}}^N - \bar{z}_n^N \mathbf{e}_n^N) - f(\mathbf{z}^N, \bar{\mathbf{z}}^N) \right) \left(\underline{\lambda} - \lambda_{z_n^N, 0}(\text{emp}(\mathbf{z}^N)) \right)^+ \mathbf{1}_{\{\bar{z}_n^N > 0\}} \right].
\end{aligned}
$$

Such couplings have been studied for continuous-time Markov chains; see, e.g., [27]. Note that, under the above Markov process, for any two initial conditions $\nu_N, \bar{\nu}_N \in \mathcal{M}_1^N(\mathcal{Z})$, the empirical measure flow associated with the first (resp. second) marginal has law $\mathbb{P}_{\nu_N}^N$ (resp. $\bar{\mathbb{P}}_{\bar{\nu}_N}^N$). Therefore, for any $t > 1$, $M > 1$, and $\beta > 0$, we have

$$
\begin{aligned}
\mathbb{P}_{\delta_0}^N(\mu^N(t) \notin \mathcal{K}_M) &\leq \bar{\mathbb{P}}_{\delta_0}^N(\mu^N(t) \notin \mathcal{K}_M) \\
&= \bar{\mathbb{P}}_{\delta_0}^N \left(\sum_{n=1}^N \vartheta(X_n^N(t)) > NM \right) \\
&\leq \exp\{-NM\beta\} \bar{\mathbb{E}}_{\delta_0}^N \left(\exp\left\{ \beta \sum_{n=1}^N \vartheta(X_n^N(t)) \right\} \right) \\
&= \exp\{-NM\beta\}(\bar{E}_0(\exp\{\beta \vartheta(X_1^N(t))\}))^N, \qquad (3.2)
\end{aligned}
$$

where the first inequality follows from the above coupling since (i) the *n*th particle under $\bar{\mathbb{P}}_{\delta_0}^N$ moves from z to $z + 1$ whenever it does so under $\mathbb{P}_{\delta_0}^N$, and (ii) the *n*th particle under $\mathbb{P}_{\delta_0}^N$ moves to 0 (i.e., a z-to-0 transition for some z) whenever it does so under $\bar{\mathbb{P}}_{\delta_0}^N$. The second inequality in (3.2) is a consequence of Chebyshev's inequality. Recall π and the laws $\bar{P}_\pi$ and $\bar{P}_0$. We couple

the laws $\bar{P}_\pi$ and $\bar{P}_0$. Consider the Markov process on $\mathcal{Z} \times \mathcal{Z}$ with the infinitesimal generator acting on functions f on $\mathcal{Z} \times \mathcal{Z}$ by

$$
\begin{aligned}
(\bar{z}_1, \bar{z}_2) \mapsto {}& (f(\bar{z}_1 + 1, \bar{z}_2 + 1) - f(\bar{z}_1, \bar{z}_2)) \left(\frac{\bar{\lambda}}{\bar{z}_1 + 1} \wedge \frac{\bar{\lambda}}{\bar{z}_2 + 1} \right) \\
& + (f(\bar{z}_1 + 1, \bar{z}_2) - f(\bar{z}_1, \bar{z}_2)) \left(\frac{\bar{\lambda}}{\bar{z}_1 + 1} - \frac{\bar{\lambda}}{\bar{z}_2 + 1} \right)^+ \\
& + (f(\bar{z}_1, \bar{z}_2 + 1) - f(\bar{z}_1, \bar{z}_2)) \left(\frac{\bar{\lambda}}{\bar{z}_2 + 1} - \frac{\bar{\lambda}}{\bar{z}_1 + 1} \right)^+ \\
& + (f(0, 0) - f(\bar{z}_1, \bar{z}_2)) \, \underline{\lambda} \mathbf{1}_{\{\bar{z}_1 > 0, \bar{z}_2 > 0\}} \\
& + (f(0, \bar{z}_2) - f(\bar{z}_1, \bar{z}_2)) \, \underline{\lambda} \mathbf{1}_{\{\bar{z}_1 > 0, \bar{z}_2 = 0\}} \\
& + (f(\bar{z}_1, 0) - f(\bar{z}_1, \bar{z}_2)) \, \underline{\lambda} \mathbf{1}_{\{\bar{z}_1 = 0, \bar{z}_2 > 0\}}.
\end{aligned}
$$

Note that, when the initial condition has law (π, δ_0), the first (resp. second) component under the above process has law $\bar{P}_\pi$ (resp. $\bar{P}_0$). Also note that if $\bar{X}_1(0) \geq \bar{X}_2(0)$, then $\bar{X}_1(s) \geq \bar{X}_2(s)$ for all s under the above coupling. Since the first component is at least the second component under the initial law (π, δ_0), it follows that $\bar{E}_0(\exp\{\beta \vartheta(X_1^N(t))\}) \leq \bar{E}_\pi(\exp\{\beta \vartheta(X_1^N(t))\})$. The latter is finite for sufficiently small $\beta > 0$, thanks to the $\exp\{-\vartheta(z)\}$ decay of the probability measure π on $\mathcal{Z}$. Thus we can choose $\bar{\beta} > 0$ small enough (independently of M) so that $\log \bar{E}_\pi(\exp\{\bar{\beta} \vartheta(X_1^N(t))\}) < 1$. Hence (3.2) implies that

$$
\mathbb{P}_{\delta_0}^N(\mu^N(t) \notin \mathscr{K}_M) \leq \exp\{-N(M\bar{\beta} - 1)\}.
$$

Therefore, for any $M > 0$ and $T \geq 1$, we get

$$
\eta_T^N(\sim \mathscr{K}_M) \leq \exp\{-N(M\bar{\beta} - 1)\}. \tag{3.3}
$$

Since $\mathscr{K}_M$ is a compact subset of $\mathcal{M}_1(\mathcal{Z})$, this shows that the family $\{\eta_T^N, T \geq 1\}$ is tight. Hence it follows that there exists an invariant probability measure $\wp^N$ for $\mathscr{L}^N$ (see, e.g., Ethier and Kurtz [16, Theorem 9.3, p. 240]). By Assumption (A1), μ^N is an irreducible Markov process; hence $\wp^N$ is the unique invariant probability measure for $\mathscr{L}^N$.

We now show the exponential tightness of the family $\{\wp^N, N \geq 1\}$. Let $M > 0$ be given, and choose $M' = (M + 1)/\bar{\beta}$. For each $N \geq 1$, since $\wp^N$ is a weak limit of the family $\{\eta_T^N, T \geq 1\}$ as $T \to \infty$, from (3.3) with M replaced by M', it follows that

$$
\wp^N(\sim \mathscr{K}_{M'}) \leq \liminf_{T \to \infty} \eta_T^N(\sim \mathscr{K}_{M'}) \leq \exp\{-NM\} \tag{3.4}
$$

for each $N \geq 1$. Hence,

$$
\limsup_{N \to \infty} \frac{1}{N} \log \wp^N(\sim \mathscr{K}_{M'}) \leq -M,
$$

which establishes that the family $\{\wp^N, N \geq 1\}$ is exponentially tight. This completes the proof of the proposition. $\qquad \square$

4. The LDP lower bound

In this section we prove the LDP lower bound for the family $\{\wp^N, N \geq 1\}$. To lower-bound the probability of a small neighbourhood of a point ξ under $\wp^N$, we first produce a trajectory that starts at $\mathscr{K}_M$ for a suitable $M > 0$, connects to ξ^* with a small cost, and then reaches ξ from ξ^* with cost arbitrarily close to $V(\xi)$, where V is the quasipotential defined in (1.3). The probability of a small neighbourhood of ξ under $\wp^N$ is then lower-bounded by the probability that the process μ^N remains in a small neighbourhood of the trajectory constructed above. The latter is then lower-bounded using the uniform LDP lower bound for μ^N, where the uniformity is over the initial condition lying in a given compact subset of $\mathcal{M}_1(\mathcal{Z})$.

Recall $K(\Delta)$ as defined in (2.2). We begin with a lemma that allows us to connect points in $K(\Delta)$ to ξ^* for small enough Δ with small cost. We omit its proof here, since it follows from a certain continuity property of V which will be shown in Lemma 5.3.

Lemma 4.1. *Given $\gamma > 0$, there exist $\Delta > 0$ and $T = T(\Delta) > 0$ such that for any $\zeta \in K(\Delta)$ there exists a trajectory φ on $[0,T]$ such that $\varphi(0) = \zeta$, $\varphi(T) = \xi^*$, and $S_{[0,T]}(\varphi|\zeta) \leq \gamma$.*

We now prove the LDP lower bound for the family $\{\wp^N, N \geq 1\}$.

Lemma 4.2. *For any $\gamma > 0$, $\delta > 0$, and $\xi \in \mathcal{M}_1(\mathcal{Z})$, there exists $N_0 \geq 1$ such that*

$$\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \xi) < \delta\} \geq \exp\{-N(V(\xi) + \gamma)\} \tag{4.1}$$

for all $N \geq N_0$.

Proof. Fix $\gamma > 0$, $\delta > 0$, and $\xi \in \mathcal{M}_1(\mathcal{Z})$. We may assume that $V(\xi) < \infty$; if $V(\xi) = \infty$ then (4.1) trivially holds for all $N \geq 1$. Choose some $M > 0$ and $N_1 \geq 1$ such that $\wp^N(\mathscr{K}_M) \geq 1/2$ for all $N \geq N_1$; this is possible from the exponential tightness of the family $\{\wp^N, N \geq 1\}$ (see Proposition 1.1). Using Lemma 4.1, choose $\varepsilon > 0$ and $T_0 > 0$ such that for any $\zeta_1 \in K(\varepsilon)$ there exists a trajectory φ_1 on $[0, T_0]$ such that $\varphi_1(0) = \zeta_1$, $\varphi_1(T_0) = \xi^*$, and $S_{[0,T_0]}(\varphi_1|\zeta_1) \leq \gamma/4$. Since ξ^* is the globally asymptotically stable equilibrium for (1.2) and since $\mathscr{K}_M$ is compact, for the above $\varepsilon > 0$, there exists a $T_1 > 0$ such that for any $\zeta \in \mathscr{K}_M$ we have $\mu_\zeta(T_1) \in K(\varepsilon)$, where μ_ζ denotes the solution to the McKean–Vlasov equation (1.2) with initial condition ζ (see Assumption (B2)). Also, by the definition of $V(\xi)$, there exists a $T_2 > 0$ and a trajectory φ_2 such that $\varphi_2(0) = \xi^*$, $\varphi_2(T_2) = \xi$, and $S_{[0,T_2]}(\varphi_2|\xi^*) \leq V(\xi) + \gamma/4$. Let $T = T_1 + T_0 + T_2$. Given $\zeta \in \mathscr{K}_M$, we construct a trajectory φ_ζ on $[0, T]$ by using the above three trajectories as follows. Let $\varphi_\zeta(0) = \zeta$; $\varphi_\zeta(t) = \mu_\zeta(t)$ for $t \in [0, T_1]$; $\varphi_\zeta(t) = \varphi_1(t - T_1)$ for $t \in (T_1, T_1 + T_0]$; and $\varphi_\zeta(t) = \varphi_2(t - (T_1 + T_0))$ for $t \in (T_1 + T_0, T]$. Note that $S_{[0,T]}(\varphi_\zeta|\zeta) \leq V(\xi) + \gamma/2$.

Recall that d is the metric on $\mathcal{M}_1(\mathcal{Z})$ and ρ is the metric on $D([0, T], \mathcal{M}_1(\mathcal{Z}))$. Note that we can choose a $\delta' > 0$ (depending on T and M) such that $\rho(\varphi, \varphi_\zeta) < \delta'$ implies that $d(\varphi(T), \varphi_\zeta(T)) < \delta$ for any $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ and $\zeta \in \mathscr{K}_M$. Indeed, if such a choice is not possible, then there exist a sequence $\{\zeta_n\} \in \mathscr{K}_M$ and a sequence of trajectories $\{\varphi_n\} \subset D([0, T], \mathcal{M}_1(\mathcal{Z}))$ such that $S_{[0,T]}(\varphi_{\zeta_n}|\zeta_n) \leq V(\xi) + \gamma/2$ and $\rho(\varphi_n, \varphi_{\zeta_n}) < 1/n$ for each $n \geq 1$, but $d(\varphi_n(T), \varphi_{\zeta_n}(T)) > \delta$. By the compactness of the level sets of $S_{[0,T]}$ in Lemma 2.1, it follows that there exists a subsequential limit for $\{\varphi_{\zeta_{n_k}}\}_{k \geq 1}$ (say, φ^*); since $\rho(\varphi_n, \varphi_{\zeta_n}) < 1/n$, φ_{n_k} also converges to φ^* in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ as $k \to \infty$. Furthermore, since $S_{[0,T]}(\varphi^*|\varphi_0^*) < \infty$, from Theorem 2.2, we have that $[0, T] \ni t \mapsto \varphi^*(t)$ is continuous. Since $D([0, T], \mathcal{M}_1(\mathcal{Z})) \ni \varphi \mapsto \varphi(T)$ is continuous at all φ such that $t \mapsto \varphi(t)$ is continuous (see, e.g, [4, p. 124]), it follows that $d(\varphi_{n_k}(T), \varphi^*(T)) \to 0$ as $k \to \infty$. This contradicts the assumption $d(\varphi_n(T), \varphi_{\zeta_n}(T)) > \delta$.

This shows that we can choose a $\delta' > 0$ such that $\rho(\varphi, \varphi_\zeta) < \delta'$ implies that $d(\varphi(T), \varphi_\zeta(T)) < \delta$ for any $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ and $\zeta \in \mathcal{K}_M$. Therefore, for each $N \geq N_1$, we have

$$
\begin{aligned}
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \xi) < \delta\} &= \int_{\mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_\zeta^N(d(\mu^N(T), \xi) < \delta) \wp^N(d\zeta) \\
&\geq \int_{\mathcal{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_\zeta^N(d(\mu^N(T), \xi) < \delta) \wp^N(d\zeta) \\
&\geq \int_{\mathcal{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_\zeta^N(\rho(\mu^N, \varphi_\zeta) < \delta') \wp^N(d\zeta) \\
&\geq \frac{1}{2} \inf_{\zeta \in \mathcal{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_\zeta^N(\rho(\mu^N, \varphi_\zeta) < \delta');
\end{aligned} \tag{4.2}
$$

here the first equality follows since $\wp^N$ is invariant to time shifts. By the uniform LDP lower bound in Theorem 2.1, there exists $N_2 \geq N_1$ such that

$$
\mathbb{P}_\zeta^N(\rho(\mu^N, \varphi) < \delta') \geq \exp\{-N(S_{[0,T]}(\varphi|\zeta) + \gamma/4)\}
$$

for all $\zeta \in \mathcal{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})$, $\varphi \in \Phi_\zeta^{[0,T]}(V(\xi) + \gamma/2)$, and $N \geq N_2$. Noting that $S_{[0,T]}(\varphi_\zeta|\zeta) \leq V(\xi) + \gamma/2$ for any $\zeta \in \mathcal{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})$, and using the above uniform LDP lower bound, (4.2) becomes

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \xi) < \delta\} \geq \frac{1}{2} \exp\{-N(V(\xi) + 3\gamma/4)\}
$$

for all $N \geq N_2$. Finally, choose $N_0 \geq N_2$ so that $1/2 \geq \exp\{-N\gamma/4\}$. Then the above becomes

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \xi) < \delta\} \geq \exp\{-N(V(\xi) + \gamma)\}
$$

for all $N \geq N_0$. This completes the proof of the LDP lower bound for the family $\{\wp^N, N \geq 1\}$. $\square$

5. Properties of the quasipotential

In this section we prove three key properties of the quasipotential V defined in (1.3). These three properties are (i) a characterisation of the set of points for which V is finite, (ii) a certain continuity property for V, and (iii) the compactness of the lower level sets of V. These properties play an important role in the proof of the LDP upper bound in Section 6.

5.1. A characterisation of finiteness of the quasipotential

Recall the function ϑ defined in (1.6) and the compact sets $\mathcal{K}_M$, $M > 0$, defined in (2.1). We start with a lemma that enables us to connect δ_0, the point mass at state 0, to a point $\xi \in \mathcal{K}_M$ for some $M > 0$. This connection is made using a trajectory of piecewise constant velocity wherein, for each $z \geq 1$, we move the mass $\xi(z)$ from state 0 to state z in z steps; in the kth step, we move the mass $\xi(z)$ from state $k - 1$ to state k with unit velocity. The lemma asserts that the cost of this piecewise-constant-velocity trajectory is bounded above by a constant that depends only on M.

