Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 88, 1–39.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1144

Large-deviation principles of switching Markov processes via Hamilton-Jacobi equations

Mark A. Peletier* Mikola C. Schlottke†

Abstract

We prove pathwise large-deviation principles of switching Markov processes by exploiting the connection to associated Hamilton-Jacobi equations, following Jin Feng's and Thomas Kurtz's method [13]. In the limit that we consider, we show how the large-deviation problem in path-space reduces to a spectral problem of finding principal eigenvalues. The large-deviation rate functions are given in action-integral form. As an application, we demonstrate how macroscopic transport properties of stochastic models of molecular motors can be deduced from an associated principal-eigenvalue problem. The precise characterization of the macroscopic velocity in terms of principal eigenvalues confirms that breaking of detailed balance is necessary for obtaining transport. In this way, we extend and unify several existing results about molecular motors and place them in the framework of stochastic processes and large-deviation theory.

Keywords: large deviations; Hamilton-Jacobi equations; Markov processes; molecular motors; eigenvalue problems; homogenization; Feng-Kurtz method.
MSC2020 subject classifications: 60F10; 92C40; 49L25; 60J25; 35B27.
Submitted to EJP on June 5, 2021, final version accepted on May 11, 2024.

1 Introduction

In this paper we investigate large deviations for switching Markov processes that are motivated by stochastic models of molecular motors. Molecular motors are proteins that are capable of moving along filaments in a living cell. Molecular motors such as kinesin and dynein drag vesicles along while moving and thereby transport them within the cell. For more background on the phenomenon of molecular motors we refer to a number of reviews [20, 18, 35, 24, 23].

Molecular motors have a *directionality*: they typically move in one direction only. A central challenge in the study of such motors is to understand the origin of this

*Department of Mathematics and Computer Science, Eindhoven University of Technology. E-mail: m.a.peletier@tue.nl
†Department of Mathematics and Computer Science, Eindhoven University of Technology.

directionality, and characterize the speed of movement. In fact, mathematical models of molecular motors typically show no energetic benefit in moving in one direction or the other; the directionality arises from a non-trivial interplay between the microscopic features of such models and the dynamics of the motor. As a result, understanding how directionality arises as symmetry breaking in a-directional models is somewhat of a puzzle.

For certain models this puzzle has been solved, at least partially. Hastings, Kinderlehrer and Mcleod studied stationary solutions of certain Fokker-Planck equations and found sufficient conditions for the occurrence of transport [16, 15]. Vorotnikov proved sufficient conditions for transport in deterministically switching [40] and randomly switching systems [41]. Perthame, Souganidis, and Mirrahimi developed a dynamic point of view on systems of molecular motors [31, 32, 29]. In particular, Mirrahimi and Souganidis prove convergence of solutions of a Fokker-Planck equation to a ballistically travelling pulse, with a velocity that is characterized by a periodic cell problem.

In this paper we extend the results of [29] to a much broader class of systems, make explicit the connection to stochastic processes, and place the treatment squarely in the context of large-deviation theory. In this way we elaborate on the work by Perthame, Souganids and Mirrahimi, which appears to be inspired by large-deviation theory, as evidenced by the title of [31] and the use of terms such as 'Hamiltonian'.

The larger class of stochastic processes that we consider is that of *switching Markov processes in a periodic setting*. This class contains different models of molecular motors as special cases, including the *continuum ratchet* and *discrete stochastic* models (see [23] and Section 2, as well as [31, 32, 29, 16, 15]).

The first mathematical results of this paper (Theorems 4.2 and 4.3; see Figure 1 below) are large-deviation theorems for such switching Markov processes. These generalize results by Kumar and Popovic [27] by focusing on pathwise large deviations, while placing more restrictive assumptions on the microscopic dynamics. Furthermore, instead of assuming the comparison principle to be satisfied as in [27, Lemma 1], we formulate conditions that imply the comparison principle. Faggionato and Silvestri establish large-deviation principles for fully discrete, 'pseudo-one-dimensional' systems [12].

A related line of research focuses on large-deviation principles for switching diffusions in a setting where the diffusion potentials do not have small-scale oscillations. Typical results provide large-deviation rate functionals that are simple sums of small-diffusion ('Freidlin-Wentzell') and occupation ('Donsker-Varadhan') rate functionals (see e.g. [14, 17, 19, 3, 26]). The rapid-scale oscillation of the potentials in this paper creates a stronger intertwining between the diffusion and switching dynamics, and consequently the rate function is not a simple sum but an expression that fully combines the dynamics of both components.

Theorems 4.2 and 4.3 recover previous convergence results such as those of Mirrahimi and Souganidis [29, Th. 1.1-1.2]. While the methods that Mirrahimi and Souganidis apply are inspired by large-deviation theory, they do not explicitly prove large deviation principles but convergence statements on the level of Fokker-Planck equations. By proving large-deviation principles instead, we are able to make a clear distinction between the contributions that come from general large-deviation theory on the one hand, and the model-specific contributions on the other hand.

For instance, our results explain from a large-deviation point-of-view why the velocity v can be characterized by a cell problem that can be interpreted as defining a large-deviation Hamiltonian $\mathcal{H}$, through $v = \mathcal{H}'(0)$. The Hamiltonian depends on the specific model, while the relation $v = \mathcal{H}'(0)$ is independent of the microscopic details. This relation then also explains the well-known fact that detailed balance (microscopic reversibility) forces zero velocity. Indeed, we prove under general conditions (Theorem 4.8) that detailed balance leads to a symmetric Hamiltonian. By the characterization of the velocity as $v = \mathcal{H}'(0)$, this means that detailed balance has to be broken in order for transport to occur.

As another example, the numerical results of Wang, Peskin and Elston suggest that there is no transport in the limit of large reaction rates [42, Section 4.3, Figure 8(a)]. We also recover this result by proving that in this limit regime the Hamiltonian becomes symmetric (Theorem 4.9).

Overview of the paper

In Section 2, we illustrate the general results by means of a concrete example of a stochastic molecular-motor model. This provides a 'running example' with which to interpret the general results that follow. We also outline with this example the relation to the papers of Perthame, Souganidis and Mirrahimi.

In Section 3, we introduce the concepts that we work with in order to rigorously formulate our results. In Section 4 we present our main results. Figure 1 summarizes the relationships between the main theorems. Theorem 4.2 provides general conditions under which the so-called spatial component of a switching Markov process satisfies a large-deviation principle. We identify the *Hamiltonian* $\mathcal{H}(p)$, a principal eigenvalue, as the central ingredient. Under the additional assumption that $p \mapsto \mathcal{H}(p)$ is convex, Theorem 4.3 establishes an *action-integral representation*. Theorems 4.2 and 4.3 highlight the arguments that come from large-deviation theory.

We then specialize to a concrete ratchet model of molecular motors. Theorems 4.6 and 4.7 establish the large-deviation theorems for two limit regimes. While Theorem 4.6 generalizes the results in [29], Theorem 4.7 characterizes yet another limit regime. We include this result to illustrate how the general structure of proof remains unaffected by the choice of scaling. Finally, we show the symmetry of Hamiltonians under detailed balance (Theorem 4.8) and in the regime of scale separation (Theorem 4.9).

Finally, we should point out that there is a sizeable literature on *piecewise deterministic Markov processes*, which are governed by deterministic ordinary differential equations with flow fields that jump randomly between a finite set of possibilities (see e.g. [10, 11]). We expect that the results of this paper might be applicable to this class of systems, but the verification of properties such as the Comparison Principle will require different methods.

2 Example—large deviations for molecular motors

2.1 Definition of the system

In this example, we consider a two-component Markov process (X^n, I^n) with values in $\mathbb{T} \times \{1, 2\}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the one-dimensional flat torus. We fix the initial condition $(X^n(0), I^n(0)) = (x_0, i_0)$ for some $(x_0, i_0) \in \mathbb{T} \times \{1, 2\}$. Let $\psi(\cdot, 1)$ and $\psi(\cdot, 2)$ be smooth functions on the torus, and we write $\psi'(x, i)$ for the derivative of $x \mapsto \psi(x, i)$. We call these functions *potentials*. The evolution of (X^n, I^n) is characterized by the stochastic differential equation

$$\mathrm{d}X_t^n = -\psi'\left(nX_t^n, I_t^n\right)\,\mathrm{d}t + \frac{1}{\sqrt{n}}\,\mathrm{d}B_t, \tag{2.1}$$

where B_t is a standard Brownian motion. The process I^n is a continuous-time Markov chain on $\{1, 2\}$, which evolves with *jump rates* $r_{ij}(\cdot)$ such that

$$\mathbb{P}\Big(I^n(t + \Delta t) = j \,|\, I^n(t) = i, X^n(t) = x\Big) = n \cdot r_{ij}\,(nx)\,\Delta t + \mathcal{O}(\Delta t^2), \quad \text{as } \Delta t \to 0. \tag{2.2}$$

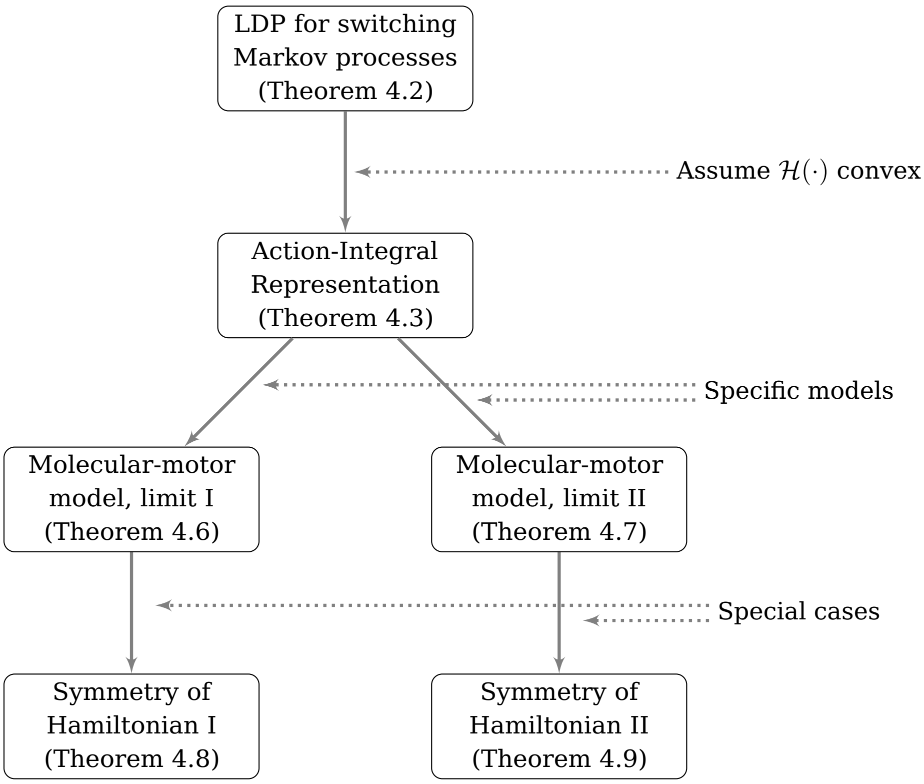


Figure 1: Overview of the results proven in this paper. From top to bottom, results become less general and more specific. Arrows indicate restrictions in passing from one context to the next.

In summary, the *spatial component* X^n is a drift-diffusion process, the *configurational component* I^n is a continuous-time Markov chain on $\{1,2\}$, and the two are coupled through their respective rates. For details about the rigorous construction of such switching drift-diffusion processes, we refer to [43, Chapter 2]. Figure 2 depicts a typical realization of (X^n, I^n), where the trajectory of the spatial component is lifted from the torus to $\mathbb{R}$.

The specific n-scaling may be motivated by starting from a process (X_t, I_t) that satisfies

$$\mathrm{d}X_t = -\psi'(X_t, I_t)\,\mathrm{d}t + \mathrm{d}B_t,$$

where the jump process I_t on $\{1,2\}$ evolves according to

$$\mathbb{P}\left(I_{t+\Delta t} = j \,|\, I_t = i, X_t = x\right) = r_{ij}(x)\Delta t + \mathcal{O}(\Delta t^2), \quad \text{as } \Delta t \to 0.$$

The *large-scale* behaviour of (X_t, I_t) is studied by considering the rescaled process (X_t^n, I_t^n) defined by $X_t^n := \frac{1}{n}X_{nt}$ and $I_t^n := I_{nt}$, and characterizing the dynamics of (X_t^n, I_t^n) for large values of n. This rescaling may be interpreted as zooming out of the x-t phase space, which is illustrated below in Figure 3. Itô calculus implies that the process (X_t^n, I_t^n) satisfies (2.1) and (2.2).