Lemma 5.1. *Given $M > 0$, there exists a constant C_M depending on M such that for any $\xi \in \mathcal{K}_M$, there exists a $T > 0$ and a trajectory φ on $[0,T]$ such that $\varphi(0) = \delta_0$, $\varphi(T) = \xi$, and $S_{[0,T]}(\varphi|\delta_0) \leq C_M$.*

Proof. Fix $M > 0$ and $\xi \in \mathcal{K}_M$. Fix $J \in \mathcal{Z} \setminus \{0\}$ and define $\mathcal{Z}_J = \{1, 2, \ldots, J\}$, $t_z = z\xi(z)$ for $z \in \mathcal{Z}_J$, and $T_z = \sum_{z' \in \mathcal{Z}_J, z' \geq z} t_{z'}$. Note that $T_J \leq T_{J-1} \leq \cdots \leq T_1$. We shall first construct a trajectory φ^J such that $\varphi^J(0) = \delta_0$, $\varphi^J(T_1)(z) = \xi(z)$ for each $z \in \mathcal{Z}_J$, and $S_{[0,T_1]}(\varphi^J | \delta_0)$ is bounded above by a constant independent of J.

Let $T_{J+1} = 0$. For each $z \in \mathcal{Z}_J$, starting with $z = J$, we move the mass $\xi(z)$ from state 0 to state z using a piecewise unit velocity trajectory over the time duration $(T_{z+1}, T_{z+1} + t_z]$. We define this trajectory φ^J on $[0, T_1]$ as follows. Let $\varphi_0^J = \delta_0$. For each $z \in \mathcal{Z}_J$ and $1 \leq k \leq z$, when $t \in (T_{z+1} + (k-1)\xi(z), T_{z+1} + k\xi(z)]$, let

$$\dot{\varphi}_t^J(l) = \begin{cases} 1 & \text{if } l = k, \\ -1 & \text{if } l = k-1, \\ 0 & \text{otherwise,} \end{cases}$$

for $l \in \mathcal{Z}$, and define $\varphi_t^J(l) = \delta_0(l) + \int_{[0,t]} \dot{\varphi}_u^J(l) du$, $l \in \mathcal{Z}$, $t \in [0, T]$.

We now calculate the cost of this trajectory. Fix $z \in \mathcal{Z}$ such that $\xi(z) > 0$, and let $1 \leq k \leq z$. For each $t \in (T_{z+1} + (k-1)\xi(z), T_{z+1} + k\xi(z))$ and $\alpha \in C_0(\mathcal{Z})$, note that

$$\langle \alpha, \dot{\varphi}_t^J - \Lambda_{\varphi_t^J}^* \varphi_t^J \rangle - \sum_{(z,z') \in \mathcal{E}} \tau(\alpha(z') - \alpha(z)) \lambda_{z,z'}(\varphi_t^J) \varphi_t^J(z)$$

$$= (\alpha(k) - \alpha(k-1)) - \sum_{(z,z') \in \mathcal{E}} (\exp\{\alpha(z') - \alpha(z)\} - 1) \lambda_{z,z'}(\varphi_t^J) \varphi_t^J(z).$$

Hence,

$$\sup_{\alpha \in C_0(\mathcal{Z})} \left\{ \langle \alpha, \dot{\varphi}_t^J - \Lambda_{\varphi_t^J}^* \varphi_t^J \rangle - \sum_{(z,z') \in \mathcal{E}} \tau(\alpha(z') - \alpha(z)) \lambda_{z,z'}(\varphi_t^J) \varphi_t^J(z) \right\}$$

$$\leq \sup_{x \in \mathbb{R}} (x - (\exp\{x\} - 1) \lambda_{k-1,k}(\varphi_t^J) \varphi_t^J(k-1))$$

$$+ \sup_{\alpha \in C_0(\mathcal{Z})} \left(- \sum_{(z,z') \in \mathcal{E}; (z,z') \neq (k-1,k)} (\exp\{\alpha(z') - \alpha(z)\} - 1) \lambda_{z,z'}(\varphi_t^J) \varphi_t^J(z) \right)$$

$$\leq \log \left(\frac{1}{\varphi_t^J(k-1) \lambda_{k-1,k}(\varphi_t^J)} \right) + 2\bar{\lambda}$$

$$\leq \log \left(\frac{1}{\varphi_t^J(k-1)} \right) + \log k + \log \left(\frac{1}{\underline{\lambda}} \right) + 2\bar{\lambda}, \tag{5.1}$$

where the last two inequalities follow from Assumption (A2). Consider the first term above. For $k > 1$, integration of this quantity over the time duration $t \in (T_{z+1} + (k-1)\xi(z), T_{z+1} + k\xi(z))$ gives

$$\int_{(T_{z+1} + (k-1)\xi(z), T_{z+1} + k\xi(z))} \log \left(\frac{1}{\varphi_t^J(k-1)} \right) dt = - \int_{\xi(z)}^0 \log \left(\frac{1}{u} \right) du$$

$$= (u \log u - u) \Big|_{\xi(z)}^0$$

$$= \xi(z) \log \left(\frac{1}{\xi(z)} \right) + \xi(z),$$

where the first equality follows from the variable change $u = \varphi_t^J(k-1)$ and the facts (i) $\dot{\varphi}_t^J(k-1) = -1$, (ii) $\varphi_t^J(k-1) = \xi(z)$ when $t = T_{z+1} + (k-1)\xi(z)$, (iii) $\varphi_t^J(k-1) = 0$ when $t = T_{z+1} + k\xi(z)$, and (iv) $du = -dt$. For $k = 1$, using the bound $\varphi_t^J(0) \geq \varphi_t^J(0) - (1 - \sum_{z'=z}^J \xi(z'))$, we get

$$\int_{(T_{z+1}, T_{z+1}+\xi(z))} \log\left(\frac{1}{\varphi_t^J(0)}\right) dt$$
$$\leq \int_{(T_{z+1}, T_{z+1}+\xi(z))} \log\left(\frac{1}{\varphi_t^J(0) - (1 - \sum_{z'=z}^J \xi(z'))}\right) dt$$
$$= -\int_{\xi(z)}^0 \log\left(\frac{1}{u}\right) du,$$

where the last equality follows from the variable change $u = \varphi_t^J(0) - (1 - \sum_{z'=z}^J \xi(z'))$ and the facts (i) $\dot{\varphi}_t^J(0) = -1$, (ii) $\varphi_t^J(0) = 1 - \sum_{z'=z+1}^J \xi(z')$ when $t = T_{z+1}$, so that $\varphi_t^J(0) - (1 - \sum_{z'=z}^J \xi(z')) = \xi(z)$ when $t = T_{z+1}$, (iii) $\varphi_t^J(0) = 1 - \sum_{z'=z}^J \xi(z')$ when $t = T_{z+1} + \xi(z)$, so that $\varphi_t^J(0) - (1 - \sum_{z'=z}^J \xi(z')) = 0$ when $t = T_{z+1} + \xi(z)$, and (iv) $du = -dt$. Thus, proceeding as before for the case $k > 1$, we arrive at

$$\int_{(T_{z+1}, T_{z+1}+\xi(z))} \log\left(\frac{1}{\varphi_t^J(0)}\right) dt \leq \xi(z) \log\left(\frac{1}{\xi(z)}\right) + \xi(z).$$

Hence, integrating (5.1) over $t \in (T_{z+1} + (k-1)\xi(z), T_{z+1} + k\xi(z))$ and summing over $1 \leq k \leq z$, we get, for each $z \in \mathcal{Z}_J$,

$$\int_{(T_{z+1}, T_{z+1}+t_z)} \sup_{\alpha \in C_0(\mathcal{Z})} \left\{ \langle \alpha, \dot{\varphi}_t^J - \Lambda_{\varphi_t^J}^* \varphi_t^J \rangle - \sum_{(z,z') \in \mathcal{E}} \tau(\alpha(z') - \alpha(z)) \lambda_{z,z'}(\varphi_t^J) \varphi_t^J(z) \right\} dt$$
$$\leq z\xi(z) \log\left(\frac{1}{\xi(z)}\right) + \tilde{C}_z, \tag{5.2}$$

where $\tilde{C}_z = (z \log z + z)\xi(z) + z\xi(z)\left(\log\left(\frac{1}{\lambda}\right) + 2\bar{\lambda}\right)$. Let $\tilde{C}^J = \sum_{z \in \mathcal{Z}_J} \tilde{C}_z$. Then, summing the above display over $z \in \mathcal{Z}_J$, we arrive at

$$S_{[0,T_1]}(\varphi^J | \delta_0) \leq \sum_{z \in \mathcal{Z}_J} z\xi(z) \log\left(\frac{1}{\xi(z)}\right) + \tilde{C}^J.$$

Note that

$$\sum_{z \in \mathcal{Z}_J} z\xi(z) \log\left(\frac{1}{\xi(z)}\right) = \sum_{\substack{z \in \mathcal{Z}_J : \\ \xi(z) \leq 1/z^3}} z\xi(z) \log\left(\frac{1}{\xi(z)}\right) + \sum_{\substack{z \in \mathcal{Z}_J : \\ \xi(z) > 1/z^3}} z\xi(z) \log\left(\frac{1}{\xi(z)}\right)$$
$$\leq \frac{1}{e} + \sum_{\substack{z \in \mathcal{Z}_J \setminus \{1\} : \\ \xi(z) \leq 1/z^3}} \frac{3 \log z}{z^2} + 3 \sum_{\substack{z \in \mathcal{Z}_J : \\ \xi(z) > 1/z^3}} (z \log z)\xi(z)$$
$$\leq \frac{1}{e} + 3 \sum_{z \in \mathcal{Z}_J} \left\{ \frac{\log z}{z^2} + (z \log z)\xi(z) \right\}, \tag{5.3}$$

where the first inequality comes from the fact that the mapping $x \mapsto x \log(1/x)$ is monotonically increasing for $x \in [0, 1/e]$. Hence,

$$S_{[0,T_1]}(\varphi^J | \delta_0) \leq \frac{1}{e} + 3 \sum_{z \in \mathcal{Z}_J} \left\{ \frac{\log z}{z^2} + (z \log z)\xi(z) \right\} + \tilde{C}^J, \qquad J \geq 1.$$

Define $T = \sum_{z \in \mathcal{Z}} z\xi(z)$. We now extend the trajectory φ^J to $(T_1, T]$ by defining $\varphi_t^J = \varphi_{T_1}^J$ for $t \in (T_1, T]$. Noting that $\dot{\varphi}_t^J(z) = 0$ for all $z \in \mathcal{Z}$ on $t \in (T_1, T]$, this extension suffers an additional cost of at most $2\bar{\lambda}T$. Hence, we get

$$S_{[0,T]}(\varphi^J | \delta_0) \leq \frac{1}{e} + 3 \sum_{z \in \mathcal{Z}_J} \left\{ \frac{\log z}{z^2} + (z \log z)\xi(z) \right\} + \tilde{C}^J + 2\bar{\lambda}T.$$

Noting that (i) the right-hand side above is bounded above by $\langle \xi, \vartheta \rangle C(\bar{\lambda}, \underline{\lambda})$, where $C(\bar{\lambda}, \underline{\lambda})$ is a constant depending on $\bar{\lambda}$ and $\underline{\lambda}$, and (ii) $\langle \xi, \vartheta \rangle \leq M$, the above display yields

$$S_{[0,T]}(\varphi^J | \delta_0) \leq C(M, \bar{\lambda}, \underline{\lambda}),$$

where $C(M, \bar{\lambda}, \underline{\lambda})$ is a constant depending on M, $\bar{\lambda}$, and $\underline{\lambda}$. Using the compactness of the level sets of $S_{[0,T]}$ (see Lemma 2.1), it follows that the sequence of trajectories $\{\varphi^J, J \geq 1\}$ has a convergent subsequence. Re-indexing the original sequence, let $\varphi^J \to \varphi$ in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ as $J \to \infty$. By construction, for each $J \in \mathcal{Z} \setminus \{0\}$, $\varphi_T^J(z) = \xi(z)$ for all $z \in \mathcal{Z}_J$; hence $\varphi_T(z) = \xi(z)$ for all $z \in \mathcal{Z}$. Recall that lower semicontinuity of $S_{[0,T]}$ was proved in the course of the proof of Lemma 2.1. Therefore, it follows that

$$S_{[0,T]}(\varphi | \delta_0) \leq \liminf_{J \to \infty} S_{[0,T]}(\varphi^J | \delta_0) \leq C(M, \bar{\lambda}, \underline{\lambda}).$$

This completes the proof of the lemma. $\qquad \square$

We are now ready to characterise the set of points ξ in $\mathcal{M}_1(\mathcal{Z})$ such that $V(\xi)$ is finite.

Lemma 5.2. *We have $V(\xi) < \infty$ if and only if $\xi \in \mathcal{K}$. Furthermore, for any $M > 0$, there exists a constant $C_M > 0$ such that $\xi \in \mathcal{K}_M$ implies $V(\xi) \leq C_M$.*

Proof. Let $\xi \in \mathcal{M}_1(\mathcal{Z})$ be such that $V(\xi) < \infty$. Then there exist a $T > 0$ and a trajectory φ on $[0, T]$ such that $\varphi(0) = \xi^*$, $\varphi(T) = \xi$, and $S_{[0,T]}(\varphi | \xi^*) \leq V(\xi) + 1$. By Theorem 2.2, there exists a measurable function h_φ on $[0, T] \times \mathcal{E}$ such that

$$\langle \varphi_t, f \rangle = \langle \varphi_0, f \rangle + \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} (f(z') - f(z))(1 + h_\varphi(u, z, z'))\lambda_{z,z'}(\varphi_u)\varphi_u(z)du \qquad (5.4)$$

holds for all $t \in [0, T]$ and $f \in C_0(\mathcal{Z})$, and $S_{[0,T]}(\varphi | \varphi(0))$ is given by

$$S_{[0,T]}(\varphi | \varphi(0)) = \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} \tau^*(h_\varphi(t, z, z'))\lambda_{z,z'}(\varphi_t)\varphi_t(z)dt.$$

For any $x \geq 0$ and $y \in \mathbb{R}$, using the convex duality relation $(x-1)y \leq \tau^*(x-1) + \tau(y)$, we get the inequality $xy \leq \tau^*(x-1) + (\exp\{y\} - 1)$. Hence, from the above non-variational representation for $S_{[0,T]}(\varphi | \varphi(0))$, (5.4) implies

$$
\begin{aligned}
\langle \varphi_t, f \rangle &\leq \langle \xi^*, f \rangle + \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} \tau^*(h_\varphi(u, z, z')) \lambda_{z,z'}(\varphi_u) \varphi_u(z) du \\
&\quad + \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} (\exp\{f(z') - f(z)\} - 1) \lambda_{z,z'}(\varphi_u) \varphi_u(z) du \\
&\leq \langle \xi^*, f \rangle + V(\xi) + 1 \\
&\quad + \int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} (\exp\{f(z') - f(z)\} - 1) \lambda_{z,z'}(\varphi_u) \varphi_u(z) du. \quad (5.5)
\end{aligned}
$$

Recall the function ϑ on $\mathcal{Z}$. For $n \geq 1$, define

$$
\vartheta_n(z) = \begin{cases} \vartheta(z), & \text{if } z \leq n, \\ 0, & \text{otherwise.} \end{cases}
$$

By convexity, note that $\vartheta_n(z+1) - \vartheta_n(z) \leq 1 + \log(z+1)$ and $\vartheta_n(0) - \vartheta_n(z) \leq 0$ for each $z \in \mathcal{Z}$. Therefore, using the upper bound for the transition rates from Assumption (A2), we have

$$
\int_{[0,t]} \sum_{(z,z') \in \mathcal{E}} (\exp\{\vartheta_n(z') - \vartheta_n(z)\} - 1) \lambda_{z,z'}(\varphi_u) \varphi_u(z) du \leq \bar{\lambda}(e - 1)t
$$

for each $t \in [0, T]$ and $n \geq 1$. It follows from (5.5) with f replaced by ϑ_n that

$$
\langle \varphi_t, \vartheta_n \rangle \leq \langle \xi^*, \vartheta_n \rangle + V(\xi) + 1 + \bar{\lambda}(e - 1)T
$$

for each $t \in [0, T]$ and $n \geq 1$. Letting $n \to \infty$ and using monotone convergence, we conclude that

$$
\sup_{t \in [0,T]} \langle \varphi_t, \vartheta \rangle = \sup_{t \in [0,T]} \lim_{n \to \infty} \langle \varphi_t, \vartheta_n \rangle \leq \langle \xi^*, \vartheta \rangle + V(\xi) + 1 + \bar{\lambda}(e - 1)T. \quad (5.6)
$$

In particular, $\langle \xi, \vartheta \rangle \leq \langle \xi^*, \vartheta \rangle + V(\xi) + 1 + \bar{\lambda}(e - 1)T$. It follows that $\xi \in \mathcal{K}$.