2.2 Large deviations for this example

We are interested in the behaviour of the spatial component X^n as $n \to \infty$. The behaviour of X^n for large n is shown in Figure 3. This figure suggests that X^n closely

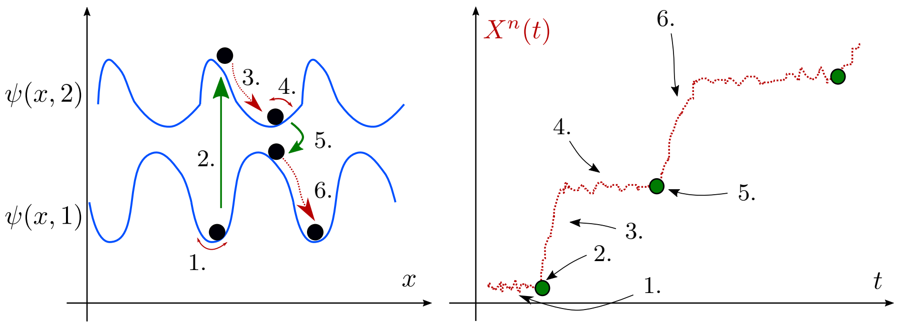


Figure 2: A typical time evolution of (X^n, I^n) satisfying (2.1) and (2.2). In the left diagram, the black bullet represents a particle that moves according to (2.1). A red arrow indicates the dynamics of the spatial component X^n. A green arrow indicates a switch of the configurational component I^n, which switches the potential in which the particle is diffusing. In the right diagram, the spatial evolution is shown in an x-t-diagram. The red dots represent the values of X^n, while a green bullet indicates a switch of the configurational component I^n. The dynamics of the particle comprises the following typical phases; 1 and 4: diffusive motion of X^n near a potential minimum; 2 and 5: configurational switch of I^n with the effect of switching to another potential; 3 and 6: flow of X^n towards a minimum of the other potential. In both diagrams, the spatial trajectory is shown lifted from the torus $\mathbb{T}$ to $\mathbb{R}$.

follows a path with a constant velocity. Indeed, when specifying the results of this paper to the example at hand—the process (X^n, I^n) defined by (2.1) and (2.2)—we find that the spatial component X^n satisfies a pathwise large-deviation principle in the limit $n \to \infty$.

To describe this fact more precisely, let $\mathcal{X} := C_{\mathbb{T}}[0,\infty)$ be the set of continuous trajectories in $\mathbb{T}$, equipped with the topology of uniform convergence on compact time intervals. The spatial component X^n is a random variable in $\mathcal{X}$, with a path distribution $\mathbb{P}(X^n \in \cdot) \in \mathcal{P}(\mathcal{X})$. We will show that there exists a *rate function* $\mathcal{I} : \mathcal{X} \to [0,\infty]$ with which $\{X^n\}_{n\in\mathbb{N}}$ satisfies a pathwise large-deviation principle in the sense of Definition 3.2 below. The gist of this statement is that for any trajectory $x \in \mathcal{X}$, we have at least intuitively

$$\mathbb{P}\left(X^n \approx x\right) \sim e^{-n\,\mathcal{I}(x)}, \quad n \to \infty. \tag{2.3}$$

The notation "$X^n \approx x$" indicates that X^n is close to x with respect to the topology on $\mathcal{X}$, and "$\sim e^{-n\,\mathcal{I}(x)}$" indicates a dominant contribution of the exponential. The rate function $\mathcal{I}$ is given by means of a *Lagrangian* $\mathcal{L} : \mathbb{R} \to [0,\infty)$ as

$$\mathcal{I}(x) = \mathcal{I}_0(x(0)) + \int_0^\infty \mathcal{L}(\partial_t x(t))\,\mathrm{d}t. \tag{2.4}$$

Here $\mathcal{I}_0 : \mathbb{T} \to [0,\infty]$ is the rate function of the initial conditions $X^n(0)$; because of the deterministic initial condition $X^n(0) = x_0$, this functional is given by $\mathcal{I}_0(x_0) = 0$ and $+\infty$ otherwise. The Lagrangian is the Legendre dual of a *Hamiltonian* $\mathcal{H} : \mathbb{R} \to \mathbb{R}$, that is $\mathcal{L}(v) = \sup_p [pv - \mathcal{H}(p)]$, and the Hamiltonian is the principal eigenvalue of an associated cell problem described in a more general context in Lemma 7.1.

Here, we focus on how this large-deviation result confirms the claim suggested by Figure 3. The rate function (2.4) has the following properties:

(i) $\mathcal{I} : \mathcal{X} \to [0,\infty]$ is nonnegative.

(ii) $\mathcal{I}(x) = 0$ if and only if $\partial_t x(t) = v$, with $v = \mathcal{H}'(0)$.

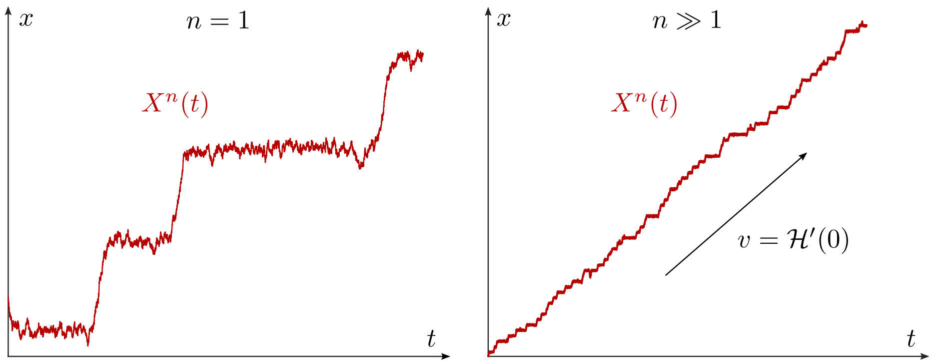


Figure 3: Two typical realizations of the spatial component X^n of the two-component process (X^n, I^n) satisfying (2.1) and (2.2). On the left, a realization is depicted for n of order one, and on the right for large n. Both graphs depict the lifted trajectory of X^n on $\mathbb{R}$. For large n, realizations of X^n closely follow a path with a constant velocity $v = \mathcal{H}'(0)$, wherein the *Hamiltonian* $\mathcal{H} = \mathcal{H}(p)$ may be derived from large-deviation theory. A more detailed illustration of the dynamics is shown in Figure 2 further above.

These two properties together characterize the unique minimizer of the rate function, and thereby in particular the typical behaviour of X^n for large n. Whenever $\mathcal{I}(x) > 0$ for a path $x \in \mathcal{X}$, then by (2.3), the probability that a realization of X^n is close to x on $\mathcal{X}$ is exponentially small in n. In fact, the large-deviation principle implies almost-sure convergence of X^n to the unique minimizer of the rate function (Theorem A.1). Uniqueness of the minimizer, item (ii), follows by strict convexity of $\mathcal{H}(p)$. For the Hamiltonian of this example, strict convexity can be proven as demonstrated in [29, Step 4 in Appendix A].

With the large-deviation principle we can investigate which sets of potentials and rates $\{\psi_1, \psi_2, r_{12}, r_{21}\}$ induce *transport*, that means a non-zero macroscopic velocity $v = \mathcal{H}'(0)$. We do not find general sufficient conditions for transport, but can draw some conclusions if the process (X^n, I^n) satisfies *detailed balance*, that is $r_{12}e^{-\psi_1} = Cr_{21}e^{-\psi_2}$ for some constant $C > 0$. Detailed balance implies that the Hamiltonian is symmetric (Theorem 4.8), and therefore $v = 0$ under detailed balance.

3 Preliminaries

In the previous section we sketched the results of this paper at the hand of an example. In this section we introduce the concepts that we use in the subsequent sections to obtain the general results of this paper in a rigorous way.

Large deviations For a Polish space E, let $\mathcal{X} := D_E[0, \infty)$ be the set of trajectories in E that are right-continuous and have left limits. We equip $\mathcal{X}$ with the Skorohod topology [9, Section 3.5]. We work with the definition of a rate function as given in [2, Chapter 1].

Definition 3.1 (Rate function)**.** *We call a map $\mathcal{I} : \mathcal{X} \to [0, \infty]$ a* rate function *if for every $C \geq 0$, the sub-level set $\{x \in \mathcal{X} : \mathcal{I}(x) \leq C\}$ is compact.* □

In particular, a rate function is lower semi-continuous. For a Borel subset $A \subseteq \mathcal{X}$, we write $\mathrm{int}(A)$ and $\mathrm{clos}(A)$ for its interior and closure.

Definition 3.2 (Large-deviation principle)**.** *For $n = 1, 2, \dots$, let P_n be a probability measure on $\mathcal{X}$, and let $\mathcal{I} : \mathcal{X} \to [0, \infty]$ be a rate function. We say that the sequence $\{\mathrm{P}_n\}_{n\in\mathbb{N}}$*

satisfies a large-deviation principle with rate function $\mathcal{I}$ *if for every Borel subset* $A \subseteq \mathcal{X}$*,*

$$-\inf_{x\in\mathrm{int}(A)} \mathcal{I}(x) \le \liminf_{n\to\infty} \frac{1}{n}\log \mathrm{P}_n(A) \le \limsup_{n\to\infty} \frac{1}{n}\log \mathrm{P}_n(A) \le -\inf_{x\in\mathrm{clos}(A)} \mathcal{I}(x) \qquad \square$$

A large-deviation principle provides an estimate of the probabilities $\mathrm{P}_n(A)$ on the logarithmic scale. At least intuitively,

$$\mathrm{P}_n(A) \approx e^{-n\,\inf_{x\in A}\mathcal{I}(x)}, \qquad n\to\infty.$$

Illustrating examples of a large-deviation principle can be found for instance in Ellis' note on Boltzmann's discoveries [8]. General introductions to the topic are also provided in [2, Chapter 1] and [13, Chapter 3].

Identifying tractable formulas for a rate function is crucial for drawing conclusions from a large-deviation principle. In this paper, we shall aim for finding action-integral representations of rate functions. Let $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ be the flat d-dimensional torus, and let $\mathcal{AC}([0,\infty);\mathbb{T}^d)$ be the set of absolutely continuous trajectories in $\mathbb{T}^d$.

Definition 3.3 (Action-integral form of rate function)**.** *We say that a rate function* $\mathcal{I} : D_{\mathbb{T}^d}[0,\infty) \to [0,\infty]$ *is of action-integral form if there is a non-trivial convex map* $\mathcal{L} : \mathbb{R}^d \to [0,\infty]$ *with which*

$$\mathcal{I}(x) = \begin{cases} \mathcal{I}_0(x(0)) + \int_0^\infty \mathcal{L}\left(\partial_t x(t)\right)\,\mathrm{d}t & \text{if } x\in\mathcal{AC}([0,\infty);\mathbb{T}^d), \\ +\infty & \text{otherwise}, \end{cases}$$

where $\mathcal{I}_0 : \mathbb{T}^d \to [0,\infty]$ *is a rate function. We refer to the map* $\mathcal{L}$ *as the* Lagrangian. $\square$

Switching Markov processes in a periodic setting We shall consider Markov processes defined by two-component stochastic processes (X^n, I^n) taking values in state spaces E_n that satisfy the following condition.

Condition 3.4 (Setting)**.** *Fix* $J \in \mathbb{N}$*. For* $n \in \mathbb{N}$*, the state space* E_n *is a product space* $E_n := E_n^X \times \{1,\dots,J\}$*, where* E_n^X *is a compact Polish space satisfying the following: there are continuous injective maps* $\iota_n : E_n^X \to \mathbb{T}^d$ *such that for all* $x\in\mathbb{T}^d$ *there exists* $x_n \in E_n^X$ *with which* $\iota_n(x_n)\to x$ *as* $n\to\infty$*.* $\square$

This condition means that the E_n^X are asymptotically dense in the torus $\mathbb{T}^d$. The typical example is the periodic lattice $(n^{-1}\mathbb{Z})^d/\mathbb{Z}^d$, where the torus is recovered in the limit of n to infinity. Another example is simply $E_n^X \equiv \mathbb{T}^d$. When it is clear from the context, we omit ι_n in the notation. In general, for a function $f = f(x,i)$ of a continuous variable x and a discrete variable i, we shall sometimes write $f(x,i) = f_i(x)$ to shorten the notation.

Let $\mathcal{X}_n := D_{E_n}[0,\infty)$. For a distribution $\mu\in\mathcal{P}(E_n)$, we identify an E_n-valued two-component process (X^n, I^n) having initial condition μ with its path distribution $\mathbb{P}_\mu^n \in \mathcal{P}(\mathcal{X}_n)$. In order to define a path distribution, we shall specify a linear map $L_n : \mathcal{D}(L_n) \subseteq C(E_n) \to C(E_n)$ on a domain $\mathcal{D}(L_n)$ and assume well-posedness of the martingale problem of the pair (L_n,μ).

We refer to [9, Chapter 4, Section 3] for a precise treatment of the martingale problem, but we briefly recall the martingale problem here. Consider a complete separable metric space Y, a linear operator $A \subset B(Y)\times B(Y)$ with domain $\mathcal{D}(A)$ in the space of bounded, Borel-measurable functions $B(Y)$, and a distribution $\mu\in\mathcal{P}(Y)$. A process Z is called a solution to the martingale problem for (A,μ) if for all $g\in\mathcal{D}(A)$, the process

$$g\left(Z(t)\right) - g\left(Z(0)\right) - \int_0^t Ag\left(Z(s)\right)ds$$

is a martingale and if $Z(0) \sim \mu$. Uniqueness holds for the martingale problem for (A, μ) if all finite-dimensional distributions of a solution are unique. A martingale problem for A is well-posed if existence and uniqueness hold for (A, μ) for any initial distribution $\mu \in \mathcal{P}(Y)$.

We shall call a linear map L_n as above a *generator* if it gives rise to a well-posed martingale problem. We specify the generators L_n of (X^n, I^n) from the following ingredients:

(1) For $i \in \{1, \dots, J\}$, we have a map $L_n^i : \mathcal{D}(L_n^i) \subseteq C(E_n^X) \to C(E_n^X)$ that is the generator of an E_n^X-valued Markov process.

(2) For $i, j \in \{1, \dots, J\}$, we have a continuous map $r_{ij}^n : E_n^X \to [0, \infty)$.

With that, define the map $L_n : \mathcal{D}(L_n) \subseteq C(E_n) \to C(E_n)$ by

$$L_n f(x, i) := L_n^i f(\cdot, i)(x) + \sum_{j=1}^{J} r_{ij}^n(x) \left[f(x, j) - f(x, i) \right], \tag{3.1}$$

where the domain is $\mathcal{D}(L_n) = \{f \in C(E_n) \,:\, f(\cdot, i) \in \mathcal{D}(L_n^i), i = 1, \dots, J\}$. For any $\mu \in \mathcal{P}(E_n)$, we denote by $\mathbb{P}_\mu^n$ the path-distribution of the solution to the martingale problem for (L_n, μ).

Condition 3.5 (Well-posedness)**.** *The martingale problem for L_n is well-posed and the map $E_n \ni z \mapsto \mathbb{P}_{\delta_z}^n \in \mathcal{P}(\mathcal{X}_n)$ is Borel measurable with respect to the weak topology on $\mathcal{P}(\mathcal{X}_n)$.* □

Definition 3.6 (Switching Markov processes in a periodic setting)**.** *We call a two-component Markov process (X^n, I^n) a* switching Markov process *if it takes values in $E_n = E_n^X \times \{1, \dots, J\}$ satisfying Condition 3.4 and if it has a generator L_n that is given by* (3.1) *and satisfies Condition 3.5.* □

Condition 3.5 is the basic assumption on the processes in [13]. We expect the martingale problem to be well-posed provided that the continuous maps r_{ij}^n are sufficiently regular. However, we do not investigate conditions on a map L_n that imply well-posedness, but assume it instead. A sufficient condition for the measurability in Condition 3.5 is given in [9, Theorem 4.4.6]. The book by Yin and Zhu, in particular Section 2.2, lists a number of references for such existence and regularity properties for switching hybrid diffusions [43].

4 Main results

In the previous section we introduced the notion of a large-deviation principle and defined switching Markov processes in a periodic setting. In this section we present our main results as depicted in the flow-diagram Figure 1 above. First, we formulate general conditions for a large-deviation principle of switching Markov processes (Theorem 4.2). Then we find an action-integral representation of the rate function under an additional convexity assumption (Theorem 4.3). The remaining theorems arise from specifications of the general setting to specific models. We prove large-deviation principles for molecular-motor models in two limit regimes (Theorems 4.6 and 4.7), and derive the fact that detailed balance and separation of scales imply symmetry of Hamiltonians (Theorems 4.8 and 4.9).

4.1 Large-deviation principle for switching Markov processes

We consider switching Markov processes (X^n, I^n) in a periodic setting in the sense of Definition 3.6, with generators of the form (3.1). The essence of this section is

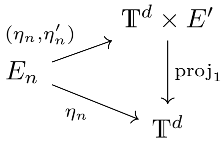
Theorem 4.2, which provides general conditions that imply a pathwise large-deviation principle of the spatial component X^n. We state the conditions in terms of *nonlinear generators* defined as follows.

Definition 4.1 (Nonlinear generators)**.** *Let* L_n *be the map defined by* (3.1)*. The* nonlinear generator *is the map* $H_n : \mathcal{D}(H_n) \subseteq C(E_n) \to C(E_n)$ *defined by*

$$H_n f(x) := \frac{1}{n} e^{-nf(x)} L_n(e^{nf(\cdot)})(x), \tag{4.1}$$

on the domain $\mathcal{D}(H_n) := \{f \in C(E_n) \,:\, e^{nf(\cdot)} \in \mathcal{D}(L_n)\}$*.* □

We shall work under the assumption that the nonlinear generators H_n converge in the limit $n \to \infty$. To formulate this convergence assumption, we need to introduce an additional state space E' for collecting up-scaled variables. The following diagram depicts the relation between the state spaces:



In the diagram, $\eta_n : E_n \to \mathbb{T}^d$ is the projection defined by $\eta_n(x,i) := \iota_n(x)$, where $\iota_n : E_n^X \to \mathbb{T}^d$ is the embedding of Condition 3.4. The map $\eta_n' : E_n \to E'$ is continuous and injective for every n. We shall assume that the E_n are asymptotically dense:

(C1) For $(x, z') \in \mathbb{T}^d \times E'$ there exists $y_n \in E_n$ such that $\eta_n(y_n) \to x$ and $\eta_n'(y_n) \to z'$ as $n \to \infty$.

A limit operator of H_n is defined by a graph $H \subseteq C(\mathbb{T}^d) \times C(\mathbb{T}^d \times E')$, a multi-valued operator. We shall assume the following convergence condition:

(C2) The domain $\mathcal{D}(H)$ satisfies $C^\infty(\mathbb{T}^d) \subseteq \mathcal{D}(H) \subseteq C^1(\mathbb{T}^d)$. For $(f, g) \in H$, there exist functions $f_n \in \mathcal{D}(H_n)$, $n \in \mathbb{N}$, such that as $n \to \infty$,

$$\|f \circ \eta_n - f_n\|_{L^\infty(E_n)} \to 0 \quad \text{and} \quad \|g \circ (\eta_n, \eta_n') - H_n f_n\|_{L^\infty(E_n)} \to 0.$$

Frequently, for any f in the domain of H, the corresponding image functions g are naturally parametrized by a set of functions on E':

(C3) There are a set $\mathcal{C} \subseteq C(E'; \mathbb{R}^k)$ and functions $H_{f,\varphi} \in C(\mathbb{T}^d \times E')$ with which

$$H = \{(f, H_{f,\varphi}) \,:\, f \in \mathcal{D}(H), \varphi \in \mathcal{C}\}.$$

Specific examples of models satisfying the above conditions will be discussed in Sections 4.3 and 7.

Theorem 4.2 (Large deviation principle for switching processes)**.** *Let* (X^n, I^n) *be a switching Markov process in the sense of Definition 3.6, with nonlinear generators* H_n *of Definition 4.1. Let* E' *be a compact metric space satisfying (C1), and let* $H \subseteq C(\mathbb{T}^d) \times C(\mathbb{T}^d \times E')$ *be a multi-valued operator satisfying (C2) and (C3) from above. Suppose the following:*

(T1) For every $\varphi \in \mathcal{C}$ *there is a map* $H_\varphi : \mathbb{R}^d \times E' \to \mathbb{R}$ *such that for all* $f \in \mathcal{D}(H)$*,*

$$H_{f,\varphi}(x, z') = H_\varphi(\nabla f(x), z'), \qquad (x, z') \in \mathbb{T}^d \times E'.$$

(T2) For every $p \in \mathbb{R}^d$, there exists a function $\varphi_p \in \mathcal{C}$ and a constant $\mathcal{H}(p) \in \mathbb{R}$ such that $H_{\varphi_p}(p, z') = \mathcal{H}(p)$ for all $z' \in E'$.

Suppose furthermore that $\{X^n(0)\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $\mathbb{T}^d$ with rate function $\mathcal{I}_0 : \mathbb{T}^d \to [0, \infty]$. Then the family of processes $\{X^n\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $D_{\mathbb{T}^d}[0, \infty)$ with a rate function $\mathcal{I} : D_{\mathbb{T}^d}[0, \infty) \to [0, \infty]$, and there exists a semigroup $V(t)$ with which the rate function is given by (5.1)*.*

We give the proof in Section 5. The formula for the rate function $\mathcal{I}$ is not important here, which is why we report it only below in (5.1) in the proof section. Condition (T1) means that the images depend on the variable $x \in \mathbb{T}^d$ only via the gradients $\nabla f(x)$. In the molecular-motor models, Condition (T2) is verified by solving a principal-eigenvalue problem, in which the constant $\mathcal{H}(p)$ is the unique principal eigenvalue of a certain cell problem.

4.2 Action-integral representation of the rate function

In the previous section, we formulated general conditions that imply a pathwise large-deviation principle. The rate function of Theorem 4.2 however is still generic (equation (5.1) below). The following theorem shows that under an additional convexity assumption, the rate function is of action-integral form in the sense of Definition 3.3 above.

Theorem 4.3. *Consider the setting of Theorem 4.2. For $p \in \mathbb{R}^d$, let $\mathcal{H}(p)$ be the constant in (T2) of Theorem 4.2. Suppose further the following:*

(T3) The map $p \mapsto \mathcal{H}(p)$ is convex and $\mathcal{H}(0) = 0$.

Then the rate function of Theorem 4.2 is of action-integral form with the Lagrangian defined by $\mathcal{L}(v) = \sup_{p\in\mathbb{R}^d} [p \cdot v - \mathcal{H}(p)]$.

Theorem 4.3 is proven in Section 6.

4.3 Large deviations for models of molecular motors

In the previous two sections we considered general switching Markov processes in a periodic setting. In this section we further specify to a class of stochastic processes motivated by molecular motors.

Definition 4.4 (A process modeling molecular motors)**.** *The pair (X^n, I^n) is a Markov process with values in $E_n = \mathbb{T}^d \times \{1, \dots, J\}$ with generator L_n acting on functions $f = f(x, i)$ as*

$$L_n f(x, i) := b_i(nx) \cdot \nabla_x f(\cdot, i)(x) + \frac{1}{n}\frac{1}{2}\Delta_x f(\cdot, i)(x) + \sum_{j\neq i} \gamma(n) r_{ij}(nx) \left[f(x, j) - f(x, i)\right], \quad (4.2)$$

where $\gamma(n) > 0$, $r_{ij}(\cdot) \in C^\infty(\mathbb{T}^d; [0, \infty))$, and $b_i(\cdot) \in C^\infty(\mathbb{T}^d)$, for $i, j = 1, \dots, J$. □

The process of Definition 4.4 is an example of a switching Markov process with generators L_n^i defined on the core $C^2(\mathbb{T}^d)$ by

$$L_n^i g(x) := b_i(nx) \cdot \nabla g(x) + \frac{1}{n}\frac{1}{2}\Delta g(x),$$

and rates $r_{ij}^n(x) = \gamma(n) r_{ij}(nx)$. The domain $\mathcal{D}(L_n^i)$ of the generators L_n^i contains the core, but is larger than $C^2(\mathbb{T}^d)$. The domain of L_n is the set given by $\mathcal{D}(L_n) = \{f(x, i) : f(\cdot, i) \in \mathcal{D}(L_n^i)\}$, and for functions f such that $f(\cdot, i) \in C^2(\mathbb{T}^d)$, the generator acts as defined in (4.2). The example of Section 2, a stochastic model of molecular motors, corresponds to the choices $d = 1$, $b_i = -\psi'(\cdot, i)$, $J = 2$ and $\gamma(n) = n$.