Conversely, let $\xi \in \mathcal{K}$. Let $M > 0$ be such that $\xi \in \mathcal{K}_M$. By Lemma 5.1, there exist a $T > 0$ and a trajectory $\varphi^{(2)}$ on $[0, T]$ such that $\varphi^{(2)}(0) = \delta_0$, $\varphi^{(2)}(T) = \xi$, and $S_{[0,T]}(\varphi^{(2)} | \delta_0) \leq C_M$ for some constant $C_M > 0$ depending on M. Let $t_0 = 0$, $t_z = \sum_{z'=1}^{z} \xi^*(z')$, $z \in \mathcal{Z} \setminus \{0\}$, and $T_1 = \sum_{z' \neq 0} \xi^*(z')$. We now construct another trajectory $\varphi^{(1)}$ on $[0, T_1]$ such that $\varphi^{(1)}(0) = \xi^*$, $\varphi^{(1)}(T_1) = \delta_0$, and $S_{[0,T_1]}(\varphi^{(1)} | \xi^*) < \infty$. This trajectory is constructed using piecewise-constant-velocity paths, and its cost $S_{[0,T_1]}(\varphi^{(1)} | \xi^*)$ is computed using arguments similar to those used in the proof of Lemma 5.1; we provide the details here for completeness. When $t \in (t_{z-1}, t_z]$ for some $z \in \mathcal{Z} \setminus \{0\}$, let

$$
\dot{\varphi}_t^{(1)}(l) = \begin{cases} -1 & \text{if } l = z, \\ 1 & \text{if } l = 0, \\ 0 & \text{otherwise,} \end{cases}
$$

for $l \in \mathcal{Z}$, and define $\varphi_t^{(1)}(l) = \varphi_0^{(1)}(l) + \int_{[0,t]} \dot{\varphi}_u^{(1)}(l)du$, $l \in \mathcal{Z}$, $t \in [0, T_1]$. Note that for each $\alpha \in C_0(\mathcal{Z})$, when $t \in (t_{z-1}, t_z)$, we have

$$\left\{ \langle \alpha, \dot{\varphi}_t^{(1)} - \Lambda^*_{\varphi_t^{(1)}} \varphi_t^{(1)} \rangle - \sum_{(z,z') \in \mathcal{E}} \tau(\alpha(z') - \alpha(z)) \lambda_{z,z'}(\varphi_t^{(1)}) \varphi_t^{(1)}(z) \right\}$$

$$= (\alpha(0) - \alpha(z)) - (\exp\{\alpha(0) - \alpha(z)\} - 1) \lambda_{z,0}(\varphi_t^{(1)}) \varphi_t^{(1)}(z)$$

$$- \sum_{(z_0, z') \in \mathcal{E} \,:\, (z_0, z') \neq (z, 0)} (\exp\{\alpha(z') - \alpha(z_0)\} - 1) \lambda_{z_0, z'}(\varphi_t^{(1)}) \varphi_t^{(1)}(z_0) \Bigg\},$$

so that optimising the left-hand side of the above display over $\alpha \in C_0(\mathcal{Z})$ yields

$$\sup_{\alpha \in C_0(\mathcal{Z})} \left\{ \langle \alpha, \dot{\varphi}_t^{(1)} - \Lambda^*_{\varphi_t^{(1)}} \varphi_t^{(1)} \rangle - \sum_{(z,z') \in \mathcal{E}} \tau(\alpha(z') - \alpha(z)) \lambda_{z,z'}(\varphi_t^{(1)}) \varphi_t^{(1)}(z) \right\}$$

$$\leq \log \left(\frac{1}{\varphi_t^{(1)}(z) \lambda_{z,0}(\varphi_t^{(1)})} \right) + 2\bar{\lambda}$$

$$\leq \log \left(\frac{1}{\varphi_t^{(1)}(z)} \right) + \log \left(\frac{1}{\underline{\lambda}} \right) + 2\bar{\lambda},$$

where the last inequality follows from the lower bound on the backward transition rates in Assumption (A2). Integrating the above over (t_{z-1}, t_z) and summing over $z \in \mathcal{Z} \setminus \{0\}$, we arrive at

$$S_{[0,T_1]}(\varphi^{(1)} | \xi^*) \leq \sum_{z \in \mathcal{Z} \setminus \{0\}} \left\{ \xi^*(z) \log \frac{1}{\xi^*(z)} + \xi^*(z) \left(1 + \log \left(\frac{1}{\underline{\lambda}} \right) + 2\bar{\lambda} \right) \right\}. \tag{5.7}$$

Since $\xi^* \in \mathcal{K}$, proceeding via the steps in (5.3), we conclude that the right-hand side of the above display is finite. We combine $\varphi^{(1)}$ and $\varphi^{(2)}$ and define a new trajectory $\tilde{\varphi}$ on $[0, T_1 + T]$ as follows: $\tilde{\varphi}(t) = \varphi^{(1)}(t)$ on $t \in [0, T_1]$; $\tilde{\varphi}(t) = \varphi^{(2)}(t - T_1)$ on $t \in (T_1, T_1 + T]$. Note that $\tilde{\varphi}(0) = \xi^*$, $\tilde{\varphi}(T_1 + T) = \xi$, and $S_{[0,T_1+T]}(\tilde{\varphi} | \xi^*) < \infty$. Hence $V(\xi) < \infty$.

To prove the second statement, we note that given any $M > 0$, for any $\xi \in \mathcal{K}_M$, the cost of the trajectory $\tilde{\varphi}$ constructed in the previous paragraph is bounded above by a constant depending only on M (and not on ξ). This completes the proof of the lemma. $\qquad \square$

5.2. Continuity

We now establish a certain continuity property of the quasipotential V. Since V has compact level sets and the space $\mathcal{M}_1(\mathcal{Z})$ is not locally compact, we cannot expect V to be continuous on $\mathcal{M}_1(\mathcal{Z})$. In fact, for any point $\xi \in \mathcal{M}_1(\mathcal{Z})$ with $V(\xi) < \infty$, one can produce a sequence $\{\xi_n, n \geq 1\}$ such that $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ as $n \to \infty$, and $\langle \xi_n, \vartheta \rangle = \infty$ for all $n \geq 1$, so that $\inf_{n \geq 1} V(\xi_n) = \infty$. We prove that V is continuous under the convergence of ϑ-moments when it is restricted to $\mathcal{K}$. That is, when $\xi_n, \xi \in \mathcal{K}$, $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$, and $\langle \xi_n, \vartheta \rangle \to \langle \xi, \vartheta \rangle$ as $n \to \infty$, we have $V(\xi_n) \to V(\xi)$ as $n \to \infty$. Towards this, we produce a trajectory that connects ξ to ξ_n by first moving the mass from all the large enough states z back to state 0, then producing a constant-velocity trajectory that fills the required mass from state 0 to all the large enough states z, and

finally adjusting mass within a finite subset of $\mathcal{Z}$ to reach ξ_n. We show that the cost of the trajectory constructed above can be made arbitrarily small for large enough n.

Lemma 5.3. *Let* $\xi_n \in \mathcal{K}$, $n \geq 1$, *and* $\xi \in \mathcal{K}$. *Suppose that* $\xi_n \to \xi$ *in* $\mathcal{M}_1(\mathcal{Z})$ *and* $\langle \xi_n, \vartheta \rangle \to \langle \xi, \vartheta \rangle$ *as* $n \to \infty$. *Then* $V(\xi_n) \to V(\xi)$ *as* $n \to \infty$.

Proof. We first prove that $\limsup_{n \to \infty} V(\xi_n) \leq V(\xi)$. Fix $\varepsilon > 0$. We shall move from ξ to ξ_n in five steps. The outline of this construction is as follows:

- $\varphi^{(0)}$: This trajectory starts with ξ and moves all the mass for all states $z > z_0$, for a suitable large enough z_0, back to state 0. This backward movement results in a cost of $O(\varepsilon)$.

- $\varphi^{(1)}$: Next, we move any additional mass, if required, from the states $\{1, 2, \ldots, z_0\}$ back to state 0 so that there is enough mass at state 0 to fill up all the states beyond z_0. Again, this backward movement results in a cost of $O(\varepsilon)$.

- $\varphi^{(2)}$: Next, we construct a trajectory of piecewise constant velocity to move the mass $\sum_{z' > z_0} \xi_n(z)$ from state 0 to state $z_0 + 1$. After this movement, state $z_0 + 1$ contains all the mass required to fill up the states beyond it. This forward movement results in a cost of $O(\varepsilon \log(1/\varepsilon))$, instead of $O(\varepsilon)$, because we move the total mass for all the states beyond z_0.

- $\varphi^{(3)}$: Then, for each $z > z_0$, we move the required mass (i.e., $\xi_n(z)$) from state 0 to state z using a trajectory of piecewise constant velocity. At the end of this procedure, for each $z > z_0$, the mass at state z becomes $\xi_n(z)$. This forward movement results in a cost of $O(\varepsilon)$.

- $\varphi^{(4)}$: Finally, we adjust the mass within the finite set $\{1, 2, \ldots, z_0\}$ to match with ξ_n. This also results in a cost of at most $O(\varepsilon \log(1/\varepsilon))$. Again, this cost is $O(\varepsilon \log(1/\varepsilon))$ instead of $O(\varepsilon)$ because we move, for each $z \in \{1, 2, \ldots, z_0\}$, the sum of the additional mass (under ξ_n as opposed to ξ) in the states $\{z, z+1, \ldots, z_0\}$ from state 0 to state z.

Therefore, the total cost of all these trajectories is at most $O(\varepsilon \log(1/\varepsilon))$, which vanishes as $\varepsilon \to 0$. We now define these trajectories in detail and evaluate their costs.

Let $z_0 \geq 2$ be such that

$$\sum_{z > z_0} \vartheta(z)\xi(z) < \varepsilon/6 \quad \text{and} \quad \sum_{z > z_0} \frac{\log z}{z^2} < \varepsilon.$$

Then choose $n_1 \geq 1$ such that $\sum_{z > z_0} \vartheta(z)\xi_n(z) < \varepsilon/3$ holds for all $n \geq n_1$; this is possible since $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ and $\langle \xi_n, \vartheta \rangle \to \langle \xi, \vartheta \rangle$ as $n \to \infty$. Let

$$t_{z_0} = 0, \quad t_z = \sum_{z' = z_0 + 1}^{z} \xi(z'), \ z > z_0, \quad \text{and} \quad T_0 = \sum_{z' > z_0} \xi(z').$$

Define the trajectory $\varphi^{(0)}$ on $[0, T_0]$ as follows. When $t \in (t_{z-1}, t_z]$ for some $z > z_0$, let

$$\dot{\varphi}_t^{(0)}(l) = \begin{cases} -1 & \text{if } l = z, \\ 1 & \text{if } l = 0, \\ 0 & \text{otherwise,} \end{cases}$$

for $l \in \mathcal{Z}$, and define

$$\varphi_t^{(0)}(l) = \xi(l) + \int_{[0,t]} \dot{\varphi}_u^{(0)}(l)du, \quad l \in \mathcal{Z}, \; t \in [0, T_0].$$

Note that $\varphi_{T_0}^{(0)}(z) = \xi(z)$ for $1 \le z \le z_0$, $\varphi_{T_0}^{(0)}(z) = 0$ for $z > z_0$, and $\varphi_{T_0}^{(0)}(0) = \xi(0) + \sum_{z>z_0} \xi(z)$. Let $M = (\sup_{n \ge n_1} \langle \xi_n, \vartheta \rangle) \vee \langle \xi, \vartheta \rangle + 1$. Using ideas similar to those used in the proof of Lemma 5.2, it can be checked that $S_{[0,T_0]}(\varphi^{(0)}|\xi) \le C_0(M, \bar{\lambda}, \underline{\lambda})\varepsilon$, for some constant $C_1(M, \bar{\lambda}, \underline{\lambda})$ depending on M, $\bar{\lambda}$, and $\underline{\lambda}$. Indeed, the cost is $O(\sum_{z>z_0} \xi(z) \log(1/\xi(z)))$, which, using the argument used to arrive at the bound (5.7) and the choice of z_0, is bounded by

$$O\left(\sum_{z>z_0} \left((z \log z)\xi(z) + \frac{\log z}{z^2} \right) \right) = O(\varepsilon).$$

Let $\varepsilon_n = \sum_{z>z_0} \xi_n(z)$. If $\varepsilon_n > \varphi_{T_0}^{(0)}(0)$, then we move the extra mass $\varepsilon_n - \varphi_{T_0}^{(0)}(0)$ from the states $\{1, 2, \ldots, z_0\}$ to state 0 as follows. Let $T_1 = T_0 + \varepsilon_n - \varphi_{T_0}^{(0)}(0)$. When t is between $T_0 + \sum_{z'=z+1}^{z_0} \varphi_{T_0}^{(0)}(z')$ and $(T_0 + \sum_{z'=z}^{z_0} \varphi_{T_0}^{(0)}(z')) \wedge T_1$ for some $z \le z_0$, let

$$\dot{\varphi}_t^{(1)}(l) = \begin{cases} -1 & \text{if } l = z, \\ 1 & \text{if } l = 0, \\ 0 & \text{otherwise}, \end{cases}$$

for $l \in \mathcal{Z}$. Define the trajectory $\varphi^{(1)}$ on $[0, T_1]$ as follows: $\varphi_t^{(1)} = \varphi_t^{(0)}$ when $t \in [0, T_0]$; $\varphi_t^{(1)}(l) = \varphi_{T_0}^{(0)}(l) + \int_{[0,t]} \dot{\varphi}_u^{(1)}(l)du$, $l \in \mathcal{Z}$, $t \in (T_0, T_1)$. Note that $\varphi^{(1)}$ depends on n, but we suppress this in the notation for ease of reading. Again, since ε_n is smaller than $\varepsilon/3$, by using calculations similar to those used in the proof of Lemma 5.2, we see that $S_{[T_0,T_1]}(\varphi^{(1)}|\varphi_{T_0}^{(0)}) \le C_1(M, \bar{\lambda}, \underline{\lambda})\varepsilon$ for some constant $C_1(M, \bar{\lambda}, \underline{\lambda})$ depending on M, $\bar{\lambda}$, and $\underline{\lambda}$. On the other hand, if $\varepsilon_n \le \varphi_{T_0}^{(0)}(0)$, we set $T_1 = T_0$ and $\varphi_t^{(1)} = \varphi_t^{(0)}$ on $[0, T_1]$. In both cases, we have $\varphi_{T_1}^{(1)}(0) \ge \varepsilon_n$.

Let $T_2 = (z_0 + 1)\varepsilon_n$. We now construct another trajectory $\varphi^{(2)}$ on $[0, T_2]$ to transfer the mass ε_n from state 0 (in $\varphi_{T_1}^{(1)}$) to state $z_0 + 1$. Let $\varphi_0^{(2)} = \varphi_{T_1}^{(1)}$. When $t \in ((z-1)\varepsilon_n, z\varepsilon_n]$ for some $z \in \{1, 2, \ldots, z_0 + 1\}$, let

$$\dot{\varphi}_t^{(2)}(l) = \begin{cases} -1 & \text{if } l = z - 1, \\ 1 & \text{if } l = z, \\ 0 & \text{otherwise}, \end{cases}$$

for $l \in \mathcal{Z}$, and define $\varphi_t^{(2)}(l) = \varphi_{T_1}^{(1)}(l) + \int_{[0,t]} \dot{\varphi}_u^{(2)}(l)du$, $l \in \mathcal{Z}$, $t \in (0, T_2]$. Note that $|x \log(\frac{1}{x}) - y \log(\frac{1}{y})| \le \delta + \delta \log(1/\delta)$ whenever $|x - y| \le \delta$, and that $\varepsilon_n \le \varepsilon/(z_0 \log z_0)$. Hence, using calculations similar to those done in the proof of Lemma 5.1, we see that $S_{[0,T_2]}(\varphi^{(2)}|\varphi_{T_1}^{(1)})$ can be bounded above by $C_2(M, \bar{\lambda}, \underline{\lambda})\varepsilon \log(1/\varepsilon)$, where $C_2(M, \bar{\lambda}, \underline{\lambda})$ is a constant depending on M, $\bar{\lambda}$, and $\underline{\lambda}$, for each $n \ge n_1$ (recall that $\varphi^{(2)}$ depends on n). Indeed, the cost is bounded by the order of

$$z_0 \varepsilon_n \log\left(\frac{1}{\varepsilon_n} \right) + (z_0 \log z_0)\varepsilon_n \le z_0 \frac{\varepsilon}{z_0 \log z_0} \log\left(\frac{z_0 \log z_0}{\varepsilon} \right) + \varepsilon$$

$$\le \varepsilon \log(1/\varepsilon) + 3\varepsilon$$

(see the bound in (5.2)), where the first inequality uses the fact that $\varepsilon_n \leq \varepsilon/(z_0 \log z_0)$, and the second inequality uses the fact that $z_0 \geq 2$, so that $z_0 \log z_0 > 1$.