The fact that such a process satisfies Condition 3.5 follows from classical theory of smooth linear parabolic systems of equations, as in e.g. [28, Sec. VII.8-10].

Definition 4.5. *Let $J \in \mathbb{N}$. We call a matrix $A \in \mathbb{R}^{J\times J}$* irreducible *if there is no decomposition of $\{1, \dots, J\}$ into two disjoint sets $\mathcal{J}_1$ and $\mathcal{J}_2$ such that $A_{ij} = 0$ whenever $i \in \mathcal{J}_1$ and $j \in \mathcal{J}_2$.* □

Theorem 4.6 (Limit I)**.** *Let (X^n_t, I^n_t) be the Markov process of Definition 4.4 with parameter $\gamma(n) = n$. Assume that the matrix R with entries $R_{ij} = \sup_{y\in\mathbb{T}^d} r_{ij}(y)$ is irreducible. Suppose furthermore that the family of initial conditions $\{X^n(0)\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $\mathbb{T}^d$ with rate function $\mathcal{I}_0 : \mathbb{T}^d \to [0,\infty]$.*

Then the family of stochastic processes $\{X^n\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $C_{\mathbb{T}^d}[0,\infty)$ with rate function of action-integral form. The Hamiltonian $\mathcal{H}(p)$ is the principal eigenvalue of an associated cell problem described in Lemma 7.1.

The irreducibility condition is imposed to solve the principal-eigenvalue problem that we obtain, and is inspired by sufficient conditions for solvability of a coupled system of elliptic PDEs [38].

The parameter $\gamma(n)$ allows to model a time-scale separation of the components. The following theorem shows that if $\gamma(n)$ scales super-linearly, then the spatial component is effectively driven by potentials averaged over the stationary measure of the fast configurational component, and the large-deviation principle is governed by an averaged Hamiltonian.

Theorem 4.7 (Limit II)**.** *Let (X^n_t, I^n_t) be the Markov process of Definition 4.4, with parameter $\gamma(n)$ such that $n^{-1}\gamma(n) \to \infty$ as $n \to \infty$. Assume that for every $y \in \mathbb{T}^d$, the matrix $R(y)$ with entries $R(y)_{ij} = r_{ij}(y)$ is irreducible. Suppose furthermore that the family of random variables $\{X^n(0)\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $\mathbb{T}^d$ with rate function $\mathcal{I}_0 : \mathbb{T}^d \to [0,\infty]$.*

Then $\{X^n\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle in $C_{\mathbb{T}^d}[0,\infty)$ with rate function of action-integral form. The Hamiltonian $\overline{\mathcal{H}}(p)$ is the principal eigenvalue of an associated averaged cell problem described in Lemma 7.2.

4.4 Detailed balance implies symmetric Hamiltonians

The large-deviation principles established by Theorems 4.6 and 4.7 can be used to analyse which sets of potentials and rates induce transport on macroscopic scales. To that end, we specify to $b_i(y) = -\nabla_y\psi_i(y)$ and $\gamma(n) = n$ in the generators defined in (4.2).

We say that the set of potentials and rates $\{r_{ij}, \psi_i\}$ satisfies *detailed balance* if for all $i, j \in \{1, \dots, J\}$ and $y \in \mathbb{T}^d$, we have

$$r_{ij}(y)e^{-2\psi_i(y)} = r_{ji}(y)e^{-2\psi_j(y)}. \tag{4.3}$$

Theorem 4.8 (Detailed balance implies a symmetric Hamiltonian)**.** *Consider the same setting and assumptions of Theorem 4.6. Suppose that the detailed-balance condition* (4.3) *is satisfied. Then the Hamiltonian $\mathcal{H}(p)$ of Theorem 4.6 satisfies $\mathcal{H}(p) = \mathcal{H}(-p)$ for all $p \in \mathbb{R}^d$.*

We give the proof of Theorem 4.8 here, since it is solely based on a suitable formula for $\mathcal{H}(p)$.

Proof of Theorem 4.8. We prove in Proposition 8.1 that under the detailed-balance condition, the principal eigenvalue $\mathcal{H}(p)$ is given by

$$\mathcal{H}(p) = \sup_{\mu\in\mathbf{P}} \left[K_p(\mu) - \mathcal{R}(\mu)\right],$$

where $\mathbf{P} \subset \mathcal{P}(E')$ is a subset of probability measures on $E' = \mathbb{T}^d \times \{1, \ldots, J\}$ specified in Proposition 8.1, $\mathcal{R}(\mu)$ is the relative Fisher information specified in (8.7), and $K_p(\mu)$ is given by

$$K_p(\mu) = \inf_{\phi} \Bigg\{ \sum_{i=1}^{J} \int_{\mathbb{T}^d} \left(\frac{1}{2} |\nabla \phi_i(x) + p|^2 - \sum_{j=1}^{J} r_{ij}(x) \right) \mathrm{d}\mu_i(x) + \sum_{i,j=1}^{J} \int_{\mathbb{T}^d} \pi_{ij}(x) \sqrt{\overline{\mu}_i(x) \overline{\mu}_j(x)} e^{\psi_j(x) + \psi_i(x)} \cosh\left(\phi(x,j) - \phi(x,i)\right) \mathrm{d}x \Bigg\},$$

where $\pi_{ij}(x) = r_{ij}(x) e^{-2\psi_i(x)}$, the infimum is taken over vectors of functions $\phi_i = \phi(\cdot, i) \in C^2(\mathbb{T}^d)$, and $\mathrm{d}\mu_i(x) = \overline{\mu}_i(x) \mathrm{d}x$.

Let $\mu \in \mathbf{P}$. We show that $K_p(\mu) = K_{-p}(\mu)$, which implies $\mathcal{H}(p) = \mathcal{H}(-p)$. The sum in which the $\cosh(\cdot)$ terms appear is symmetric in the sense that

$$C(\phi) := \sum_{i,j=1}^{J} \int_{\mathbb{T}^d} \pi_{ij}(x) \sqrt{\overline{\mu}_i(x) \overline{\mu}_j(x)} e^{\psi_j(x) + \psi_i(x)} \cosh\left(\phi(x,j) - \phi(x,i)\right) \mathrm{d}x$$

satisfies $C(\phi) = C(-\phi)$. The bijective transformation $\phi \to (-\phi)$ together with the sign change $p \to (-p)$ leaves the infimum in $K_p(\mu)$ invariant, and hence symmetry of $C(\phi)$ implies the claimed symmetry $K_p(\mu) = K_{-p}(\mu)$. □

With a similar analysis, we can study the behaviour of molecular motors under external forces. Let (X^n, I^n) be the stochastic process of Theorem 4.6 in dimension $d = 1$ with drift $b_i(y) = F - \psi'(y, i)$, where F is a constant (modeling an external force) and $\psi(\cdot, i) \in C^\infty(\mathbb{T})$ are smooth periodic potentials, $i = 1, \ldots, J$. The process (X^n, I^n) is $\mathbb{T} \times \{1, \ldots, J\}$-valued and satisfies

$$\mathrm{d}X_t^n = (F - \psi'(nX_t^n, I_t^n))\,\mathrm{d}t + \frac{1}{\sqrt{n}}\,\mathrm{d}B_t,$$

where I_t^n a jump process on $\{1, \ldots, J\}$ with jump rates $nr_{ij}(nx)$.

For example, this model predicts a *positive force-velocity feedback* under detailed balance: $F > 0$ implies $\partial_p \mathcal{H}(0) > 0$, and $F < 0$ implies $\partial_p \mathcal{H}(0) < 0$. The positive force-velocity feedback may be derived from the following properties:

(a) $\mathcal{H}(-F - p) = \mathcal{H}(-F + p)$ for all p,

(b) $\mathcal{H}(0) = 0$, and

(c) $\mathcal{H}(\cdot)$ is strictly convex.

Indeed, specializing (a) to $p = F$, we find $\mathcal{H}(-2F) = \mathcal{H}(-F + F) = \mathcal{H}(0) = 0$. Since the Hamiltonian is also strictly convex (c), the Hamiltonian must have a positive slope at $p = 0$ if $F > 0$ and a negative slope at $p = 0$ if $F < 0$.

The symmetry property (a) is a consequence of detailed balance and can be shown similarly as Theorem 4.8; the only difference is that in the formula for $K_p(\mu)$, the term $|\nabla \phi_i(x) + p|^2$ gets replaced by the term $|\nabla \phi_i(x) + F + p|^2$, with which we find

$$\mathcal{H}(-F - p) = \sup_{\mu \in \mathbf{P}} \left[K_{-F-p}(\mu) - \mathcal{R}(\mu) \right] = \sup_{\mu \in \mathbf{P}} \left[K_{-F+p}(\mu) - \mathcal{R}(\mu) \right] = \mathcal{H}(-F + p).$$

Property (b) is proven further below in Section 7.1 under *Verification of (T3) of Theorem 4.3.* Property (c) is proven in [29, Lemma 2.1] in a different but similar context. We do not include a proof of (c) here.

Theorem 4.9 (Separation of time scales implies a symmetric Hamiltonian)**.** *Let the stochastic process (X_t^n, I_t^n) of Definition 4.4, with $b_i = -\nabla\psi_i$, satisfy the assumptions of Theorem 4.7. Suppose in addition that the rates $r_{ij}(\cdot)$ are constant on $\mathbb{T}^d$. Then $\overline{\mathcal{H}}(p) = \overline{\mathcal{H}}(-p)$, where $\overline{\mathcal{H}}(p)$ is the Hamiltonian in Theorem 4.7.*

Since the derivation of the required formula for $\overline{\mathcal{H}}(p)$ is similar to the derivation of $\mathcal{H}(p)$, we omit the details and only give a sketch of the argument here.

Sketch of proof of Theorem 4.9. The principal eigenvalue $\overline{\mathcal{H}}(p)$ is given by

$$\overline{\mathcal{H}}(p) = \sup_{\mu\in\mathbf{P}} \left[K_p(\mu) - \mathcal{R}(\mu)\right], \quad K_p(\mu) = \inf_{\varphi\in C^\infty(\mathbb{T}^d)} \frac{1}{2}\int_{\mathbb{T}^d} |\nabla\varphi + p|^2 \,\mathrm{d}\mu,$$

with $\mathbf{P}$ and $\mathcal{R}$ specified below. The bijective transformation $\varphi \to (-\varphi)$ leaves the infimum in $K_p(\mu)$ invariant, and therefore we have $K_p(\mu) = K_{-p}(\mu)$ for all $\mu \in \mathbf{P}$. This implies $\overline{\mathcal{H}}(p) = \overline{\mathcal{H}}(-p)$.

In the formula for $\overline{\mathcal{H}}(p)$, the set of probability measures $\mathbf{P} \subset \mathcal{P}(\mathbb{T}^d)$ is

$$\mathbf{P} = \left\{\mu \in \mathcal{P}(\mathbb{T}^d) \,:\, \mu \ll \mathrm{d}x \text{ and } \mathrm{d}\mu = \overline{\mu}\mathrm{d}x \text{ with } \nabla\left(\log\overline{\mu}\right) \in L^2_\mu(\mathbb{T}^d)\right\}.$$

The map $\mathcal{R}$ is the relative Fisher information; with the stationary measure ν of the jump process on $\{1,\dots,J\}$ with rates r_{ij},

$$\mathcal{R}(\mu) = \frac{1}{8}\int_{\mathbb{T}^d} \left|\nabla\log\left(\frac{\overline{\mu}}{e^{-2\overline{\psi}}}\right)\right|^2 \mathrm{d}\mu, \quad \overline{\psi}(x) = \sum_i \nu_i\,\psi_i(x).$$

□

5 Proof of large-deviation principle for switching Markov processes

The main point of this section is to prove Theorem 4.2, the large-deviation principle for switching Markov processes in a periodic setting. The proof is based on a connection between large deviations and Hamilton-Jacobi equations that we first make explicit in Section 5.1 by adapting theorems of [13] to our setting.