Note that $\varphi^{(2)}_{T_2}(z_0 + 1) = \varepsilon_n$. We now construct a trajectory that distributes this mass ε_n from state $z_0 + 1$ to all the states $z \geq z_0 + 1$ to match with $\xi_n(z)$. Let $t'_z = z\xi_n(z)$ for $z \geq z_0 + 2$ and $T_3 = \sum_{z \geq z_0 + 2} t'_z$. Similarly to the construction in the proof of Lemma 5.1, we can now construct a trajectory $\varphi^{(3)}$ on $[0, T_3]$ such that $\varphi^{(3)}_0 = \varphi^{(2)}_{T_2}$, $\varphi^{(3)}_{T_3}(z) = \xi_n(z)$ for each $z \geq z_0 + 1$, and $S_{[0,T_3]}(\varphi^{(3)}|\varphi^{(2)}_{T_2}) \leq C_3(M, \overline{\lambda}, \underline{\lambda})\varepsilon$ for some constant $C_3(M, \overline{\lambda}, \underline{\lambda})$ depending on M, $\overline{\lambda}$, and $\underline{\lambda}$, for all $n \geq n_1$. Indeed, using the bounds in (5.2) and (5.3), the total cost is bounded by the order of

$$\sum_{z > z_0 + 1} \left((z \log z)\xi_n(z) + \frac{\log z}{z^2} \right) \leq \frac{\varepsilon}{3} + \varepsilon,$$

where the inequality follows from the choice of z_0.

Finally, we construct a trajectory that connects $\varphi^{(3)}_{T_3}$ to ξ_n by adjusting the mass within the states $\{0, 1, \ldots, z_0\}$. Note that $\varphi^{(3)}_{T_3}(z) = \xi_n(z)$ for each $z \geq z_0 + 1$. Let $\mathcal{Z}_0 \subset \{1, 2, \ldots, z_0\}$ denote the set of all $z \in \{1, 2, \ldots, z_0\}$ such that $\varphi^{(3)}_{T_3}(z) > \xi_n(z)$. Similarly to the construction of $\varphi^{(1)}$, for each $z \in \mathcal{Z}_0$, we move the mass $\varphi^{(3)}_{T_3}(z) - \xi_n(z)$ from state z to state 0 using unit velocity over a time duration $\varphi^{(3)}_{T_3}(z) - \xi_n(z)$. Once these mass transfers are complete, starting with $z = 1$, we move the mass

$$\sum_{z' \geq z, z' \notin \mathcal{Z}_0, z' \leq z_0} (\xi_n(z') - \varphi^{(3)}_{T_3}(z'))$$

from state $z - 1$ to state z at unit rate. Let

$$T_4 = \sum_{z \in \mathcal{Z}_0} (\varphi^{(3)}_{T_3}(z) - \xi_n(z)) + \sum_{z \notin \mathcal{Z}_0, z \leq z_0} (\xi_n(z) - \varphi^{(3)}_{T_3}(z)),$$

and let $\varphi^{(4)}$ denote this piecewise-constant-velocity trajectory. Let $\tilde{\varepsilon}_n = \sum_{z \notin \mathcal{Z}_0, z \leq z_0} (\xi_n(z) - \varphi^{(3)}_{T_3}(z))$. At each step of $\varphi^{(4)}$, since we move a mass of at most $\tilde{\varepsilon}_n$ from state $z - 1$ to state z, the cost of $\varphi^{(4)}$ is at most of the order of

$$z_0 \tilde{\varepsilon}_n \log \left(\frac{1}{\tilde{\varepsilon}_n} \right) + (z_0 \log z_0)\tilde{\varepsilon}_n$$

(see (5.2)). Since $\tilde{\varepsilon}_n \to 0$ as $n \to \infty$, we may choose $n_2 \geq n_1$ so that $\tilde{\varepsilon}_n \leq \varepsilon/(z_0 \log z_0)$ for all $n \geq n_2$. Therefore, for $n \geq n_2$, the above display is bounded by

$$z_0 \frac{\varepsilon}{z_0 \log z_0} \log \left(\frac{z_0 \log z_0}{\varepsilon} \right) + \varepsilon \leq \varepsilon \log(1/\varepsilon) + 3\varepsilon,$$

which is $O(\varepsilon \log(1/\varepsilon))$. Therefore, $S_{[0,T_4]}(\varphi^{(4)}|\varphi^{(3)}_{T_3}) \leq C_4(M, \overline{\lambda}, \underline{\lambda})\varepsilon \log(1/\varepsilon)$ for all $n \geq n_2$, for some constant $C_4(M, \overline{\lambda}, \underline{\lambda})$ depending on M, $\overline{\lambda}$, and $\underline{\lambda}$.

Let $T = \sum_{i=1}^4 T_i$. We now append the four paths $\varphi^{(i)}$, $1 \leq i \leq 4$, constructed in the previous paragraphs over the time duration $[0, T]$ to get a path φ such that $\varphi_0 = \xi$, $\varphi_T = \xi_n$, and

$S_{[0,T]}(\varphi|\xi) \le C(M, \overline{\lambda}, \underline{\lambda})\varepsilon \log(1/\varepsilon)$, where $C(M, \overline{\lambda}, \underline{\lambda})$ is a constant depending on M, $\overline{\lambda}$ and $\underline{\lambda}$. Hence, for each $n \ge n_2$, we have

$$V(\xi_n) \le V(\xi) + S_{[0,T_4]}(\varphi|\xi) \le V(\xi) + C(M, \overline{\lambda}, \underline{\lambda})\varepsilon \log(1/\varepsilon).$$

Therefore, $\limsup_{n\to\infty} V(\xi_n) \le V(\xi) + C(M, \overline{\lambda}, \underline{\lambda})\varepsilon \log(1/\varepsilon)$. Letting $\varepsilon \to 0$ and noting that $\varepsilon \log(1/\varepsilon) \to 0$, we arrive at $\limsup_{n\to\infty} V(\xi_n) \le V(\xi)$.

To prove $\liminf_{n\to\infty} V(\xi_n) \ge V(\xi)$, we reverse the role of ξ_n and ξ in the above argument. That is, we construct a trajectory φ on $[0, T]$ such that $\varphi_0 = \xi_n$, $\varphi_T = \xi$, and $S_{[0,T]}(\varphi|\xi_n) \le \varepsilon_n$ for all $n \ge 1$, where $\varepsilon_n \to 0$ as $n \to \infty$. Thus, we get

$$V(\xi) \le V(\xi_n) + \varepsilon_n.$$

Letting $n \to \infty$, we conclude that $\liminf_{n\to\infty} V(\xi_n) \ge V(\xi)$. This completes the proof of the lemma. $\qquad\square$

Remark 5.1. The choice of n_1 in the above proof suggests that the inequality $\limsup_{n\to\infty} V(\xi_n) \le V(\xi)$ can be proved as long as $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ as $n \to \infty$ and $\limsup_{n\to\infty}\langle \xi_n, \vartheta \rangle \le \langle \xi, \vartheta \rangle$ holds. Similarly, the inequality $\liminf_{n\to\infty} V(\xi_n) \ge V(\xi)$ can be proved as long as $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ and $\liminf_{n\to\infty}\langle \xi_n, \vartheta \rangle \ge \langle \xi, \vartheta \rangle$ holds. This observation will be used later in the proof of the compactness of the lower level sets of V.

5.3. Compactness of the lower level sets of the quasipotential

Define the level sets of V by

$$\Xi(s) := \{\xi \in \mathcal{M}_1(\mathcal{Z}) : V(\xi) \le s\}, \qquad s > 0.$$

In this section we establish the compactness of $\Xi(s)$ for each $s > 0$.

Lemma 5.4. *For each $s > 0$, $\Xi(s)$ is a compact subset of $\mathcal{M}_1(\mathcal{Z})$.*

Proof. We first prove an inclusion property of the level sets of V, namely, that given $M > 0$ there exists $M' > 0$ such that

$$\{\xi \in \mathcal{M}_1(\mathcal{Z}) : V(\xi) \le M\} \subset \mathscr{K}_{M'}. \tag{5.8}$$

On one hand, using Proposition 1.1 on the exponential tightness of the family $\{\wp^N, N \ge 1\}$, we can choose $M' > 0$ (see (3.4)) such that

$$\limsup_{N\to\infty} \frac{1}{N} \log \wp^N(\sim\mathscr{K}_{M'}) \le -(M+1).$$

On the other hand, using the LDP lower bound established in Lemma 4.2 and the compactness of $\mathscr{K}_{M'}$, we have

$$\liminf_{N\to\infty} \frac{1}{N} \log \wp^N(\sim\mathscr{K}_{M'}) \ge - \inf_{\xi \notin \mathscr{K}_{M'}} V(\xi).$$

Combining the above two displays, we get

$$- \inf_{\xi \notin \mathscr{K}_{M'}} V(\xi) \le \liminf_{N\to\infty} \frac{1}{N} \log \wp^N(\sim\mathscr{K}_{M'}) \le \limsup_{N\to\infty} \frac{1}{N} \log \wp^N(\sim\mathscr{K}_{M'}) \le -(M+1).$$

That is, $\xi \notin \mathscr{K}_{M'}$ implies $V(\xi) \geq M + 1 > M$. This shows (5.8). By Prokhorov's theorem, $\mathscr{K}_M$ is a compact subset of $\mathcal{M}_1(\mathcal{Z})$; hence (5.8) shows that $\Xi(s)$ is precompact for each $s > 0$.

We now show that $\Xi(s)$ is closed in $\mathcal{M}_1(\mathcal{Z})$. Let $\xi_n \in \Xi(s)$ for each $n \geq 1$, and let $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ as $n \to \infty$. By Fatou's lemma, we have $\liminf_{n \to \infty} \langle \xi_n, \vartheta \rangle \geq \langle \xi, \vartheta \rangle$. Hence, by Remark 5.1, we have $\liminf_{n \to \infty} V(\xi_n) \geq V(\xi)$. Thus, $\xi \in \Xi(s)$. This completes the proof of the lemma. $\qquad \square$

6. The LDP upper bound

Recall $\mathscr{K}_M$ as defined in (2.1) and $K(\Delta)$ as defined in (2.2). For $m \in \mathbb{N}$, define

$$\mathscr{S}_m(\Delta, M) = \{\varphi \in D([0, m], \mathcal{M}_1(\mathcal{Z})) : \varphi(0) \in \mathscr{K}_M, \varphi(n) \notin K(\Delta) \text{ for all } n = 1, 2, \ldots, m\}.$$

That is, $\mathscr{S}_m(\Delta, M)$ denotes the set of all trajectories that start at $\mathscr{K}_M$ and do not intersect $K(\Delta)$ at all integer time points in $[0, m]$. We begin with a lemma that asserts that the elements of $\mathscr{S}_m(\Delta, M)$ for large enough m must have non-trivial cost. The key idea used in the proof comes from the compactness of level sets of the process-level large deviations rate function $S_{[0,T]}(\cdot|\nu)$, $\nu \in K$, for any compact subset K of $\mathcal{M}_1(\mathcal{Z})$ (see Lemma 2.1).

Lemma 6.1. *For any $s > 0$, $M > 0$, and $\Delta > 0$, there exists $m_0 \in \mathbb{N}$ such that*

$$\inf\{S_{[0,m_0]}(\varphi|\varphi(0)), \varphi \in \mathscr{S}_{m_0}(\Delta, M)\} > s. \tag{6.1}$$

Proof. Suppose not. Then there exist $s > 0$, $M > 0$, $\Delta > 0$, a sequence of positive numbers $\{\varepsilon_m, m \geq 1\}$ such that $\varepsilon_m \to 0$ as $m \to \infty$, and a sequence of trajectories $\{\varphi_m, m \geq 1\}$ such that $\varphi_m \in \mathscr{S}_m(\Delta, M)$, and $S_{[0,m]}(\varphi_m|\varphi_m(0)) \leq s + \varepsilon_m$ for each $m \geq 1$.

Note that there exists an $M_1 > 0$ such that $\varphi_m(t) \in \mathscr{K}_{M_1}$ for each $t \in [0, m]$ and each $m \geq 1$. Indeed, by Lemma 5.2, there exists $C_M > 0$ such that $\zeta \in K(\Delta)$ implies $V(\zeta) \leq C_M$. Thus, for each $m \geq 1$, there exist a $\bar{T}_m > 0$ and a trajectory $\bar{\varphi}_m$ on $[0, \bar{T}_m]$ such that $\bar{\varphi}_m(0) = \xi^*$, $\bar{\varphi}_m(\bar{T}_m) = \zeta \in K(\Delta)$, and $S_{[0,\bar{T}_m]}(\bar{\varphi}_m|\xi^*) \leq C_M + 1$. We extend this trajectory $\bar{\varphi}_m$ to $(\bar{T}_m, \bar{T}_m + m]$ by defining $\bar{\varphi}_m(t) = \varphi_m(t - \bar{T}_m)$ on $t \in (\bar{T}_m, \bar{T} + m]$. Note that $S_{[0,\bar{T}_m+m]}(\bar{\varphi}_m|\xi^*) \leq C_M + 1 + s + \varepsilon_m$, so that $V(\varphi_m(t)) \leq C_M + 1 + s + \varepsilon_m$ for each $t \in [0, m]$ and each $m \geq 1$. Thus, we can find an $M_1 > 0$ such that (5.8) holds with M replaced by $C_M + s + \sup_{m \geq 1} \varepsilon_m + 2$ and M' replaced by M_1. It follows that $\varphi_m(t) \in \mathscr{K}_{M_1}$ for each $t \in [0, m]$ and each $m \geq 1$.

For the above choice of M_1, using Assumption (B2), choose $T_1 > 1$ such that $\mu_\zeta(t) \in K(\Delta/2)$ for each $t \geq T_1$ and each $\zeta \in \mathscr{K}_{M_1}$, where μ_ζ is the solution to the McKean–Vlasov equation (1.2) with initial condition ζ. Note that the closure of the set of all trajectories φ on $[0, T_1]$ in $D([0, T_1], \mathcal{M}_1(\mathcal{Z}))$ with initial condition $\varphi(0) \in \mathscr{K}_{M_1}$ and $\varphi(T_1) \notin K(\Delta)$ does not contain any trajectory of the McKean–Vlasov equation (1.2). It follows from Lemma 2.1 that

$$\beta := \inf\{S_{[0,T_1]}(\varphi|\varphi(0)), \varphi(0) \in \mathscr{K}_{M_1}, \varphi(n) \notin K(\Delta) \text{ for each } n = 1, 2, \ldots, \lfloor T_1 \rfloor\} > 0.$$

Therefore, noting that $\varphi_m(t) \in \mathscr{K}_{M_1}$ for each $t \in [0, m]$ and $m \geq 1$, we see that

$$\begin{aligned} S_{[0,m]}(\varphi_m|\varphi_m(0)) &\geq \sum_{n=1}^{\lfloor m/T_1 \rfloor} S_{[(n-1)T_1, nT_1]}(\varphi_m|\varphi_m((n-1)T_1)) \\ &\geq \left\lfloor \frac{m}{T_1} \right\rfloor \beta \\ &\to \infty \quad \text{as } m \to \infty, \end{aligned}$$

which contradicts our assumption. This completes the proof of the lemma. $\qquad \square$

With a slight abuse of notation, given $A \subset \mathcal{M}_1(\mathcal{Z})$, $s > 0$, and $T > 0$, define

$$\Phi_A^{[0,T]}(s) := \{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi(0) \in A, S_{[0,T]}(\varphi|\varphi(0)) \leq s\}.$$

We now prove a certain containment property for elements of $\mathcal{M}_1(\mathcal{Z})$ that can arise as endpoints of trajectories in $\Phi_{K(\Delta)}^{[0,T]}(s)$, $s > 0$ and $\Delta > 0$, i.e., points $\xi \in \mathcal{M}_1(\mathcal{Z})$ such that there exists a trajectory φ with $\varphi_0 \in K(\Delta)$ and $S_{[0,T]}(\varphi|\varphi_0) \leq s$. We prove that such points are not far from the lower level sets of V in $\mathcal{M}_1(\mathcal{Z})$. This connection between trajectories over finite time horizons and the level sets of the quasipotential V is the key to transferring the process-level LDP upper bound in Theorem 2.1 to the LDP upper bound for the family of invariant measures $\{\wp^N, N \geq 1\}$.