5.1 Strategy of proof

Viscosity solutions and the comparison principle We adapt [13, Definitions 6.1 and 7.1] to the compact setting. For a Banach space B, we identify operators with graphs $H \subseteq B \times B$, with domain $\mathcal{D}(H) := \{f \,:\, \exists(f,g) \in H\}$ and range $\mathcal{R}(H) := \{g \,:\, \exists(f,g) \in H\}$, and refer to them as *multi-valued operators*. For the following definition, E and E' are compact Polish spaces with metrics d_E and $d_{E'}$, $B(E\times E')$ is the set of measurable and bounded functions on $E\times E'$, equipped with the uniform norm, and $M(E\times E')$ is the set of measurable functions.

Definition 5.1 (Viscosity solutions)**.** *Let $H \subseteq C(E)\times M(E\times E')$ be a multi-valued operator with domain $\mathcal{D}(H) \subseteq C(E)$. Let $h \in C(E)$ and $\tau > 0$.*

i) A function $u_1 : E \to \mathbb{R}$ is a viscosity subsolution of $(1-\tau H)u = h$ if it is bounded and upper semicontinuous, and if for all $(f,g) \in H$ there exists a point $(x,z') \in E\times E'$ such that

$$(u_1 - f)(x) = \sup(u_1 - f) \quad \textit{and} \quad u_1(x) - \tau g(x,z') - h(x) \le 0.$$

ii) A function $u_2 : E \to \mathbb{R}$ is a viscosity supersolution of $(1 - \tau H)u = h$ if it is bounded and lower semicontinuous, and if for all $(f, g) \in H$ there exists a point $(x, z') \in E \times E'$ such that

$$(f - u_2)(x) = \sup(f - u_2) \quad \textit{and} \quad u_2(x) - \tau g(x, z') - h(x) \geq 0.$$

iii) A function $u_1 : E \to \mathbb{R}$ is a strong viscosity subsolution of $(1 - \tau H)u = h$ if it is bounded and upper semicontinuous, and if for all $(f, g) \in H$ and $x \in E$, whenever

$$(u_1 - f)(x) = \sup(u_1 - f),$$

then there exists a $z' \in E'$ such that

$$u_1(x) - \tau g(x, z') - h(x) \leq 0.$$

Similarly for strong viscosity supersolutions.

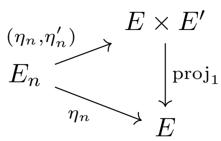
A function $u \in C(E)$ is called a viscosity solution of $(1 - \tau H)u = h$ if it is both a viscosity sub- and supersolution. □

Let us briefly highlight the adaptations we made with respect to [13]. First, formulating viscosity solutions via sequences as in [13, Definition 7.1] is only required when working with non-compact spaces, while in the context of this paper we only work in compact spaces. Second, the product space $E \times E'$ in this paper corresponds to the set E' in [13].

Definition 5.2 (Comparison Principle)**.** *The* comparison principle *holds for viscosity sub- and supersolutions of $(1 - \tau H)u = h$ if for any viscosity subsolution u_1 and viscosity supersolution u_2, we have $u_1 \leq u_2$ on E. For two operators $H_\dagger, H_\ddagger \subseteq C(E) \times C(E \times E')$, we say that the comparision principle holds if for any viscosity subsolution u_1 of $(1 - \tau H_\dagger)u = h$ and viscosity supersolution u_2 of $(1 - \tau H_\ddagger)u = h$, we have $u_1 \leq u_2$ on E.* □

If the comparison principle holds, then viscosity solutions are unique, since two viscosity solutions u, v satisfy $u \leq v$ and $v \leq u$.

A general large-deviation theorem Just as in Theorem 4.2, we work with compact Polish spaces E_n, E and E' that are related via continuous embeddings η_n and η'_n by



such that for any $x \in E$, there exist $x_n \in E_n$ such that $\eta_n(x_n) \to x$ as $n \to \infty$. The following theorem is an adaptation of [13, Theorem 7.18] to our setting. This adaptation is obtained by collecting in one place assumptions that are mentioned in several places in [13], and specializing them to the compact setting.

Theorem 5.3. *Let L_n be the generator of an E_n-valued process Y^n such that Condition 3.5 is satisfied, and let H_n be the nonlinear generators defined by $H_n f = \frac{1}{n} e^{-nf} L_n e^{nf}$, for $n \in \mathbb{N}$. Let the compact Polish spaces E_n, E and E' be related as in the above diagram. In addition, suppose:*

(i) (Condition 7.9 of [13] on the state spaces) There exists an index set Q and approximating state spaces $A_n^q \subseteq E_n$, $q \in Q$, such that the following holds:

(a) For $q_1, q_2 \in Q$, there exists $q_3 \in Q$ such that $A_n^{q_1} \cup A_n^{q_2} \subseteq A_n^{q_3}$.

(b) For each $x \in E$, there exists $q \in Q$ and $y_n \in A_n^q$ such that $\eta_n(y_n) \to x$ as $n \to \infty$.

(c) For each $q \in Q$, there exist compact sets $K_1^q \subseteq E$ and $K_2^q \subseteq E \times E'$ such that

$$\sup_{y \in A_n^q} \inf_{x \in K_1^q} d_E(\eta_n(y), x) \xrightarrow{n\to\infty} 0,$$

and

$$\sup_{y \in A_n^q} \inf_{(x,z) \in K_2^q} \left[d_E(\eta_n(y), x)) + d_{E'}(\eta_n'(y), z) \right] \xrightarrow{n\to\infty} 0.$$

(d) For each compact $K \subseteq E$, there exists $q \in Q$ such that $K \subseteq \liminf \eta_n(A_n^q)$.

(ii) (Convergence Condition 7.11 of [13]) There exist multi-valued operators $H_\dagger, H_\ddagger \subseteq C(E) \times C(E \times E')$ which are the limit of the H_n's in the following sense:

(a) For each $(f,g) \in H_\dagger$, there exist $f_n \in \mathcal{D}(H_n)$ such that

$$\sup_n \left(\sup_{x \in E_n} |f_n(x)| + \sup_{x \in E_n} |H_n f_n(x)| \right) < \infty,$$

and for each $q \in Q$, $\lim_{n\to\infty} \sup_{y \in A_n^q} |f_n(y) - f(\eta_n(y))| = 0$. Furthermore, for each $q \in Q$ and every sequence $y_n \in A_n^q$ such that $\eta_n(y_n) \to x \in E$ and $\eta_n'(y_n) \to z' \in E'$, we have $\limsup_{n\to\infty} H_n f_n(y_n) \le g(x, z')$.

(b) For each $(f,g) \in H_\ddagger$, there exist $f_n \in \mathcal{D}(H_n)$ (not necessarily the same as above in (a)) such that

$$\sup_n \left(\sup_{x \in E_n} |f_n(x)| + \sup_{x \in E_n} |H_n f_n(x)| \right) < \infty,$$

and for each $q \in Q$, $\lim_{n\to\infty} \sup_{y \in A_n^q} |f_n(y) - f(\eta_n(y))| = 0$. Furthermore, for each $q \in Q$ and every sequence $y_n \in E_n$ such that $\eta_n(y_n) \to x \in E$ and $\eta_n'(y_n) \to z' \in E'$, we have $\liminf_{n\to\infty} H_n f_n(y_n) \ge g(x, z')$.

(iii) (Comparison principle) For each $h \in C(E)$ and $\tau > 0$, the comparison principle holds for viscosity subsolutions of $(1 - \tau H_\dagger)u = h$ and viscosity supersolutions of $(1 - \tau H_\ddagger)u = h$.

Let $X_t^n := \eta_n(Y_t^n)$ be the corresponding E-valued process. Suppose that $\{X^n(0)\}_{n \in \mathbb{N}}$ satisfies a large-deviation principle in E with rate function $\mathcal{I}_0 : E \to [0,\infty]$.

Then $\{X^n\}_{n\in\mathbb{N}}$ satisfies the large-deviation principle with a rate function $\mathcal{I} : C_E[0,\infty) \to [0,\infty]$. Furthermore, there exists a semigroup $V(t) : C(E) \to C(E)$ with which the rate function is given by

$$\mathcal{I}(x) = \mathcal{I}_0(x(0)) + \sup_{k \in \mathbb{N}} \sup_{(t_1,\dots,t_k)} \sum_{i=1}^k \mathcal{I}_{t_i - t_{i-1}}(x(t_i)|x(t_{i-1})), \tag{5.1}$$

where for $z, y \in E$,

$$\mathcal{I}_t(z|y) = \sup_{f \in C(E)} \left[f(z) - V(t)f(y) \right]. \tag{5.2}$$

The semigroup $V(t)$ is defined via the Crandall-Liggett Theorem—for details we refer to [13, Chapter 5].

5.2 Proof of Theorem 4.2

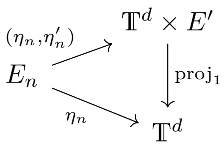
We prove Theorem 4.2 by verifying the conditions of Theorem 5.3, which are convergence of nonlinear generators (Proposition 5.4) and the comparison principle (Proposition 5.5). The rest of this section below the proof of Theorem 4.2 is devoted to proving the propositions. We point out that the main challenge is to prove the comparison principle using only (T1) and (T2) of Theorem 4.2. In the following propositions and lemmas, we assume that conditions (C1), (C2), and (C3) are satisfied, as in the setting of Theorem 4.2.

Proposition 5.4. *In the setting of Theorem 4.2, condition (i) of Theorem 5.3 is satisfied. Let $H \subseteq C^1(\mathbb{T}^d) \times C(\mathbb{T}^d \times E')$ be a multi-valued operator satisfying (T1). Then H satisfies the convergence condition (ii) of Theorem 5.3.*

Proposition 5.5. *In the setting of Theorem 4.2, let $H \subseteq C^1(\mathbb{T}^d) \times C(\mathbb{T}^d \times E')$ be a multi-valued operator satisfying conditions (T1) and (T2). Then for $\tau > 0$ and $h \in C(\mathbb{T}^d)$, the comparison principle is satisfied for viscosity sub- and supersolutions of $(1 - \tau H)u = h$.*

Proof of Theorem 4.2. By Proposition 5.4, conditions (i) and (ii) of Theorem 5.3 hold with the single operator $H = H_\dagger = H_\ddagger$. By Proposition 5.5, the comparison principle is satisfied for $(1 - \tau H)u = h$, and hence condition (iii) of Theorem 5.3 holds with a single operator $H = H_\dagger = H_\ddagger$. Therefore the large-deviation principle follows by Theorem 5.3. □

Proof of Proposition 5.4. We recall that with $E_n = E_n^X \times \{1, \dots, J\}$ and $\iota_n : E_n^X \to \mathbb{T}^d$ of Condition 3.4, the state spaces are related as in the following diagram,



where $\eta_n : E_n \to \mathbb{T}^d$ is defined by $\eta_n(x, i) = \iota_n(x)$ and $\eta_n' : E_n \to E'$ is a continuous map. In the notation of Theorem 5.3, we have $E = \mathbb{T}^d$.

For verifying the general condition (i) of Theorem 5.3 on the approximating state spaces A_n^q, we take the singleton $Q = \{q\}$ and set $A_n^q := E_n$. Then part (a) holds, and parts (b) and (d) are a consequence of Condition 3.4 on E_n, which says that for any $x \in \mathbb{T}^d$, there exist $x_n \in E_n^X$ such that $\iota_n(x_n) \to x$. Part (c) follows by taking the compact sets $K_1^q := \mathbb{T}^d$ and $K_2^q := \mathbb{T}^d \times E'$, because then

$$\inf_{x \in K_1^q} d_E(\eta_n(y), x) = \inf_{x \in \mathbb{T}^d} d_E(\eta_n(y), x) = 0$$

for any $n \in \mathbb{N}$ and any $y \in E_n$. Hence, for all $n \in N$,

$$\sup_{y \in E_n} \inf_{x \in K_1^q} d_E(\eta_n(y), x) = 0.$$

The other convergence condition in part (c) follows similarly.