Lemma 6.2. *For any $s > 0$ and $\delta > 0$ there exist $\Delta > 0$ and $T_1 \geq 1$ such that for all $T \geq T_1$,*

$$\{\varphi(T) : \varphi \in \Phi_{K(\Delta)}^{[0,T]}(s)\} \subset \{\xi \in \mathcal{M}_1(\mathcal{Z}) : d(\xi, \Xi(s)) \leq \delta\}. \tag{6.2}$$

Proof. Suppose not. Then there exist $s > 0$, $\delta > 0$, sequences $\{\Delta_n, n \geq 1\}$ and $\{T_n, n \geq 1\}$ such that $\Delta_n \downarrow 0$ and $T_n \uparrow \infty$ as $n \to \infty$, and trajectories $\varphi_n \in \Phi_{K(\Delta_n)}^{[0,T_n]}(s)$ such that $d(\varphi_n(T_n), \Xi(s)) > \delta$ for each $n \geq 1$. Let $\xi_n = \varphi_n(T_n)$, $n \geq 1$. By Lemma 5.3, there exist a $T' > 0$ and a sequence $\{\varepsilon_n, n \geq 1\}$, with $\varepsilon_n \to 0$ as $n \to \infty$, such that for any $\zeta' \in K(\Delta_n)$ there exists a trajectory $\bar{\varphi}^{\zeta'}$ on $[0, T']$ such that $\bar{\varphi}^{\zeta'}(0) = \xi^*$, $\bar{\varphi}^{\zeta'}(T') = \zeta'$, and $S_{[0,T']}(\bar{\varphi}^{\zeta'}|\xi^*) \leq \varepsilon_n$. For each $n \geq 1$, let $\tilde{\varphi}_n$ be the trajectory on $[0, T' + T_n]$ defined as follows. Let $\tilde{\varphi}_n(0) = \xi^*$; $\tilde{\varphi}_n(t) = \bar{\varphi}^{\varphi_n(0)}(t)$ on $t \in [0, T']$; $\tilde{\varphi}_n(t) = \varphi_n(t - T')$ on $t \in (T', T' + T_n]$. In particular, $\tilde{\varphi}_n(T' + T_n) = \xi_n$. Clearly, $S_{[0,T'+T_n]}(\tilde{\varphi}_n|\xi^*) \leq s + \varepsilon_n$. It follows that $V(\xi_n) \leq s + \varepsilon_n$. Using the compactness of the lower level sets of V (see Lemma 5.4), we can find a convergent subsequence of $\{\xi_n, n \geq 1\}$; after re-indexing and denoting this convergent subsequence by $\{\xi_n, n \geq 1\}$, let $\xi_n \to \xi$ in $\mathcal{M}_1(\mathcal{Z})$ as $n \to \infty$. By assumption, $d(\xi_n, \Xi(s)) > \delta$ for each $n \geq 1$, and hence $d(\xi, \Xi(s)) \geq \delta$. Using the lower semicontinuity of V, we see that

$$V(\xi) \leq \liminf_{n \to \infty} V(\xi_n) \leq \liminf_{n \to \infty} (s + \varepsilon_n) = s.$$

Hence $\xi \in \Xi(s)$. This contradicts $d(\xi, \Xi(s)) \geq \delta$, which is a consequence of our assumption. This proves the lemma. $\quad\square$

We are now ready to prove the LDP upper bound for the family $\{\wp^N, N \geq 1\}$. The proof relies on the uniform LDP upper bound in Theorem 2.1, the exponential tightness of the family $\{\wp^N, N \geq 1\}$, the containment property established in Lemma 6.2, an estimate on the probability that μ^N lies in $\mathscr{S}_m(M, \Delta)$ (which uses the process-level uniform LDP upper bound in Theorem 2.1 and the result of Lemma 6.1), and finally the strong Markov property of μ^N.

Lemma 6.3. *For any $\gamma > 0$, $\delta > 0$, and $s > 0$, there exists $N_0 \geq 1$ such that*

$$\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \Xi(s)) \geq \delta\} \leq \exp\{-N(s - \gamma)\}$$

for all $N \geq N_0$.

Proof. Fix $\gamma > 0$, $\delta > 0$, and $s > 0$. Choose $M > 0$ and $N_1 \geq 1$ such that $\wp^N(\sim\mathscr{K}_M) \leq \exp\{-Ns\}$ for all $N \geq N_1$; this is possible from the exponential tightness of the family $\{\wp^N, N \geq 1\}$ (see Proposition 1.1). For the given $s > 0$ and $\delta > 0$, from Lemma 6.2, choose $\Delta > 0$ and $T_1 > 0$ such that (6.2) holds for all $T \geq T_1$. For the above choice of $\Delta > 0$ and $M > 0$, by

Lemma 6.1, choose $m_0 \in \mathbb{N}$ such that (6.1) holds. By (6.1) and the compactness of $\Phi_{\mathscr{K}_M}^{[0,m_0]}(s)$ in $D([0, m_0], \mathcal{M}_1(\mathcal{Z}))$ (which follows from Lemma 2.1), the closure of $\mathscr{S}_{m_0}(\Delta, M)$ does not intersect $\Phi_{\mathscr{K}_M}^{[0,m_0]}(s)$. It follows that there exists a $\delta_0 > 0$ such that $\varphi \in \mathscr{S}_{m_0}(\Delta, M)$ implies $\rho(\varphi, \Phi_{\mathscr{K}_M}^{[0,m_0]}(s)) \geq \delta_0$. Hence by the uniform LDP upper bound in Theorem 2.1, there exists $N_2 \geq N_1$ such that

$$\mathbb{P}_{\zeta}^N(\mu^N \in \mathscr{S}_{m_0}(\Delta, M)) \leq \mathbb{P}_{\zeta}^N(\rho(\mu^N, \Phi_{\mathscr{K}_M}^{[0,m_0]}(s)) \geq \delta_0)$$
$$\leq \exp\{-N(s - \gamma/2)\} \tag{6.3}$$

for all $\zeta \in \mathscr{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})$ and $N \geq N_2$. Thus, with $T = m_0 + T_1$ and $N \geq N_2$, we have

$$\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \Xi(s)) \geq \delta\}$$
$$= \int_{\mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta}^N(d(\mu^N(T), \Xi(s)) \geq \delta)\wp^N(d\zeta)$$
$$\leq \exp\{-Ns\} + \int_{\mathscr{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta}^N(d(\mu^N(T), \Xi(s)) \geq \delta)\wp^N(d\zeta)$$
$$\leq \exp\{-Ns\} + \sup_{\zeta \in \mathscr{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta}^N(\mu^N \in \mathscr{S}_{m_0}(\Delta, M))$$
$$+ \int_{\mathscr{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta}^N(\mu^N \notin \mathscr{S}_{m_0}(\Delta, M), d(\mu^N(T), \Xi(s)) \geq \delta)\wp^N(d\zeta)$$
$$\leq \exp\{-Ns\} + \exp\{-N(s - \gamma/2)\}$$
$$+ \int_{\mathscr{K}_M \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta}^N(\mu^N \notin \mathscr{S}_{m_0}(\Delta, M), d(\mu^N(T), \Xi(s)) \geq \delta)\wp^N(d\zeta); \tag{6.4}$$

here the first equality follows since $\wp^N$ is invariant to time shifts, the first inequality follows from the choice of M, and the third inequality follows from (6.3).

To bound the integrand in the third term above, let $T' \geq T_1$ and $\zeta' \in K(\Delta)$. Choose $0 < \delta' < \delta$ (depending on T and s, and not on ζ' and T') such that $\rho(\varphi_1, \varphi_2) < \delta'/2$ implies $d(\varphi_1(T'), \varphi_2(T')) < \delta/2$ whenever $\varphi_1 \in D([0, T'], \mathcal{M}_1(\mathcal{Z}))$ and $\varphi_2 \in \Phi_{\zeta'}^{[0,T']}(s)$. The existence of such a δ' can be justified via arguments similar to those used in the proof of Lemma 4.2; see the paragraph before (4.2). Note that if a trajectory φ on $[0, T']$ with initial condition $\varphi(0) = \zeta'$ is such that $\rho(\varphi, \Phi_{\zeta'}^{[0,T']}(s)) < \delta'/2$, then there exists a trajectory $\varphi' \in \Phi_{\zeta'}^{[0,T']}(s)$ such that $\rho(\varphi, \varphi') < \delta'/2$. By the choice of δ', we have $d(\varphi(T'), \varphi'(T')) < \delta/2$. By Lemma 6.2, we find that $d(\varphi'(T'), \Xi(s)) \leq \delta'/2$. Hence, by the triangle inequality, $d(\varphi(T'), \Xi(s)) < \delta/2 + \delta'/2 < \delta$. The contrapositive of the above statement is

$$d(\varphi(T'), \Xi(s)) \geq \delta \Rightarrow \rho(\varphi, \Phi_{\zeta'}^{[0,T']}(s)) \geq \delta'/2.$$

We therefore conclude that

$$\mathbb{P}_{\zeta'}^N(d(\mu^N(T'), \Xi(s)) \geq \delta) \leq \mathbb{P}_{\zeta'}^N(\rho(\mu^N, \Phi_{\zeta'}^{[0,T']}(s)) \geq \delta'/2) \tag{6.5}$$

for all $T' \geq T_1$, $\zeta' \in \mathscr{K}(\Delta) \cap \mathcal{M}_1^N(\mathcal{Z})$, and $N \geq 1$.

Note that the integrand in the last term of (6.4) can be bounded above by

$$
\begin{aligned}
&\mathbb{P}_\zeta^N(\mu^N \notin \mathscr{S}_{m_0}(\Delta, M), d(\mu^N(T), \Xi(s)) \geq \delta) \\
&\quad = \mathbb{P}_\zeta^N(\mu^N(m) \in K(\Delta) \text{ for some } m = 1, 2, \ldots, m_0, \ d(\mu^N(T), \Xi(s)) \geq \delta) \\
&\quad \leq \sum_{m=1}^{m_0} \sup_{\zeta' \in K(\Delta) \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta'}^N(d(\mu^N(T-m), \Xi(s)) \geq \delta) \\
&\quad \leq \sum_{m=1}^{m_0} \sup_{\zeta' \in K(\Delta) \cap \mathcal{M}_1^N(\mathcal{Z})} \mathbb{P}_{\zeta'}^N(\rho((\mu^N(t), 0 \leq t \leq T-m), \Phi_{\zeta'}^{[0,T-m]}(s)) \geq \delta'/2), \quad (6.6)
\end{aligned}
$$

where the first inequality follows from the strong Markov property of μ^N and the second inequality follows from (6.5) by the choice of T. By the uniform LDP upper bound in Theorem 2.1, for each $m = 1, 2, \ldots m_0$, there exists $N(m) \geq N_2$ such that

$$
\mathbb{P}_{\zeta'}^N(\rho((\mu^N(t), 0 \leq t \leq T-m), \Phi_{\zeta'}^{[0,T-m]}(s)) \geq \delta'/2) \leq \exp\{-N(s - \gamma/2)\}
$$

for all $\zeta' \in \mathscr{H}(\Delta) \cap \mathcal{M}_1^N(\mathcal{Z})$ and $N \geq N(m)$. Put $N_3 = \max\{N(m), m = 1, 2, \ldots, m_0, N_1, N_2\}$. Then (6.6) yields

$$
\mathbb{P}_\zeta^N(\mu^N \notin \mathscr{S}_{m_0}(\Delta, M), d(\mu^N(T), \Xi(s)) \geq \delta) \leq m_0 \exp\{-N(s - \gamma/2)\}
$$

for all $\zeta \in \mathscr{H}_M \cap \mathcal{M}_1^N(\mathcal{Z})$ and $N \geq N_3$. Substitution of this back in (6.4) yields

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \Xi(s)) \geq \delta\} \leq \exp\{-Ns\} + (m_0 + 1) \exp\{-N(s - \gamma/2)\}
$$

for all $N \geq N_3$. Finally, choose $N_0 \geq N_3$ such that $1 + (m_0 + 1) \exp\{N\gamma/2\} \leq \exp\{N\gamma\}$ for all $N \geq N_0$. Then the above display becomes

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \Xi(s)) \geq \delta\} \leq \exp\{-N(s - \gamma)\}
$$

for all $N \geq N_0$. This completes the proof of the lemma. $\qquad \square$

7. Proof of Theorem 1.1

We now complete the proof of Theorem 1.1.

- (Compactness of level sets.) For any $s > 0$, by Lemma 5.4, the set $\Xi(s) = \{\xi \in \mathcal{M}_1(\mathcal{Z}) : V(\xi) \leq s\}$ is a compact subset of $\mathcal{M}_1(\mathcal{Z})$.

- (LDP lower bound.) Given $\gamma > 0$, $\delta > 0$, and $\xi \in \mathcal{M}_1(\mathcal{Z})$, by Lemma 4.2, there exists $N_0 \geq 1$ such that

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \xi) < \delta\} \geq \exp\{-N(V(\xi) + \gamma)\}
$$

for all $N \geq N_0$.

- (LDP upper bound.) Given $\gamma > 0$, $\delta > 0$, and $s > 0$, by Lemma 6.3, there exists $N_0 \geq 1$ such that

$$
\wp^N\{\zeta \in \mathcal{M}_1(\mathcal{Z}) : d(\zeta, \Xi(s)) \geq \delta\} \leq \exp\{-N(s - \gamma)\}
$$

for all $N \geq N_0$.

This completes the proof of Theorem 1.1.

8. Two counterexamples

In this section, for the two non-interacting counterexamples described in Section 1.1, we prove that the quasipotential is not equal to the relative entropy with respect to the corresponding globally asymptotically stable equilibrium. These two counterexamples are (i) a system of non-interacting M/M/1 queues, and (ii) a system of non-interacting nodes in a wireless local area network (WLAN) with constant forward transition rates. We detail the proofs in the case of non-interacting M/M/1 queues. Similar arguments carry over to the case of a non-interacting WLAN system with constant forward transition rates as well.

8.1. A system of non-interacting M/M/1 queues

Recall the system of non-interacting M/M/1 queues described in Section 1.1.1. Recall the relative entropy from (1.4) and the process-level large deviations rate function from (2.8). Also recall the function ϑ defined in (1.6) and the compact sets $\mathscr{K}_M$, $M > 0$, defined in (2.1). Define the quasipotential

$$V_Q(\xi) := \inf\{S^Q_{[0,T]}(\varphi|\xi^*_Q), \varphi(0) = \xi^*_Q, \varphi(T) = \xi, T > 0\}, \quad \xi \in \mathcal{M}_1(\mathcal{Z}),$$

where S^Q is defined by (2.8) with $\mathcal{E}$ replaced by $\mathcal{E}_Q$ and L_ζ replaced by L^Q for each $\zeta \in \mathcal{M}_1(\mathcal{Z})$.

We first prove that the quasipotential V_Q is not finite outside $\mathscr{K}$. The key property used for this is the fact that the attractor ξ^*_Q has geometric decay. As a consequence, $\langle \xi^*_Q, \vartheta \rangle < \infty$. Using this property, we first show that if $\xi \notin \mathscr{K}$, then the associated quasipotential evaluated at ξ cannot be finite. This is shown by producing a lower bound for the cost of any trajectory starting at ξ^*_Q and ending at $\xi \notin \mathscr{K}$ from the rate function in (2.8).

Lemma 8.1. *If $\xi \in \mathcal{M}_1(\mathcal{Z})$ is such that $\xi \notin \mathscr{K}$, then $V_Q(\xi) = \infty$.*

Proof. Fix $\xi \in \mathcal{M}_1(\mathcal{Z})$. Let $T > 0$ and $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ be such that $\varphi_0 = \xi^*_Q$ and $\varphi_T = \xi$. For each $n \geq 1$, define f_n by

$$f_n(z) = \begin{cases} z, & \text{if } z \leq n, \\ 2n - z, & \text{if } n + 1 \leq z \leq 2n, \\ 0, & \text{if } z > 2n, \end{cases}$$

and define $f_\infty(z) = z$ for each $z \in \mathcal{Z}$. Note that the purpose of f_n is to approximate f_∞ using $C_0(\mathcal{Z})$ functions so that we can insert them into (2.8). We first assume that $\langle \xi, f_\infty \rangle = \infty$. In particular, $\xi \notin \mathscr{K}$. Using the function f_n in place of f in the right-hand side of (2.8), we have

$$S^Q_{[0,T]}(\varphi|\xi^*_Q) \geq \langle \varphi_T, f_n \rangle - \langle \xi^*_Q, f_n \rangle - \int_{[0,T]} \langle \varphi_u, L^Q f_n \rangle - \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}_Q} \tau(f_n(z') - f_n(z))\lambda_{z,z'}\varphi_u(z)du$$

$$= \langle \varphi_T, f_n \rangle - \langle \xi^*_Q, f_n \rangle - \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}_Q} (\exp\{f_n(z') - f_n(z)\} - 1)\lambda_{z,z'}\varphi_u(z)du,$$

where $\lambda_{z,z+1} = \lambda_f$, $z \in \mathcal{Z}$, and $\lambda_{z,z-1} = \lambda_b$, $z \in \mathcal{Z} \setminus \{0\}$. Noting that $f_n(z') - f_n(z)$ is either 1, 0, or -1 for each $(z, z') \in \mathcal{E}_Q$, we have

$$\sum_{(z,z') \in \mathcal{E}_Q} (\exp\{f_n(z') - f_n(z)\} - 1)\lambda_{z,z'}\varphi_u(z) \leq 2(e - 1)\lambda_b$$

for each $u \in [0, T]$. Hence the above becomes

$$S^Q_{[0,T]}(\varphi|\xi^*_Q) \geq \langle \varphi_T, f_n \rangle - \langle \xi^*_Q, f_n \rangle - 2(e - 1)\lambda_b T.$$

Note that $\langle \xi_Q^*, f_\infty \rangle < \infty$. Hence, letting $n \to \infty$ and using the monotone convergence theorem, we conclude that $S_{[0,T]}^Q(\varphi | \xi_Q^*) = \infty$.

We now assume that $\xi \notin \mathcal{K}$ is such that $\langle \xi, f_\infty \rangle < \infty$. Let $T > 0$ and $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ be such that $\varphi_0 = \xi_Q^*$ and $\varphi_T = \xi$. Without loss of generality, we can assume that $\sup_{t \in [0,T]} \langle \varphi_t, f_\infty \rangle < \infty$; otherwise the argument in the above paragraph shows that $S_{[0,T]}^Q(\varphi | \xi_Q^*) = \infty$. Define

$$\vartheta_n(z) = \begin{cases} \vartheta(z), & \text{if } z \leq n, \\ \vartheta(2n - z), & \text{if } n + 1 \leq z \leq 2n, \\ 0, & \text{if } z > 2n. \end{cases}$$

Using ϑ_n in the right-hand side of (2.8), we get

$$S_{[0,T]}^Q(\varphi | \xi_Q^*) \geq \langle \xi, \vartheta_n \rangle - \langle \xi_Q^*, \vartheta_n \rangle - \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}_Q} (\exp\{\vartheta_n(z') - \vartheta_n(z)\} - 1)\lambda_{z,z'}\varphi_u(z)du.$$

Noting that $\vartheta_n(z') - \vartheta_n(z)$ can be bounded above by $1 + \log(z + 1)$ for each $(z, z') \in \mathcal{E}_Q$, it follows that

$$\sum_{(z,z') \in \mathcal{E}_Q} (\exp\{\vartheta_n(z') - \vartheta_n(z)\} - 1)\lambda_{z,z'}\varphi_u(z) \leq 2\lambda_b(e(\sup_{t \in [0,T]} \langle \varphi_t, f_\infty \rangle + 1) - 1)$$

for each $u \in [0, T]$. Hence the above display becomes

$$S_{[0,T]}^Q(\varphi | \xi_Q^*) \geq \langle \xi, \vartheta_n \rangle - \langle \xi_Q^*, \vartheta_n \rangle - 2\lambda_b(e(\sup_{t \in [0,T]} \langle \varphi_t, f_\infty \rangle + 1) - 1)T.$$

As before, letting $n \to \infty$, using the monotone convergence theorem, and noting that $\xi_Q^* \in \mathcal{K}$, we conclude that $S_{[0,T]}^Q(\varphi | \xi_Q^*) = \infty$.

Since $\xi \notin \mathcal{K}$, $T > 0$, and $\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$ such that $\varphi_0 = \xi_Q^*$ and $\varphi_T = \xi$ are arbitrary, the proof of the lemma is complete. $\square$

We now prove the main result of this section, namely, that the quasipotential V_Q is not equal to the relative entropy $I(\cdot \| \xi_Q^*)$.

Proposition 8.1. *Let $\xi \in \mathcal{M}_1(\mathcal{Z})$ be such that $\langle \xi, f_\infty \rangle < \infty$ and $\xi \notin \mathcal{K}$. Then $I(\xi \| \xi_Q^*) < \infty$ and $V(\xi) = \infty$. In particular, $V \neq I(\cdot \| \xi_Q^*)$.*

Proof. By the Donsker–Varadhan variational formula (see Donsker and Varadhan [14, Lemma 2.1]), for any $\xi \in \mathcal{M}_1(\mathcal{Z})$ and any bounded function f on $\mathcal{Z}$, we have

$$I(\xi \| \xi_Q^*) \geq \langle \xi, f \rangle - \log\left(\sum_{z \in \mathcal{Z}} \exp\{f(z)\}\xi_Q^*(z)\right).$$

Recall the definition of f_n and f_∞ from the proof of Lemma 8.1. Let $\bar{\beta} > 0$ be such that $\sum_{z \in \mathcal{Z}} \exp\{\bar{\beta}z\}\xi_Q^*(z) < \infty$. Replacing f by $\bar{\beta}f_n$ in the above display, letting $n \to \infty$ and using the monotone convergence theorem, we arrive at

$$\bar{\beta}\langle \xi, f_\infty \rangle \leq I(\xi \| \xi_Q^*) + \log\left(\sum_{z \in \mathcal{Z}} \exp\{\bar{\beta}z\}\xi_Q^*(z)\right).$$

It follows that

$$\{\xi \in \mathcal{M}_1(\mathcal{Z}) : I(\xi \| \xi_Q^*) < \infty\} \subset \{\xi \in \mathcal{M}_1(\mathcal{Z}) : \langle \xi, f_\infty \rangle < \infty\}.$$

On the other hand, since $\langle \xi_Q^*, f_\infty \rangle < \infty$, it is easy to check that $\{\xi \in \mathcal{M}_1(\mathcal{Z}) : I(\xi \| \xi_Q^*) < \infty\} \supset \{\xi \in \mathcal{M}_1(\mathcal{Z}) : \langle \xi, f_\infty \rangle < \infty\}$.

Let $\xi \in \mathcal{M}_1(\mathcal{Z})$ be such that $\langle \xi, \vartheta \rangle = \infty$ and $\langle \xi, f_\infty \rangle < \infty$. Then the above yields $I(\xi \| \xi_Q^*) < \infty$. By Lemma 8.1, we see that $V_Q(\xi) = \infty$. This completes the proof of the proposition. □

8.2. A non-interacting WLAN system with constant forward rates

Recall the model described in Section 1.1.2. Define the quasipotential

$$V_W(\xi) := \inf\{S_{[0,T]}^W(\varphi | \xi_W^*), \varphi_0 = \xi_W^*, \varphi_T = \xi, T > 0\}, \quad \xi \in \mathcal{M}_1(\mathcal{Z}),$$

where S^W is defined by (2.8) with $\mathcal{E}$ replaced by $\mathcal{E}_W$ and L_ζ replaced by L^W for each $\zeta \in \mathcal{M}_1(\mathcal{Z})$. We now state the main result for this non-interacting WLAN.

Proposition 8.2. *Let $\xi \in \mathcal{M}_1(\mathcal{Z})$ be such that $\langle \xi, f_\infty \rangle < \infty$ and $\xi \notin \mathcal{K}$. Then $I(\xi \| \xi_W^*) < \infty$ and $V(\xi) = \infty$. In particular, $V_W \neq I(\cdot \| \xi_W^*)$.*

We have the following lemma. The proof is similar to the proof of Lemma 8.1, noting that $\langle \xi_W^*, \vartheta \rangle < \infty$, and it is left to the reader.

Lemma 8.2. *If $\xi \in \mathcal{M}_1(\mathcal{Z})$ is such that $\xi \notin \mathcal{K}$, then $V_W(\xi) = \infty$.*

Using the above lemma, we can now prove Proposition 8.2 along similar lines to the proof of Proposition 8.1 in the previous section.

Appendix A. Proofs of Section 2

A.1. Compactness of level sets of $S_{[0,T]}$

Proof of Lemma 2.1. Fix $T > 0$, $s > 0$, and $K \subset \mathcal{M}_1(\mathcal{Z})$ compact. Given $\nu \in K$, $\varphi \in \Phi_\nu^{[0,T]}(s)$, and a finite set $B \subset \mathcal{Z}$, if we choose $f(t, z) = \mathbf{1}_{\{z \in B\}}$ for all $t \in [0, T]$, then (2.6) yields

$$\varphi_t(B) - \varphi_r(B) = \int_{[r,t]} \sum_{(z,z') \in \mathcal{E}} (f(z') - f(z))(1 + h_\varphi(u, z, z'))\lambda_{z,z'}(\varphi_u)\varphi_u(z)du$$

for all $0 \leq r < t \leq T$. Note that we may take $h_\varphi \geq -1$; otherwise the rate function would be infinite as per (2.7) and the definition of τ^* in (2.4). Therefore, we get

$$|\varphi_t(B) - \varphi_r(B)| \leq \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} (1 + h_\varphi(u, z, z')) \times \mathbf{1}_{\{u \in [r,t]\}}\lambda_{z,z'}(\varphi_u)\varphi_u(z)du. \tag{A.1}$$

Noting that

$$\sup\left\{\int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} \tau^*(h_\varphi(u, z, z'))\lambda_{z,z'}(\varphi_u)\varphi_u(z)du, \varphi \in \Phi_\nu^{[0,T]}(s), \nu \in K\right\} \leq s,$$

it follows that the family $\{1 + h_\varphi, \varphi \in \Phi_\nu^{[0,T]}(s), \nu \in K\}$ is uniformly integrable. That is,

$$\sup\left\{\int_{[0,T]} (1 + h_\varphi(u, z, z')) \times \mathbf{1}_{\{1+h_\varphi \geq M\}}\lambda_{z,z'}(\varphi_u)\varphi_u(z)du, \varphi \in \Phi_\nu^{[0,T]}(s), \nu \in K\right\} \to 0$$

as $M \to \infty$. Hence for any $M > 0$, using the boundedness of the transition rates (from Assumption (A2)), (A.1) yields

$$
|\varphi_t(B) - \varphi_r(B)|
$$
$$
\leq 2M\bar{\lambda}(t-r) + \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} (1 + h_\varphi(u, z, z')) \times \mathbf{1}_{\{1+h_\varphi \geq M\}} \lambda_{z,z'}(\varphi_u)\varphi_u(z) du
$$

for all $0 \leq r < t \leq T$, and $B \subset \mathcal{M}_1(\mathcal{Z})$. It follows that

$$
\sup_{\varphi \in \cup_{v \in K} \Phi_v^{[0,T]}(s)} \sup_{t,r \,:\, |t-r| \leq \delta} d(\varphi_t, \varphi_r)
$$
$$
\leq 2M\bar{\lambda}\delta + \sup_{\varphi \in \cup_{v \in K} \Phi_v^{[0,T]}(s)} \sup_{t,r \,:\, |t-r| \leq \delta} \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} (1 + h_\varphi(u, z, z'))
$$
$$
\times \mathbf{1}_{\{1+h_\varphi \geq M\}} \lambda_{z,z'}(\varphi_u)\varphi_u(z) du.
$$

Letting $\delta \to 0$ first and then $M \to \infty$, we arrive at

$$
\lim_{\delta \downarrow 0} \sup_{\varphi \in \cup_{v \in K} \Phi_v^{[0,T]}(s)} \sup_{t,r \,:\, |t-r| \leq \delta} d(\varphi_t, \varphi_r) = 0.
$$

Hence it follows that $\cup_{v \in K} \Phi_v^{[0,T]}(s)$ is precompact in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ (see, e.g., Billingsley [4, Theorem 12.3]).

To show that $\cup_{v \in K} \Phi_v^{[0,T]}(s)$ is closed, let $\{\varphi_n, n \geq 1\} \subset \cup_{v \in K} \Phi_v^{[0,T]}(s)$ and suppose that $\varphi_n \to \bar{\varphi}$ in $D([0, T], \mathcal{M}_1(\mathcal{Z}))$. Note that, for any $f \in C_0^1([0, T] \times \mathcal{M}_1(\mathcal{Z}))$, the mapping

$$
\varphi \mapsto \left\{ \langle \varphi_T, f_T \rangle - \langle \varphi_0, f_0 \rangle - \int_{[0,T]} \langle \varphi_u, \partial_u f_u \rangle du \right.
$$
$$
\left. - \int_{[0,T]} \langle \varphi_u, L_{\varphi_u} f_u \rangle du - \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} \tau(f_u(z') - f_u(z))\lambda_{z,z'}(\varphi_u)\varphi_u(z) du \right\}
$$

is continuous on $D([0, T], \mathcal{M}_1(\mathcal{Z}))$, and hence the mapping

$$
\varphi \mapsto \sup_{f \in C_0^1([0,T] \times \mathcal{Z})} \left\{ \langle \varphi_T, f_T \rangle - \langle \varphi_0, f_0 \rangle - \int_{[0,T]} \langle \varphi_u, \partial_u f_u \rangle du \right.
$$
$$
\left. - \int_{[0,T]} \langle \varphi_u, L_{\varphi_u} f_u \rangle du - \int_{[0,T]} \sum_{(z,z') \in \mathcal{E}} \tau(f_u(z') - f_u(z))\lambda_{z,z'}(\varphi_u)\varphi_u(z) du \right\}
$$

is lower semicontinuous on $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ (see, e.g., Berge [1, Theorem 1, p. 115]). Hence

$$
S_{[0,T]}(\bar{\varphi}|\bar{\varphi}(0)) \leq \liminf_{n \to \infty} S_{[0,T]}(\varphi_n|\varphi_n(0)) \leq s,
$$

and it follows that $\cup_{v \in K} \Phi_v^{[0,T]}(s)$ is closed. Consequently, $\cup_{v \in K} \Phi_v^{[0,T]}(s)$ is a compact subset of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$. $\qquad \square$

A.2. Uniform LDP for $\mu_{\nu_N}^N$: proof of Theorem 2.1

In this section we prove Theorem 2.1. In the case of a finite state space (i.e., when $\mathcal{Z}$ is a finite set), the LDP for the family $\{\mu_{\nu_N}^N, N \geq 1\}$, whenever $\nu_N \to \nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$, was proved in [7, Theorem 3.1] under suitable assumptions. The main assumption required in the proof of [7, Theorem 3.1] was the boundedness of the 'total outgoing jump rate' across all the states, which also holds in our countable-state-space case under Assumptions (A1)–(A3). So, to prove the LDP for the family $\{\mu_{\nu_N}^N, N \geq 1\}$, whenever $\nu_N \to \nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$, one can go through the steps in [7, Section 5] verbatim; we reproduce the important steps here for the sake of completeness. Once this LDP is proved, we then show the uniform LDP over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ using [8, Propositions 1.12 and 1.14].