We verify the convergence Condition (ii) of Theorem 5.3. By (T1), part (C2), there exist $f_n \in \mathcal{D}(H_n)$ such that

$$\|f \circ \eta_n - f_n\|_{L^\infty(E_n)} \xrightarrow{n\to\infty} 0 \quad \text{and} \quad \|H_{f,\varphi} \circ (\eta_n, \eta_n') - H_n f_n\|_{L^\infty(E_n)} \xrightarrow{n\to\infty} 0.$$

With these f_n, both conditions (ii)(a) and (ii)(b) are simultaneously satisfied for the operator $H = H_\dagger = H_\ddagger$. For example, regarding (ii)(a):

- The existence of f_n with the required boundedness,

$$\sup_{n\in\mathbb{N}}\left(\sup_{y\in E_n}|f_n(y)|+\sup_{y\in E_n}|H_nf_n(y)|\right)<\infty,$$

  follows from the uniform-convergence condition (C2). We have a singleton $Q=\{q\}$ and $A_n^q=E_n$, so that

$$\sup_{y\in A_n^q}|f_n(y)-f(\eta_n(y))|=\|f_n-f\circ\eta_n\|_{L^\infty(E_n)}\xrightarrow{n\to\infty}0,$$

  also by condition (C2).

- Condition (C1) guarantees that for any point $(x,z')\in\mathbb{T}^d\times E'$ there exist $y_n\in E_n$ such that both $\eta_n(y_n)\to x$ and $\eta'_n(y_n)\to z'$ For any such sequence y_n, the bound $\limsup_{n\to\infty}H_nf_n(y_n)\leq g(x,z')$ follows as well from the uniform-convergence condition (C2), where $g=H_{f,\varphi}$.

Part (ii)(b) follows analogously. □

For proving Proposition 5.5, we use two operators H_1,H_2 that are derived from a multi-valued limit H. Define $H_1,H_2:C(E)\to M(E)$ by

$$H_1f(x):=\inf_{\varphi}\sup_{z'\in E'}H_{f,\varphi}(x,z')\quad\text{and}\quad H_2f(x):=\sup_{\varphi}\inf_{z'\in E'}H_{f,\varphi}(x,z'),$$

with equal domains $\mathcal{D}(H_1)=\mathcal{D}(H_2):=\mathcal{D}(H)$. Since the images of H are of the form $H_{f,\varphi}(x,z')=H_\varphi(\nabla f(x),z')$, the operators H_1 and H_2 are as well of the form $H_1f(x)=\mathcal{H}_1(\nabla f(x))$ and $H_2f(x)=\mathcal{H}_2(\nabla f(x))$, with two maps $\mathcal{H}_1,\mathcal{H}_2:\mathbb{R}^d\to\mathbb{R}$. We prove Proposition 5.5 with the following lemmas.

Lemma 5.6 (Local operators admit strong solutions)**.** *Let $H\subseteq C^1(\mathbb{T}^d)\times C(\mathbb{T}^d\times E')$ be a multi-valued limit operator satisfying (T1) of Theorem 4.2. Then for any $\tau>0$ and $h\in C(\mathbb{T}^d)$, viscosity solutions of $(1-\tau H)u=h$ coincide with strong viscosity solutions in the sense of Definition 5.1.*

Lemma 5.7 (H_1 and H_2 are viscosity extensions)**.** *Let H be a multi-valued operator satisfying (T1) and (T2) of Theorem 4.2. For all $h\in C(\mathbb{T}^d)$ and $\tau>0$, strong viscosity subsolutions u_1 of $(1-\tau H)u=h$ are strong viscosity subsolutions of $(1-\tau H_1)u=h$, and strong viscosity supersolutions u_2 of $(1-\tau H)u=h$ are strong viscosity supersolutions of $(1-\tau H_2)u=h$.*

Lemma 5.8 (H_1 and H_2 are ordered)**.** *Let H be a multi-valued operator satisfying (T1) and (T2) of Theorem 4.2. Then $\mathcal{H}_1(p)\leq\mathcal{H}_2(p)$ for all $p\in\mathbb{R}^d$.*

The lemmas are proven further below.

Proof of Proposition 5.5. Let u_1 be a subsolution and u_2 be a supersolution of the equation $(1-\tau H)u=h$. By Lemma 5.6, u_1 is a strong subsolution and u_2 a strong supersolution of $(1-\tau H)u=h$, respectively. By Lemma 5.7, u_1 is a strong subsolution of $(1-\tau H_1)u=h$, and u_2 is a strong supersolution of H_2.

With that, we establish below the inequality

$$\max_{\mathbb{T}^d}(u_1-u_2)\leq\tau\left[\mathcal{H}_1(p_\delta)-\mathcal{H}_2(p_\delta)\right]+h(x_\delta)-h(x'_\delta),\tag{5.3}$$

with some $x_\delta,x'_\delta\in\mathbb{T}^d$ such that $\mathrm{dist}(x_\delta,x'_\delta)\to0$ as $\delta\to0$, and certain $p_\delta\in\mathbb{R}^d$. Then using that $h\in C(\mathbb{T}^d)$ is uniformly continuous since $\mathbb{T}^d$ is compact, and that $\mathcal{H}_1(p_\delta)\leq\mathcal{H}_2(p_\delta)$ by Lemma 5.8, we can further estimate as

$$\max_{\mathbb{T}^d}(u_1-u_2)\leq h(x_\delta)-h(x'_\delta)\leq\omega_h(\mathrm{dist}(x_\delta,x'_\delta)),$$

where $\omega_h : [0,\infty) \to [0,\infty)$ is a modulus of continuity satisfying $\omega_h(r_\delta) \to 0$ for $r_\delta \to 0$. Then $(u_1 - u_2) \le 0$ follows by taking the limit $\delta \to 0$.

We are left with proving (5.3). Define $\Phi_\delta : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}$ by

$$\Phi_\delta(x,x') := u_1(x) - u_2(x') - \frac{\Psi(x,x')}{2\delta},$$

where

$$\Psi(x,x') := \sum_{j=1}^d \sin^2\left(\pi(x_j - x_j')\right), \qquad \text{for all } x,x' \in \mathbb{T}^d. \tag{5.4}$$

Then $\Psi \ge 0$, and $\Psi(x,x') = 0$ holds if and only if $x = x'$, and

$$\nabla_1\left[\Psi(\cdot,x')\right](x) = -\nabla_2\left[\Psi(x,\cdot)\right](x') \qquad \text{for all } x,x' \in \mathbb{T}^d. \tag{5.5}$$

By boundedness and upper semicontinuity of u_1 and $(-u_2)$, and compactness of $\mathbb{T}^d \times \mathbb{T}^d$, for each $\delta > 0$ there exists a pair $(x_\delta, x_\delta') \in \mathbb{T}^d \times \mathbb{T}^d$ such that

$$\Phi_\delta(x_\delta,x_\delta') = \max_{x,x'} \Phi_\delta(x,x').$$

Since $\Phi_\delta(x_\delta,x_\delta) \le \Phi_\delta(x_\delta,x_\delta')$ and u_2 is bounded, we obtain

$$\Psi(x_\delta,x_\delta') \le 2\delta\left(u_2(x_\delta) - u_2(x_\delta')\right) \le 4\delta\|u_2\|_{L^\infty(\mathbb{T}^d)} = \mathcal{O}(\delta).$$

Hence $\Psi(x_\delta,x_\delta') \to 0$ as $\delta \to 0$. In order to use the sub- and supersolution properties of u_1 and u_2, introduce the smooth test functions f_1^δ and f_2^δ as

$$f_1^\delta(x) := u_2(x_\delta') + \frac{\Psi(x,x_\delta')}{2\delta} \quad \text{and} \quad f_2^\delta(x') := u_1(x_\delta) - \frac{\Psi(x_\delta,x')}{2\delta}.$$

Since Ψ is smooth, $f_1^\delta, f_2^\delta \in C^\infty(\mathbb{T}^d) \subseteq \mathcal{D}(H)$ are both in the domain of H, and hence in the domain of H_1 and H_2, respectively. Furthermore, $(u_1 - f_1^\delta)$ has a maximum at $x = x_\delta$, and $(f_2^\delta - u_2)$ has a maximum at $x' = x_\delta'$, by definition of (x_δ, x_δ') and Φ_δ. Since u_1 is a strong subsolution of $(1-\tau H_1)u = h$,

$$u_1(x_\delta) - \tau H_1 f_1^\delta(x_\delta) - h(x_\delta) \le 0,$$

and since u_2 is a strong supersolution of $(1-\tau H_2)u = h$,

$$u_2(x_\delta') - \tau H_2 f_2^\delta(x_\delta') - h(x_\delta') \ge 0.$$

Thereby, we can estimate $\max(u_1 - u_2)$ as

$$\begin{aligned}
\max_{\mathbb{T}^d}(u_1 - u_2) &= \max_{x\in\mathbb{T}^d} \Phi_\delta(x,x) \\
&\le \Phi_\delta(x_\delta,x_\delta') = u_1(x_\delta) - u_2(x_\delta') - \frac{\Psi(x_\delta,x_\delta')}{2\delta} \\
&\le u_1(x_\delta) - u_2(x_\delta') \\
&\le \tau\left[H_1 f_1^\delta(x_\delta)) - H_2 f_2^\delta(x_\delta')\right] + h(x_\delta) - h(x_\delta') \\
&= \tau\left[\mathcal{H}_1(\nabla f_1^\delta(x_\delta)) - \mathcal{H}_2(\nabla f_2^\delta(x_\delta'))\right] + h(x_\delta) - h(x_\delta').
\end{aligned}$$

By (5.5), $\nabla f_1^\delta(x_\delta) = \nabla f_2^\delta(x_\delta') =: p_\delta \in \mathbb{R}^d$, which establishes (5.3), and thereby finishes the proof. ☐

The rest of the section, we prove lemmas 5.6, 5.7 and 5.8. Regarding Lemma 5.6, a proof for single valued operators is given in [13, Lemma 9.9].

Proof of Lemma 5.6. Let $\tau > 0$, $h \in C(\mathbb{T}^d)$. We verify that subsolutions are strong subsolutions. Let u_1 be a subsolution of $(1-\tau H)u = h$ and $(f, H_{f,\varphi}) \in H$, and let $x \in \mathbb{T}^d$ be such that $(u_1 - f)(x) = \sup(u_1 - f)$.

The function $\tilde{f}$ defined by $\tilde{f}(x') := \Psi(x', x)$, with $\Psi(x', x)$ defined by (5.4), is smooth and therefore $\tilde{f}$ is in the domain $\mathcal{D}(H)$. Then x is the unique maximal point of $(u_1 - (f + \tilde{f}))$,

$$(u_1 - (f + \tilde{f}))(x) = \sup_{\mathbb{T}^d}(u_1 - (f + \tilde{f})).$$

Since u_1 is a subsolution, there exists a point $z' \in E'$ such that

$$u_1(x) - \tau H_{f+\tilde{f},\varphi}(x, z') - h(x) \leq 0.$$

Using $\nabla \tilde{f}(x) = 0$ and that H depends only on gradients by (T1), we obtain

$$H_{f+\tilde{f},\varphi}(x, z') = H_\varphi \left((\nabla f + \nabla \tilde{f})(x), z' \right) = H_\varphi(\nabla f(x), z') = H_{f,\varphi}(x, z').$$

Hence

$$u_1(x) - \tau H_{f,\varphi}(x, z') - h(x) \leq 0.$$

Thus u_1 is a strong subsolution. The argument is similar for the supersolution case, where one can use $(-\tilde{f})$.

Vice versa, when given a strong sub- or supersolution u_1 or u_2, for every $f \in \mathcal{D}(H)$, $(u_1 - f)$ and $(f - u_2)$ attain their suprema at some $x_1, x_2 \in \mathbb{T}^d$ due to the continuity assumptions on the domain of H, the semi-continuity properties of u_1 and u_2, and compactness of $\mathbb{T}^d$. By the strong solution properties, the sub- and supersolution inequalities follow. □

Proof of Lemma 5.7. Let u_1 be a strong subsolution of $(1 - \tau H)u = h$, that is for any $(f, H_{f,\varphi})$, if $(u_1 - f)(x) = \sup(u_1 - f)$ for a point $x \in \mathbb{T}^d$, then there exists a point $z' \in E'$ such that

$$u_1(x) - \tau H_{f,\varphi}(x, z') - h(x) \leq 0. \tag{5.6}$$

Let $f \in \mathcal{D}(H_1) = \mathcal{D}(H)$ and $x \in \mathbb{T}^d$ be such that $(u_1 - f)(x) = \sup(u_1 - f)$. For any φ there exists a point $z' \in E'$ such that the above subsolution inequality (5.6) holds. Therefore for all x,

$$u_1(x) - h(x) \leq \tau \sup_{z' \in E'} H_{f,\varphi}(x, z').$$

Since the point $x \in \mathbb{T}^d$ is independent of φ, we obtain

$$u_1(x) - \tau H_1 f(x) - h(x) \overset{\text{def}}{=} u_1(x) - \tau \inf_\varphi \sup_{z' \in E'} H_{f,\varphi}(x, z') - h(x) \leq 0.$$

The argument is similar for supersolutions. □

Proof of Lemma 5.8. By assumption, for every $p \in \mathbb{R}^d$ there exists a function $\varphi_p \in C(E')$ such that for all $z' \in E'$,

$$H_{\varphi_p}(p, z') = \mathcal{H}(p).$$

Thus

$$\sup_{z' \in E'} H_{\varphi_p}(p, z') = \mathcal{H}(p) = \inf_{z' \in E'} H_{\varphi_p}(p, z').$$

Taking the infimum and supremum over φ, we find

$$\begin{aligned} \mathcal{H}_1(p) &= \inf_\varphi \sup_{z'} H_\varphi(p, z') \\ &\leq \sup_{z'} H_{\varphi_p}(p, z') = \mathcal{H}(p) = \inf_{z'} H_{\varphi_p}(p, z') \\ &\leq \sup_\varphi \inf_{z'} H_\varphi(p, z') = \mathcal{H}_2(p), \end{aligned}$$

which finishes the proof. □

6 Proof of action-integral representation

In this section we prove Theorem 4.3, the action-integral representation of the rate function of Theorem 4.2, by following the strategy outlined in [13, Chapter 8]. We first briefly summarize the strategy in Section 6.1, specialized to our setting.

6.1 Strategy of proof

Let $\mathcal{H} = \mathcal{H}(p)$ be the Hamiltonian of Theorem 4.3 and let $\mathcal{L} = \mathcal{L}(v)$ be the associated Lagrangian defined by

$$\mathcal{L}(v) := \sup_{p\in\mathbb{R}^d} \left[p \cdot v - \mathcal{H}(p)\right]. \tag{6.1}$$

Define $V_{\mathrm{NS}}(t) : C(\mathbb{T}^d) \to C(\mathbb{T}^d)$ by

$$V_{\mathrm{NS}}(t)f(x) = \sup_{\substack{\gamma\in \mathrm{AC}_{\mathbb{T}^d}[0,\infty)\\ \gamma(0)=x}} \left[f(\gamma(t)) - \int_0^t \mathcal{L}(\partial_s\gamma(s))\,\mathrm{d}s\right], \tag{6.2}$$

where $\mathrm{AC}_{\mathbb{T}^d}[0,\infty)$ is the set of absolutely continuous paths in the torus. The map $V_{\mathrm{NS}}(t)$ is the *Nisio semigroup* with cost function $\mathcal{L}$. In Definition 8.1 and Equation (8.10) in [13], the Nisio semigroup is defined by means of relaxed controls in order to cover a general class of possible cost functions. Since the Lagrangian $\mathcal{L}(v)$ is convex, the semigroup $V_{\mathrm{NS}}(t)$ equals the semigroup given in (8.10) of [13], which can be seen by using that $\lambda_s = \delta_{\partial_s x(s)}$ is an admissible control and by applying Jensen's inequality. Such an argument is given for example in Theorem 10.22 in [13].

The rate function $\mathcal{I}$ of Theorem 4.2 is given in terms of a limiting semigroup $V(t)$ as shown in equations (5.1) and (5.2). The desired action-integral representation follows if the semigroup $V(t)$ of Theorem 4.2 is equal to the Nisio semigroup $V_{\mathrm{NS}}(t)$ defined by (6.2). In [13, Chapter 8], the equality of semigroups is traced back to conditions on their generators. In our case, the generator of the limiting seimgroup is the limiting multi-valued operator H of Theorem 4.2, and the generator of the Nisio semigroup is an operator $\mathbf{H}$ defined by the Hamiltonian $\mathcal{H}(p)$. We summarize in Proposition 6.1 below that the generators satisfy the required conditions of [13, Chapter 8] and show that these conditions suffice to prove the action-integral representation.

6.2 Proof of Theorem 4.3

In this section, we first prove Theorem 4.3 by means of Proposition 6.1 below. The rest of the section is then devoted to proving Proposition 6.1.

Proposition 6.1. *Under the same assumptions of Theorems 4.2 and 4.3, define the operator $\mathbf{H} : \mathcal{D}(\mathbf{H}) \subseteq C^1(\mathbb{T}^d) \to C(\mathbb{T}^d)$ on the domain $\mathcal{D}(\mathbf{H}) = \mathcal{D}(H)$ by setting $\mathbf{H}f(x) := \mathcal{H}(\nabla f(x))$. Let $\tau > 0$ and $h \in C(\mathbb{T}^d)$. Then:*

- *(i) The Lagrangian* (6.1) *and the operator $\mathbf{H}$ satisfy Conditions 8.9, 8.10 and 8.11 of [13], with the set of controls $U = \mathbb{R}^d$, operator $Af(x,u) = \nabla f(x) \cdot u$, cost function $L(x,u) = \mathcal{L}(u)$, and $\mathbf{H}_\dagger = \mathbf{H}_\ddagger = \mathbf{H}$.*
- *(ii) The comparison principle (Definition 5.2) holds for viscosity sub- and supersolutions of $(1-\tau\mathbf{H})u = h$.*
- *(iii) Every viscosity solution u of $(1-\tau H)u = h$ is also a viscosity solution of $(1-\tau\mathbf{H})u = h$.*

Proof of Theorem 4.3. Let $V(t)$ be the semigroup obtained in Theorem 4.2 and let $V_{\mathrm{NS}}(t)$ bet the Nisio semigroup (6.2). We shall verify that $V(t) = V_{\mathrm{NS}}(t)$. Then by [13, Theorem 8.14], the rate function of Theorem 4.2 (given by (5.1)) satisfies the control representation (8.18) of [13]. The action-integral representation follows from this control representation by applying Jensen's inequality.

By [13, Theorem 8.27], we obtain $V_{\mathrm{NS}}(t) = \mathbf{V}(t)$, where the semigroup $\mathbf{V}(t)$ is defined by

$$\mathbf{V}(t) = \lim_{m\to\infty}\left[\left(1 - \frac{t}{m}\mathbf{H}\right)^{-1}\right]^m. \tag{6.3}$$

The conditions of Theorem 8.27 are satisfied since Conditions 8.9, 8.10 and 8.11 of [13] are satisfied by Item (i), and since the comparison principle holds by Item (ii).

By [13, Corollary 8.29], we obtain $V(t) = \mathbf{V}(t)$. The conditions of Corollary 8.29 are satisfied: Item (iii) above corresponds to Item a) of Corollary 8.29, the conditions of [13, Theorem 6.14] are satisfied under the assumptions of our Theorem 4.2, the conditions of [13, Theorem 8.27] are satisfied for the same reasons as mentioned above, and $D_\alpha = \mathcal{D}(H)$. □

Proof of (i) in Proposition 6.1. We first show that the following Items (a), (b), (c) imply Conditions 8.9, 8.10 and 8.11 of [13], which are formulated in order to cover a more general and non-compact setting.

(a) The function $\mathcal{L} : \mathbb{R}^d \to [0,\infty]$ is lower semicontinuous and for every $C \geq 0$, the level set $\{v \in \mathbb{R}^d \,:\, \mathcal{L}(v) \leq C\}$ is relatively compact in $\mathbb{R}^d$.

(b) For all $f \in \mathcal{D}(H)$ there exists a right continuous, nondecreasing function $\psi_f : [0,\infty) \to [0,\infty)$ such that for all $(x_0, v) \in \mathbb{T}^d \times \mathbb{R}^d$,

$$|\nabla f(x_0)\cdot v| \leq \psi_f(\mathcal{L}(v)) \qquad \text{and} \qquad \lim_{r\to\infty}\frac{\psi_f(r)}{r} = 0.$$

(c) For each $x_0 \in E$ and every $f \in \mathcal{D}(\mathbf{H})$, there exists an absolutely continuous path $x : [0,\infty) \to \mathbb{T}^d$ such that

$$\int_0^t \mathcal{H}(\nabla f(x(s)))\,ds = \int_0^t \left[\nabla f(x(s))\cdot\dot{x}(s) - \mathcal{L}(\dot{x}(s))\right] ds. \tag{6.4}$$

Regarding Items (1)-(5) of [13, Condition 8.9], the operator $Af(x,v) := \nabla f(x)\cdot v$ defined on the domain $\mathcal{D}(A) = \mathcal{D}(H)$ satisfies Item (1). For Item (2), we can take $\Gamma = \mathbb{T}^d \times \mathbb{R}^d$, and for $x_0 \in \mathbb{T}^d$, take the pair (x,λ) with $x(t) = x_0$ and $\lambda(dv\times dt) = \delta_0(dv)\times dt$. Item (3) is a consequence of the above Item (a). Item (4) holds since $\mathbb{T}^d$ is compact. Item (5) is implied by the above Item (b). Condition 8.10 is implied by Condition 8.11 and the fact that $\mathbf{H}1 = 0$, see Remark 8.12 (e) in [13]. Finally, Condition 8.11 is implied by the above Item (c), with the control $\lambda(dv\times dt) = \delta_{\partial_t x(t)}(dv)\times dt$.

We turn to verifying Items (a), (b) and (c). Since $\mathcal{H}(0) = 0$, we have $\mathcal{L} \geq 0$. The Legendre-transform $\mathcal{L}$ is convex, and lower semicontinuous since the map $\mathcal{H}(p)$ is convex and finite-valued, hence in particular continuous. For $C \geq 0$, we prove that the set $\{v \in \mathbb{R}^d \,:\, \mathcal{L}(v) \leq C\}$ is bounded, and hence is relatively compact. For any $p \in \mathbb{R}^d$ and $v \in \mathbb{R}^d$, we have $p\cdot v \leq \mathcal{L}(v) + \mathcal{H}(p)$. Thereby, if $\mathcal{L}(v) \leq C$, then $|v| = \sup_{|p|=1} p\cdot v \leq \sup_{|p|=1}\left[\mathcal{L}(v) + \mathcal{H}(p)\right] \leq C + C_1$, where C_1 exists due to continuity of $\mathcal{H}$. Then for $R := C + C_1$, $\{v \,:\, \mathcal{L}(v) \leq C\} \subseteq \{v \,:\, |v| \leq R\}$, thus $\{\mathcal{L} \leq C\}$ is a bounded subset in $\mathbb{R}^d$.

Item (b) can be proven as in [13, Lemma 10.21]. We give the proof here. Let $f \in \mathcal{D}(H)$. There exists a constant C_f such that for all (x_0, v), we have

$$|\nabla f(x_0)\cdot v| \leq C_f \cdot |v|.$$

For $s \geq 0$, define the map $\varphi(s)$ by

$$\varphi(s) := s \inf_{|v| \geq s} \frac{\mathcal{L}(v)}{|v|}.$$

Let $\psi_f(r) := C_f \cdot \varphi^{-1}(r)$ with $\varphi^{-1}(r) = \inf\{w \,:\, \varphi(w) \geq r\}$. By monotonicity of φ,

$$\varphi(C_f^{-1} |\nabla f(x_0) \cdot v|) \leq \varphi(|v|) \leq \mathcal{L}(v).$$

Hence by monotonicity of ψ_f, we find $|\nabla f(x_0) \cdot v| \leq \psi_f(\mathcal{L}(v))$. The map $\mathcal{L}(v)$ is superlinear, because $\mathcal{H}(p)$ is convex. Therefore $s^{-1}\varphi(s) \to +\infty$ as $s \to \infty$, and consequently $r^{-1}\psi_f(r) \to 0$ as $r \to \infty$.