A.2.1. *LDP for* $\{\mu_{\nu_N}^N, N \geq 1\}$ *when* $\nu_N \to \nu$ *in* $\mathcal{M}_1(\mathcal{Z})$. We first introduce some notation. Let $\{(X_n^N(t), t \in [0, T]), 1 \leq n \leq N\}$ denote the joint evolution of the states of all the particles. This is a Markov process on $\mathcal{Z}^N$ with the infinitesimal generator acting on functions f on $\mathcal{Z}^N$ given by

$$(z_1, \ldots, z_N) \mapsto \sum_{n=1}^N \sum_{z_n' \in \{z_n+1, 0\}} (f(z_1, \ldots, z_n', \ldots, z_N) - f(z_1, \ldots, z_N)) \lambda_{z_n, z_n'}(\text{emp}(z_1, \ldots, z_N)),$$

where $\text{emp}(z_1, \ldots, z_N) := \frac{1}{N} \sum_{n=1}^N \delta_{z_n} \in \mathcal{M}_1^N(\mathcal{Z})$. Define the empirical measure

$$\Theta^N := \frac{1}{N} \sum_{n=1}^N \delta_{X_n^N(\cdot)};$$

Θ^N is an $\mathcal{M}_1(D([0, T], \mathcal{Z}))$-valued random variable. Let $\sigma : \mathcal{M}_1(D([0, T], \mathcal{Z})) \to D([0, T], \mathcal{M}_1(\mathcal{Z}))$ denote the canonical projection map. Note that $\mu^N(t) = \sigma_t(\Theta^N)$, $t \in [0, T]$. Similarly, let $\{(\bar{X}_n^N(t), t \in [0, T]), 1 \leq n \leq N\}$ denote the evolution of the independent particles, where each particle executes a Markov process with the infinitesimal generator $\bar{L}$ defined in (3.1). Define the corresponding empirical measure $\bar{\Theta}^N$ by

$$\bar{\Theta}^N := \frac{1}{N} \sum_{n=1}^N \delta_{\bar{X}_n^N(\cdot)}.$$

Let $\mathcal{P}_{\nu_N}^N$ (resp. $\bar{\mathcal{P}}_{\nu_N}^N$) denote the law of Θ^N (resp. $\bar{\Theta}^N$) with initial condition $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$ (i.e., $\frac{1}{N} \sum_{n=1}^N \delta_{X_n^N(0)} = \nu_N$). These are probability measures on $\mathcal{M}_1(D([0, T], \mathcal{Z}))$; i.e., $\mathcal{P}_{\nu_N}^N, \bar{\mathcal{P}}_{\nu_N}^N \in \mathcal{M}_1(\mathcal{M}_1(D([0, T], \mathcal{Z})))$.

Note that $\mathcal{P}_{\nu_N}^N \ll \bar{\mathcal{P}}_{\nu_N}^N$. For $x \in D([0, T], \mathcal{Z})$ and $\mu \in D([0, T], \mathcal{M}_1(\mathcal{Z}))$, define

$$h(x; \mu) := \sum_{t \in [0, T]} \mathbf{1}_{\{x(t) \neq x(t-)\}} \log \left(\frac{\lambda_{x(t-), x(t)}(\mu_t)}{\widetilde{\lambda}_{x(t-), x(t)}} \right) - \int_{[0, T]} \sum_{\substack{z' \in \mathcal{Z}: \\ (x(t-), z') \in \mathcal{E}}} \left(\lambda_{x(t-), z'}(\mu_t) - \widetilde{\lambda}_{x(t-), z'} \right) dt,$$

where $\{\widetilde{\lambda}_{z, z'}, (z, z') \in \mathcal{E}\}$ are the non-interacting rates defined by

$$\widetilde{\lambda}_{z, z'} := \begin{cases} \overline{\lambda}/(z+1) & \text{if } z' = z+1, \\ \underline{\lambda} & \text{if } z' = 0, z \geq 1. \end{cases}$$

Also, define

$$h(Q) := \int_{D([0,T],\mathcal{Z})} h(\,\cdot\,;\sigma(Q))\,dQ, \qquad Q \in \mathcal{M}_1(D([0,T],\mathcal{Z})). \tag{A.2}$$

Using Girsanov's theorem, it is straightforward to check that

$$\frac{d\mathcal{P}^N_{\nu_N}}{d\bar{\mathcal{P}}^N_{\nu_N}}(Q) = \exp\{Nh(Q)\}, \quad Q \in \mathcal{M}_1(D([0,T],\mathcal{Z})).$$

We now introduce some notation related to path spaces. Define $\psi : D([0,T],\mathcal{Z}) \to \{0,1,\ldots\}$ by

$$\psi(x) = \sum_{t \in [0,T]} \mathbf{1}_{\{x(t) \neq x(t-)\}};$$

$\psi(x)$ is the number of discontinuities in x. Since $\mathcal{Z}$ is a countable set, it follows that $\psi(x) < \infty$ for all $x \in D([0,T],\mathcal{Z})$ ([4, Chapter 3, Lemma 1]). Define

$$\mathcal{X} := \{x \in D([0,T],\mathcal{Z}) : \psi(x) < \infty, (x(t-),x(t)) \in \mathcal{E} \text{ whenever } x(t) \neq x(t-),\ t \in [0,T]\},$$

and equip $\mathcal{X}$ with the subspace topology. Since $\mathcal{Z}$ is countable, we have that ψ is continuous on $\mathcal{X}$. Define

$$\|f\|_\psi := \sup_{x \in \mathcal{X}} \frac{|f(x)|}{1 + \psi(x)}, \qquad \text{for } f : \mathcal{X} \to \mathbb{R}.$$

Then, define

$$\mathcal{C}_\psi(\mathcal{X}) := \{f : \mathcal{X} \to \mathbb{R} \text{ such that } f \text{ is continuous and } \|f\|_\psi < \infty\}$$

and

$$\mathcal{M}_{1,\psi}(\mathcal{X}) := \left\{ Q \in \mathcal{M}_1(\mathcal{X}) : \int_{\mathcal{X}} \psi \, dQ < \infty \right\}.$$

$\mathcal{M}_{1,\psi}(\mathcal{X})$ is a subset of $\mathcal{C}_\psi(\mathcal{X})^*$, the algebraic dual of $\mathcal{C}_\psi(\mathcal{X})$, and we equip it with the weak* topology. This is the coarsest topology on $\mathcal{M}_{1,\psi}(\mathcal{X})$ where we say $Q_N \to Q$ in $\mathcal{M}_{1,\psi}(\mathcal{X})$ as $N \to \infty$ if and only if

$$\int_{\mathcal{X}} f \, dQ_N \to \int_{\mathcal{X}} f \, dQ \text{ as } N \to \infty, \quad \text{for all } f \in \mathcal{C}_\psi(\mathcal{X}).$$

Recall $\bar{P}_z$, $z \in \mathcal{Z}$, from Section 3. For each $\nu \in \mathcal{M}_1(\mathcal{Z})$, define $J : \mathcal{M}_1(\mathcal{X}) \to [0,\infty]$ by

$$J(Q) := \sup_{f \in \mathcal{C}_\psi(\mathcal{X})} \left[\int_{\mathcal{X}} f \, dQ - \sum_{z \in \mathcal{Z}} \nu(z) \log \int_{\mathcal{X}} \exp\{f\} \, d\bar{P}_z \right]. \tag{A.3}$$

By [7, Lemma 5.3], we also have

$$J(Q) = \sup_{f \in \mathcal{C}_b(\mathcal{X})} \left[\int_{\mathcal{X}} f \, dQ - \sum_{z \in \mathcal{Z}} \nu(z) \log \int_{\mathcal{X}} \exp\{f\} \, d\bar{P}_z \right], \tag{A.4}$$

where $\mathcal{C}_b(\mathcal{X})$ is the space of bounded and continuous functions on $\mathcal{X}$ equipped with the supremum norm.

We first state a lemma for the LDP for the family $\{\bar{\mathcal{P}}_{\nu_N}^N, N \geq 1\}$ on $\mathcal{M}_{1,\psi}(\mathcal{X})$ whenever $\nu_N \to \nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$. Its proof follows verbatim from that of [7, Lemma 5.1].

Lemma A.1. (LDP for the non-interacting system [7, Lemma 5.1].) *Let $\nu_N \to \nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$. Then the family $\{\bar{\mathcal{P}}_{\nu_N}^N, N \geq 1\}$ satisfies the LDP on $\mathcal{M}_{1,\psi}(\mathcal{X})$ with rate function J defined in (A.3).*

Next, we provide two necessary conditions for the finiteness of J as defined in (A.3).

Lemma A.2. (Finiteness of J [7, Lemma 5.2].) *If $J(Q) < \infty$, then we have $Q \in \mathcal{M}_{1,\psi}(\mathcal{X})$ and $Q \circ \sigma_0^{-1} = \nu$.*

Proof. Let Q be such that $J(Q) < \infty$. The proof of $Q \circ \sigma_0^{-1} = \nu$ follows verbatim from [7, Lemma 5.2]. For the first assertion, since $\psi \in \mathcal{C}_\psi(\mathcal{X})$, from the definition of J in (A.3), we have

$$J(Q) \geq \int_{\mathcal{X}} \psi \, dQ - \sum_{z \in \mathcal{Z}} \nu(z) \log \int_{\mathcal{X}} \exp\{\psi\} \, d\bar{P}_z. \tag{A.5}$$

Note that, for each $z \in \mathcal{Z}$, under $\bar{P}_z$ (see (3.1)), the number of jumps on $[0, T]$ is stochastically dominated by a Poisson random variable with parameter $(\underline{\lambda} + \bar{\lambda})T \leq 2\bar{\lambda}T$. Therefore,

$$\int_{\mathcal{X}} \exp\{\psi\} \, d\bar{P}_z \leq \sum_{k \geq 0} \exp\{k\} \frac{\exp\{-2\bar{\lambda}T\}(2\bar{\lambda}T)^k}{k!} = c_1 < \infty,$$

where c_1 is some constant independent of z. Therefore,

$$\sum_{z \in \mathcal{Z}} \nu(z) \log \int_{\mathcal{X}} \exp\{\psi\} d\bar{P}_z < \infty.$$

Hence, from (A.5), using $J(Q) < \infty$, we conclude that

$$\int_{\mathcal{X}} \psi \, dQ < \infty.$$

It follows that $Q \in \mathcal{M}_{1,\psi}(\mathcal{X})$. $\qquad\qquad\square$

The next lemma is required to prove the continuity of h on $\mathcal{M}_{1,\psi}(\mathcal{X})$.

Lemma A.3. (See [7, Lemma 5.7] for the finite-state-space case.) *Suppose that $Q \in \mathcal{M}_1(\mathcal{X})$ is such that $J(Q) < \infty$. Then*

$$\lim_{\alpha \to 0} \sup_{t \in [0,T]} \int_{\mathcal{X}} \sup_{u \in [t-\alpha, t+\alpha] \cap [0,T]} \mathbf{1}_{\{X(u) \neq X(u-)\}} \, dQ(X) = 0.$$

Proof. Let $P \in \mathcal{M}_1(\mathcal{X})$ denote the mixture distribution defined by $dP := \sum_{z \in \mathcal{Z}} \nu(z) d\bar{P}_z$. Since $J(Q) < \infty$, it follows that $Q \ll P$. Indeed, using Jensen's inequality, we have

$$\sum_{z \in \mathcal{Z}} \log \int_{\mathcal{X}} \exp\{f\} \, dP_z \leq \log \int_{\mathcal{X}} \exp\{f\} \, dP \qquad \text{for any } f \in \mathcal{C}_b(\mathcal{X}),$$

and hence, from (A.4) and the Donsker–Varadhan variational formula for $I(Q\|P)$, we conclude that

$$I(Q\|P) \leq J(Q). \tag{A.6}$$

Since $J(Q) < \infty$, the above implies that $I(Q\|P) < \infty$. This shows $Q \ll P$. Hence, with $K_{t,\alpha} = \{x \in \mathcal{X} : x(u) \neq x(u-) \text{ for some } u \in [t-\alpha, t+\alpha] \cap [0, T]\}$, we have

$$\int_{\mathcal{X}} \sup_{u \in [t-\alpha, t+\alpha] \cap [0,T]} \mathbf{1}_{\{X(u) \neq X(u-)\}} \, dQ(X) = Q(K_{t,\alpha})$$
$$= \int_{\mathcal{X}} \left(\frac{dQ}{dP}\right) \mathbf{1}_{\{K_{t,\alpha}\}} \, dP$$
$$\leq \left\|\left(\frac{dQ}{dP}\right)\right\|_{\tau^*,P} \|\mathbf{1}_{\{K_{t,\alpha}\}}\|_{\tau,P}, \tag{A.7}$$

where the last inequality follows from the Hölder inequality in Orlicz spaces. Here, $\|\cdot\|_{\tau,P}$ is the Orlicz norm defined by

$$\|f\|_{\tau,P} := \inf\left\{a > 0 : \int_{\mathcal{X}} \tau\left(\frac{|f(x)|}{a}\right) \, dP(x) \leq 1\right\}.$$

Similarly, $\|f\|_{\tau^*,P}$ is defined as above with τ replaced by τ^*.

Consider $\left\|\left(\frac{dQ}{dP}\right)\right\|_{\tau^*,P}$. Note that there exists a $u_0 \geq 1$ such that $\tau^*(u) \leq 2u \log u$ for all $u \geq u_0$. Therefore,

$$\int_{\mathcal{X}} \tau^*\left(\frac{dQ}{dP}\right) \, dP \leq \tau^*(u_0) + 2\int_{\mathcal{X}} \left(\frac{dQ}{dP}\right) \log\left(\frac{dQ}{dP}\right) \, dP$$
$$= \tau^*(u_0) + 2I(Q\|P)$$
$$\leq \tau^*(u_0) + 2J(Q)$$
$$< \infty,$$

where the second inequality follows from (A.6) and the third inequality follow from the assumption that $J(Q) < \infty$. Since $\tau^*(u/a) \leq \tau^*(u)/a$ for $a \geq 1$ (by Jensen's inequality), this shows that

$$\left\|\left(\frac{dQ}{dP}\right)\right\|_{\tau^*,P} < c_2 < \infty \quad \text{for some } c_2 \text{ that does not depend on } t. \tag{A.8}$$

Next, consider $\|\mathbf{1}_{\{K_{t,\alpha}\}}\|_{\tau^*,P}$. Note that, under P, the number of jumps in $[t-\alpha, t+\alpha] \cap [0, T]$ is stochastically dominated by a Poisson random variable with parameter $2\alpha(\bar{\lambda} + \underline{\lambda}) \leq 4\alpha\bar{\lambda}$. Therefore, $P(K_{t,\alpha}) \leq 1 - \exp\{-4\alpha\bar{\lambda}\} \leq 4\alpha\bar{\lambda}$. Since $\tau(\mathbf{1}_{\{K_{t,\alpha}\}}/a) = \tau(1/a)\mathbf{1}_{\{K_{t,\alpha}\}}$ for any $a > 0$, we have

$$\int_{\mathcal{X}} \tau\left(\mathbf{1}_{\{K_{t,\alpha}\}}/a\right) \, dP = \tau(1/a)P(K_{t,\alpha}) \leq \tau(1/a)4\alpha\bar{\lambda}.$$

Therefore, if we choose $a = 1/(\tau^{-1}(1/4\alpha\bar{\lambda}))$, the right-hand side of the above display becomes 1. This shows that

$$\|\mathbf{1}_{\{K_{t,\alpha}\}}\|_{\tau^*,P} \leq \frac{1}{\tau^{-1}(1/4\alpha\bar{\lambda})}$$

for all t. Hence, by (A.7), (A.8), and the previous display, we get

$$\sup_{t\in[0,T]}\int_{\mathcal{X}}\sup_{u\in[t-\alpha,t+\alpha]\cap[0,T]}\mathbf{1}_{\{X(u)\neq X(u-)\}}\,dQ(X)\leq\frac{c_2}{\tau^{-1}(1/4\alpha\overline{\lambda})}\to 0\quad\text{as }\alpha\to 0.$$

This completes the proof of the lemma.   □

Next, we argue the continuity of the projection map σ.

Lemma A.4. (Continuity of σ [7, Lemma 5.8].) *Let $Q\in\mathcal{M}_1(\mathcal{X})$ be such that $J(Q)<\infty$. Then $\sigma:\mathcal{M}_1(D([0,T],\mathcal{Z}))\to D([0,T],\mathcal{M}_1(\mathcal{Z}))$ is continuous at Q.*

Proof. Let $Q\in\mathcal{M}_1(\mathcal{X})$ be such that $J(Q)<\infty$. By Lemma A.2, it follows that $Q\in\mathcal{M}_{1,\psi}(\mathcal{X})$. In [21, Lemma 2.8], for the case when $\nu=\delta_{z_0}$ for some $z_0\in\mathcal{Z}$, it is shown that $\sigma:\mathcal{M}_1(D([0,T],\mathcal{Z}))\to D([0,T],\mathcal{M}_1(\mathcal{Z}))$ is continuous at Q whenever $Q\in\mathcal{M}_{1,\psi}(\mathcal{X})$. (This continuity is shown in [21] when $\mathcal{M}_1(D([0,T],\mathcal{Z}))$ is equipped with the usual weak topology and $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ is equipped with the stronger uniform topology. Since the Skorokhod topology on $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ is coarser than the uniform topology, it follows that σ is continuous.) For general $\nu\in\mathcal{M}_1(\mathcal{Z})$, by using the result of Lemma A.3 and following the proof of [21, Lemma 2.8] verbatim, we arrive at the continuity of σ.   □

Finally, we have that h is continuous on $\mathcal{M}_{1,\psi}(\mathcal{X})$.

Lemma A.5. *Assume* (A1)*,* (A2)*, and* (A3)*. Then the mapping h defined in* (A.2) *is continuous on $\mathcal{M}_{1,\psi}(\mathcal{X})$.*

Proof. Using Lemma A.3, Lemma A.4, and Assumptions (A1)–(A3), the proof of [21, Lemma 2.9] holds verbatim.   □

The above lemmas give us the LDP for the family $\{\mu_{\nu_N}^N, N\geq 1\}$ on $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ whenever $\nu_N\to\nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N\to\infty$.

Proposition A.1. *Assume* (A1)*,* (A2)*, and* (A3)*. Suppose that $\nu_N\to\nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N\to\infty$. Then the family $\{\mu_{\nu_N}^N, N\geq 1\}$ satisfies the LDP on $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ with rate function $S_{[0,T]}(\cdot|\nu)$ defined in* (2.5)*.*

Proof. Let $\nu_N\to\nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N\to\infty$. By Lemma A.1, we have that $\{\bar{\mathcal{P}}_{\nu_N}^N, N\geq 1\}$ satisfies the LDP on $\mathcal{M}_{1,\psi}(\mathcal{X})$ with rate function J. Since h is continuous on the set $\{Q\in\mathcal{M}_{1,\psi}(\mathcal{X}):J(Q)<\infty\}$ (by Lemma A.5), from Varadhan's lemma, one can conclude (see the proof of [7, Theorem 3.1]) that the family $\{\mathcal{P}_{\nu_N}^N\}$ satisfies the LDP on $\mathcal{M}_{1,\psi}(\mathcal{X})$ with rate function $Q\mapsto J(Q)-h(Q)$. By Lemma A.4, since σ is continuous (with the usual weak topology on $\mathcal{M}_1(D([0,T],\mathcal{Z}))$) at Q when $J(Q)<\infty$, it follows that the restriction of σ to $\mathcal{M}_{1,\psi}(\mathcal{X})$ is also continuous (with respect to the stronger topology on $\mathcal{M}_{1,\psi}(\mathcal{X})$) at Q when $J(Q)<\infty$. Therefore, using the generalised contraction principle (e.g., [13, Theorem 4.2.23]), the LDP for the family $\{\mu_{\nu_N}^N, N\geq 1\}$ on $D([0,T],\mathcal{M}_1(\mathcal{Z}))$ follows. The rate function for this LDP can be shown to admit the form given in (2.5) (see, e.g., the proof of [21, Theorem 3.1]).   □

A.2.2. *Uniform LDP for $\{\mu_{\nu_N}^N, N\geq 1\}$ over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$.* Proposition A.1 establishes the LDP for the family $\{\mu_{\nu_N}^N, N\geq 1\}$ whenever $\nu_N\to\nu$ in $\mathcal{M}_1(\mathcal{Z})$ as $N\to\infty$. We now extend this to the uniform LDP on the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$. Towards this, we rely on [8, Propositions 1.12 and 1.14]. Although our definition of the uniform LDP (Definition 2.2) has initial conditions lying in $A\cap\mathcal{M}_1^N(\mathcal{Z})$ (unlike the definition of the uniform LDP in [8, Definition 1.13], where the initial conditions do not depend on the

parameter N), we can use straightforward modifications of the arguments in [8, Propositions 1.12 and 1.14] to prove the desired uniform LDP. We provide an outline of these arguments here.

We first provide a definition of the uniform Laplace principle over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$. Recall the definition of the rate function $S_{[0,T]}$ in (2.5). For $v \in \mathcal{M}_1(\mathcal{Z})$ and $g \in C_b(D([0, T], \mathcal{M}_1(\mathcal{Z})))$, define

$$F(v, g) := - \inf_{\varphi \in D([0,T], \mathcal{M}_1(\mathcal{Z}))} \left[g(\varphi) + S_{[0,T]}(\varphi|v) \right].$$

Definition A.1. We say that the family $\{\mu_{v_N}^N, N \geq 1\}$ of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$-valued random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies the uniform Laplace principle over the class $\mathcal{A}$ of subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|v), v \in \mathcal{M}_1(\mathcal{Z})\}$, $S_{[0,T]}(\cdot|v): D([0, T], \mathcal{M}_1(\mathcal{Z})) \to [0, +\infty]$, $v \in \mathcal{M}_1(\mathcal{Z})$, if the following hold:

- (Compactness of level sets.) For each $K \subset \mathcal{M}_1(\mathcal{Z})$ compact and $s \geq 0$, $\bigcup_{v \in K} \Phi_v(s)$ is a compact subset of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$, where $\Phi_v(s) := \{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi_0 = v, S_{[0,T]}(\varphi|v) \leq s\}$;

- (Uniform Laplace asymptotics.) For any $A \in \mathcal{A}$ and $g \in C_b(D([0, T], \mathcal{M}_1(\mathcal{Z})))$, we have

$$\lim_{N \to \infty} \sup_{v_N \in A \cap \mathcal{M}_1^N(\mathcal{Z})} \left| \frac{1}{N} \log \mathbb{E}_{v_N} \left[\exp\{-Ng(\mu_{v_N}^N)\} \right] - F(v_N, g) \right| = 0.$$

This is a modification of [8, Definition 1.11] to the case when the initial conditions are only allowed to lie in $A \cap \mathcal{M}_1^N(\mathcal{Z})$. We have the following result.

Lemma A.6. (See [8, Proposition 1.12].) *Assume* (A1)*,* (A2)*, and* (A3)*. Then the family* $\{\mu_{v_N}^N, N \geq 1\}$ *satisfies the uniform Laplace principle over the class of compact subsets of* $\mathcal{M}_1(\mathcal{Z})$ *with the family of rate functions* $\{S_{[0,T]}(\cdot|v), v \in \mathcal{M}_1(\mathcal{Z})\}$*,* $S_{[0,T]}(\cdot|v): D([0, T], \mathcal{M}_1(\mathcal{Z})) \to [0, +\infty]$*,* $v \in \mathcal{M}_1(\mathcal{Z})$*.*

Proof. By Lemma 2.1, we have that for each $K \subset \mathcal{M}_1(\mathcal{Z})$ compact and $s \geq 0$, $\bigcup_{v \in K} \Phi_v(s)$ is a compact subset of $D([0, T], \mathcal{M}_1(\mathcal{Z}))$, where $\Phi_v(s) := \{\varphi \in D([0, T], \mathcal{M}_1(\mathcal{Z})) : \varphi_0 = v, I_v(\varphi) \leq s\}$.

To show the uniform Laplace asymptotics, let $g \in C_b(D([0, T], \mathcal{M}_1(\mathcal{Z})))$. By Proposition A.1, whenever $v_N \to v$ in $\mathcal{M}_1(\mathcal{Z})$ as $N \to \infty$, we have that the family $\{\mu_{v_N}^N, N \geq 1\}$ satisfies the LDP on $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ with rate function $S_{[0,T]}(\cdot|v)$. Therefore, by Varadhan's lemma (e.g., [13, Theorem 4.3.1]), we have

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}_{v_N} \left[\exp\{-Ng(\mu_{v_N}^N)\} \right] = F(v, g). \tag{A.9}$$

Define

$$F^N(v_N', g) := \frac{1}{N} \log \mathbb{E}_{v_N'} \left[\exp\{-Ng(\mu_{v_N'}^N)\} \right], \quad v_N' \in \mathcal{M}_1^N(\mathcal{Z}).$$

Using (A.9), we now show that the mapping $v \mapsto F(v, g)$ is continuous. To show this continuity, it suffices to show that given any $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $v' \in \mathcal{M}_1(\mathcal{Z})$ such that $d(v', v) < \delta$ and $v_N' \in \mathcal{M}_1^N(\mathcal{Z})$ such that $v_N' \to v'$ as $N \to \infty$, we have

$$|F^N(v_N', g) - F(v, g)| < \varepsilon \quad \text{for all large enough } N.$$

Indeed, if this is true, sending $N \to \infty$ in the above display and using (A.9), we arrive at $|F(\nu', g) - F(\nu, g)| < \varepsilon$, which shows the continuity of $\nu \mapsto F(\nu, g)$. We now prove the above statement using contraposition. Suppose the above statement is not true. Then there exist $\varepsilon > 0$ and a sequence $\{\nu_N\}$ with $\nu_N \in \mathcal{M}_1^N(\mathcal{Z})$ and $\nu_N \to \nu$ as $N \to \infty$ such that $|F^N(\nu_N, g) - F(\nu, g)| > \varepsilon$. Using (A.9), we get $|F(\nu, g) - F(\nu, g)| > \varepsilon > 0$, which is a contradiction. This establishes the continuity of the mapping $\mathcal{M}_1(\mathcal{Z}) \ni \nu \mapsto F(\nu, g)$.

Since $\nu \mapsto F(\nu, g)$ is continuous, using (A.9), by the same arguments as in [8, Proposition 1.12], one can show that for any compact subset K of $\mathcal{M}_1(\mathcal{Z})$, we have $\sup_{\nu_N \in K \cap \mathcal{M}_1^N(\mathcal{Z})} |F^N(\nu_N, g) - F(\nu, g)| \to 0$ as $N \to \infty$. This shows that the family $\{\mu_{\nu_N}^N, N \geq 1\}$ satisfies the uniform Laplace principle over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|\nu), \nu \in \mathcal{M}_1(\mathcal{Z})\}$. □

We can now complete the proof of Theorem 2.1 using the arguments in [8, Proposition 1.14].

Proof of Theorem 2.1. By Lemma A.6, the family $\{\mu_{\nu_N}^N, N \geq 1\}$ satisfies the uniform Laplace principle $D([0, T], \mathcal{M}_1(\mathcal{Z}))$ over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|\nu), \nu \in \mathcal{M}_1(\mathcal{Z})\}$. Restricting the initial conditions to $\mathcal{M}_1^N(\mathcal{Z})$ and following the proof of [8, Proposition 1.14] verbatim, we conclude that the family $\{\mu_{\nu_N}^N, N \geq 1\}$ satisfies the uniform LDP on $\mathcal{M}_1(D([0, T], \mathcal{Z}))$ over the class of compact subsets of $\mathcal{M}_1(\mathcal{Z})$ with the family of rate functions $\{S_{[0,T]}(\cdot|\nu), \nu \in \mathcal{M}_1(\mathcal{Z})\}$. □

Acknowledgements

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BERGE, C. (1997). *Topological Spaces: Including a Treatment of Multi-valued Functions, Vector Spaces, and Convexity*. Dover, Mineola, NY.

[2] BERTINI, L. *et al.* (2002). Macroscopic fluctuation theory for stationary non-equilibrium states. *J. Statist. Phys.* **107**, 635–675.

[3] BERTINI, L. *et al.* (2003). Large deviations for the boundary driven symmetric simple exclusion process. *Math. Phys. Anal. Geom.* **6**, 231–267.

[4] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.

[5] BODINEAU, T. AND GIACOMIN, G. (2004). From dynamic to static large deviations in boundary driven exclusion particle systems. *Stoch. Process. Appl.* **110**, 67–81.

[6] BORDENAVE, C., MCDONALD, D. AND PROUTIERE, A. (2010). A particle system in interaction with a rapidly varying environment: mean field limits and applications. *Networks Heterog. Media* **5**, 31–62.

[7] BORKAR, V. S. AND SUNDARESAN, R. (2012). Asymptotics of the invariant measure in mean field models with jumps. *Stoch. Systems* **2**, 322–380.

[8] BUDHIRAJA, A. AND DUPUIS, P. (2019). *Analysis and Approximation of Rare Events*. Springer, New York.

[9] CERRAI, S. AND PASKAL, N. (2022). Large deviations principle for the invariant measures of the 2D stochastic Navier–Stokes equations with vanishing noise correlation. *Stoch. Partial Differential Equat.* **10**, 1651–1681.

[10] CERRAI, S. AND RÖCKNER, M. (2004). Large deviations for stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term. *Ann. Prob.* **32**, 1100–1139.

[11] CERRAI, S. AND RÖCKNER, M. (2005). Large deviations for invariant measures of stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term. *Ann. Inst. H. Poincaré Prob. Statist.* **41**, 69–105.

[12] DAWSON, D. A. AND GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20**, 247–308.

[13] DEMBO, A. AND ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications*, 2nd edn. Springer, Berlin, Heidelberg.

[14] DONSKER, M. D. AND VARADHAN, S. R. S. (1975). Asymptotic evaluation of certain Markov process expectations for large time, I. *Commun. Pure Appl. Math.* **28**, 1–47.

[15] DURRETT, R. (2019). *Probability: Theory and Examples*, 5th edn. Cambridge University Press.

[16] ETHIER, S. N. AND KURTZ, T. G. (2005). *Markov Processes: Characterization and Convergence*. John Wiley, New York.

[17] FARFÁN, J., LANDIM, C. AND TSUNODA, K. (2019). Static large deviations for a reaction–diffusion model. *Prob. Theory Relat. Fields* **174**, 49–101.

[18] FREIDLIN, M. I. AND WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd edn. Springer, Berlin, Heidelberg.

[19] KHASMINSKII, R. (2012). *Stochastic Stability of Differential Equations*. Springer, Berlin, Heidelberg.

[20] KUMAR, A., ALTMAN, E., MIORANDI, D. AND GOYAL, M. (2006). New insights from a fixed point analysis of single cell IEEE 802.11 WLANs. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, pp. 1550–1561.

[21] LÉONARD, C. (1995). Large deviations for long range interacting particle systems with jumps. *Ann. Inst. H. Poincaré Prob. Statist.* **31**, 289–323.

[22] LÉONARD, C. (1995). On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. *Prob. Theory Relat. Fields* **101**, 1–44.

[23] LIPTSER, R. (1996). Large deviations for two scaled diffusions. *Prob. Theory Relat. Fields* **106**, 71–104.

[24] MARTIROSYAN, D. (2017). Large deviations for stationary measures of stochastic nonlinear wave equations with smooth white noise. *Commun. Pure Appl. Math.* **70**, 1754–1797.

[25] MCKEAN, H. P. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations* (Lecture Series in Differential Equations 7), Catholic University, Washington, DC, pp. 41–57.

[26] MEYN, S. P. *et al.* (2015). Ancillary service to the grid using intelligent deferrable loads. *IEEE Trans. Automatic Control* **60**, 2847–2862.

[27] MUFA, C. (1994). Optimal Markovian couplings and applications. *Acta Math. Sinica* **10**, 260–275.

[28] PUHALSKII, A. (2019). Large deviations of the long term distribution of a non Markov process. *Electron. Commun. Prob.* **24**, article no. 35.

[29] PUHALSKII, A. A. (2016). On large deviations of coupled diffusions with time scale separation. *Ann. Prob.* **44**, 3111–3186.

[30] PUHALSKII, A. A. (2020). *Large deviation limits of invariant measures*. Preprint. Available at https://arxiv.org/abs/2006.16456.

[31] SALINS, M., BUDHIRAJA, A. AND DUPUIS, P. (2019). Uniform large deviation principles for Banach space valued stochastic differential equations. *Trans. Amer. Math. Soc.* **372**, 8363–8421.

[32] SALINS, M. AND SPILIOPOULOS, K. (2021). Metastability and exit problems for systems of stochastic reaction–diffusion equations. *Ann. Prob.* **49**, 2317–2370.

[33] SOWERS, R. (1992). Large deviations for the invariant measure of a reaction–diffusion equation with non-Gaussian perturbations. *Prob. Theory Relat. Fields* **92**, 393–421.

[34] SOWERS, R. B. (1992). Large deviations for a reaction–diffusion equation with non-Gaussian perturbations. *Ann. Prob.* **20**, 504–537.

[35] VERETENNIKOV, A. Y. (2000). On large deviations for SDEs with small diffusion and averaging. *Stoch. Process. Appl.* **89**, 69–79.

[36] YASODHARAN, S. AND SUNDARESAN, R. (2023). Large time behaviour and the second eigenvalue problem for finite state mean-field interacting particle systems. *Adv. Appl. Prob.* **55**, 85–125.