We finish the proof by verifying Item (c). This is shown in [25, Lemma 3.2.3] under the assumption of continuous differentiability of $\mathcal{H}(p)$, by solving a differential equation under a global-boundedness assumption. Here, we verify Item (c) under the milder assumption of convexity of $\mathcal{H}(p)$ by solving a suitable subdifferential equation. For $p_0 \in \mathbb{R}^d$, define the subdifferential $\partial\mathcal{H}(p_0)$ at p_0 as the set

$$\partial\mathcal{H}(p_0) := \{\xi \in \mathbb{R}^d \mid \forall p \in \mathbb{R}^d \,:\, \mathcal{H}(p) \geq \mathcal{H}(p_0) + \langle \xi, p - p_0 \rangle\}.$$

We shall solve for any $f \in C^1(\mathbb{T}^d)$ the subdifferential equation $\dot{x} \in \partial\mathcal{H}(\nabla f(x))$. This means we show that for any initial condition $x_0 \in \mathbb{T}^d$, there exists an absolutely continuous path $x : [0,\infty) \to \mathbb{T}^d$ satisfying both $x(0) = x_0$ and $\dot{x}(t) \in \partial\mathcal{H}(\nabla f(x(t)))$ almost everywhere on $[0,\infty)$. Then (6.4) follows by noting that $\mathcal{H}(\nabla f(y)) \geq \nabla f(y) \cdot v - \mathcal{L}(v)$ for all $y \in \mathbb{T}^d$ and $v \in \mathbb{R}^d$, by convex duality. In particular, $\mathcal{H}(\nabla f(x(s))) \geq \nabla f(x(s)) \cdot \dot{x}(s) - \mathcal{L}(\dot{x}(s))$, and integrating gives one inequality in (6.4). Regarding the other inequality, since $\dot{x} \in \partial\mathcal{H}(\nabla f(x))$, we know that for almost every $t \in [0,\infty)$ and for all $p \in \mathbb{R}^d$, we have $\mathcal{H}(p) \geq \mathcal{H}(\nabla f(x(t))) + \dot{x}(t) \cdot (p - \nabla f(x(t)))$. Therefore, a.e. on $[0,\infty)$,

$$\begin{aligned}\mathcal{H}(\nabla f(x(t))) &\leq \nabla f(x(t)) \cdot \dot{x}(t) - \sup_{p \in \mathbb{R}^d} [p \cdot \dot{x}(t) - \mathcal{H}(p)] \\ &= \nabla f(x(t)) \cdot \dot{x}(t) - \mathcal{L}(\dot{x}(t)),\end{aligned}$$

and integrating gives the other inequality.

For solving the subdifferential equation, define $F : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ by $F(x) := \partial\mathcal{H}(\nabla f(x))$, where the function $f \in C^1(\mathbb{T}^d)$ is regarded as a periodic function on $\mathbb{R}^d$. We apply Lemma 5.1 in [4] for solving $\dot{x} \in F(x)$. The conditions of Lemma 5.1 in the case of $\mathbb{R}^d$ are satisfied if the following holds: $\sup_{x \in \mathbb{R}^d} \|F(x)\|$ is finite, for all $x \in \mathbb{R}^d$, the set $F(x)$ is non-empty, closed and convex, and the map $x \mapsto F(x)$ is upper semicontinuous.

For $\xi \in F(x)$, note that for all $p \in \mathbb{R}^d$ $\xi \cdot (p - \nabla f(x)) \leq \mathcal{H}(p) - \mathcal{H}(\nabla f(x))$. Therefore, by shifting $p = p' + \nabla f(x)$, we obtain for all $p' \in \mathbb{R}^d$ that $\xi \cdot p' \leq \mathcal{H}(p' + \nabla f(x)) - \mathcal{H}(\nabla f(x))$. By continuous differentiability and periodicity of f, and continuity of $\mathcal{H}$, the right-hand side is bounded in x, and we obtain

$$\begin{aligned}\sup_{x \in \mathbb{R}^d} \sup_{\xi \in F(x)} |\xi| &= \sup_{x \in \mathbb{R}^d} \sup_{\xi \in F(x)} \sup_{|p'|=1} \xi \cdot p' \\ &\leq \sup_{x \in \mathbb{R}^d} \sup_{\xi \in F(x)} \sup_{|p'|=1} [\mathcal{H}(p' + \nabla f(x)) - \mathcal{H}(\nabla f(x))] < \infty.\end{aligned}$$

For any $x \in \mathbb{R}^d$, the set $F(x)$ is non-empty, since the subdifferential of a proper convex function $\mathcal{H}(\cdot)$ is nonempty at points where $\mathcal{H}(\cdot)$ is finite and continuous (see e.g. [36, Th. 23.4]). Furthermore, $F(x)$ is convex and closed, which follows from the properties of a subdifferential set.

Regarding upper semicontinuity, recall the definition from [4]: the map $F : \mathbb{R}^d \to 2^{\mathbb{R}^d} \setminus \{\emptyset\}$ is upper semicontinuous if for all closed sets $A \subseteq \mathbb{R}^d$, the set $F^{-1}(A) \subseteq \mathbb{R}^d$ is closed, where $F^{-1}(A) = \{x \in \mathbb{R}^d \mid F(x) \cap A \neq \emptyset\}$. Let $A \subseteq \mathbb{R}^d$ be closed and $x_n \to x$ in $\mathbb{R}^d$, with $x_n \in F^{-1}(A)$. That means for all $n \in \mathbb{N}$ that the sets $\partial\mathcal{H}(\nabla f(x_n)) \cap A$ are non-empty, and consequently, there exists a sequence $\xi_n \in F(x_n) \cap A$. We proved above that the set $F(y) \cap A$ is uniformly bounded in $y \in \mathbb{R}^d$. Hence the sequence ξ_n is bounded, and passing to a subsequence if necessary, it converges to some ξ. By definition of $F(x_n)$, for all $p \in \mathbb{R}^d$,

$$\xi_n(p - \nabla f(x_n)) \leq \mathcal{H}(p) - \mathcal{H}(\nabla f(x_n)).$$

Passing to the limit, we obtain that for all $p \in \mathbb{R}^d$,

$$\xi(p - \nabla f(x)) \leq \mathcal{H}(p) - \mathcal{H}(\nabla f(x)).$$

This implies by definition that $\xi \in \partial\mathcal{H}(\nabla f(x))$. Since $\xi_n \in A$ and A is closed, we have $\xi \in A$. Hence $x \in F^{-1}(A)$, and $F^{-1}(A)$ is indeed closed. □

Proof of (ii) in Proposition 6.1. The comparison principle for the operator $\mathbf{H}$ follows from the fact that $\mathbf{H}f = \mathcal{H}(\nabla f)$ depends on x only via gradients. Indeed, for subsolutions u_1 and supersolutions u_2 of $(1 - \tau\mathbf{H})u = h$, we have $\max(u_1 - u_2) \leq \tau[\mathcal{H}(\nabla f_1(x_\delta)) - \mathcal{H}(\nabla f_2(x'_\delta))] + h(x_\delta) - h(x'_\delta)$, with test functions $f_1, f_2 \in \mathcal{D}(H)$ satisfying $\nabla f_1(x_\delta) = \nabla f_2(x'_\delta)$, and $\mathrm{dist}(x_\delta, x'_\delta) \to 0$ as $\delta \to 0$. Therefore $\mathcal{H}(\nabla f_1(x_\delta)) - \mathcal{H}(\nabla f_2(x'_\delta)) = 0$, and $\max(u_1 - u_2) \leq 0$ follows by taking the limit $\delta \to 0$. □

Proof of (iii) in Proposition 6.1. Let $u \in C(\mathbb{T}^d)$ be a viscosity solution of the equation $(1-\tau H)u = h$. By Lemmas 5.6 and 5.7, u is a strong viscosity subsolution of $(1 - \tau H_1)u = h$ and a strong viscosity supersolution of $(1 - \tau H_2)u = h$. In the proof of Lemma 5.8 we obtained $\mathcal{H}_1 \leq \mathcal{H} \leq \mathcal{H}_2$, which in particular implies the inequalities $-H_1 \geq -\mathbf{H} \geq -H_2$. With that, we find that u is both a strong viscosity sub- and supersolution of $(1 - \tau\mathbf{H})u = h$. □

7 Proof of large deviations for molecular motors

In this section, we consider the stochastic process (X^n, I^n) of Definition 4.4 and prove Theorems 4.6 and 4.7. The generator L_n of (X^n, I^n) is given by

$$L_n f(x,i) \;=\; \frac{1}{n}\frac{1}{2}\Delta f(\cdot,i)(x) + b_i(nx)\cdot\nabla f(\cdot,i)(x) + \sum_{j=1}^J \gamma(n) r_{ij}(nx)\left[f(x,j) - f(x,i)\right],$$

with state space $E_n = \mathbb{T}^d \times \{1,\ldots,J\} = \{(x,i)\}$, drifts $b_i \in C^\infty(\mathbb{T}^d)$, jump rates $r_{ij} \in C^\infty(\mathbb{T}^d;[0,\infty))$, and $\gamma(n) > 0$. We frequently write $f(x,i) = f_i(x)$. The nonlinear generators defined by $H_n f = \frac{1}{n}e^{-nf}L_n e^{nf(\cdot)}$ are given (for $f \in C^2(\mathbb{T}^d) \subset \mathcal{D}(H_n)$) by

$$\begin{aligned} H_n f(x,i) = \frac{1}{n}\frac{1}{2}\Delta f_i(x) + \frac{1}{2}|\nabla f_i(x)|^2 + b_i(nx)\nabla f_i(x) \\ + \frac{1}{n}\gamma(n)\sum_{j=1}^J r_{ij}(nx)\left[e^{n(f(x,j)-f(x,i))} - 1.\right]. \end{aligned} \tag{7.1}$$

7.1 Proof of Theorem 4.6

Verification of (T1) of Theorem 4.2. We have the scaling $\gamma(n) = n$. Choosing the functions $f_n(x,i) = f(x) + \frac{1}{n}\,\varphi\,(nx,i)$, we find

$$H_n f_n(x,i) = \frac{1}{n}\frac{1}{2}\Delta f(x) + \frac{1}{2}\Delta_y \varphi_i\,(nx) + \frac{1}{2}\left|\nabla f(x) + \nabla_y \varphi_i\,(nx)\,\right|^2 + b_i\,(nx)\,(\nabla f(x) + \nabla_y\varphi_i\,(nx)) + \sum_{j=1}^{J} r_{ij}\,(nx)\left[e^{\varphi(nx,j)-\varphi(nx,i)} - 1\right],$$

where ∇_y and Δ_y denote the gradient and Laplacian with respect to the variable $y = nx$, instead of the variable x. The only term of order $\frac{1}{n}$ that remains is $\frac{1}{n}\,\Delta f(x)/2$. This suggests to take the remainder terms as the definition of the multi-valued operator H. In the notation of Theorem 4.2, we choose $E' = \mathbb{T}^d \times \{1,\dots,J\}$ as the state space of the macroscopic variables, and define

$$H := \left\{(f, H_{f,\varphi})\,:\, f \in C^2(\mathbb{T}^d),\ H_{f,\varphi} \in C(\mathbb{T}^d \times E') \text{ and } \varphi \in C^2(E')\right\}, \tag{7.2}$$

with the image functions $H_{f,\varphi} : \mathbb{T}^d \times E' \to \mathbb{R}$ defined by

$$H_{f,\varphi}(x,y,i) := \frac{1}{2}\Delta_y\varphi_i(y) + \frac{1}{2}\left|\nabla f(x) + \nabla_y\varphi_i(y)\right|^2 + b_i(y)(\nabla f(x) + \nabla_y\varphi_i(y)) + \sum_{j=1}^{J} r_{ij}(y)\left[e^{\varphi(y,j)-\varphi(y,i)} - 1\right], \tag{7.3}$$

where we write $\varphi = (\varphi_1,\dots,\varphi_J)$ via the identification $C^2(E') \simeq (C^2(\mathbb{T}^d))^J$. □

Verification of (C1), (C2) and (C3). For verifying (C1), define the maps $\eta_n' : E_n \to E'$ by $\eta_n'(x,i) := (nx,i)$, and recall that the maps $\eta_n : E_n \to \mathbb{T}^d$ are the projections $\eta_n(x,i) := x$. For any $(x,y,i) \in \mathbb{T}^d \times E'$, we search for elements $(y_n,i_n) \in \mathbb{T}^d \times \{1,\dots,J\}$ such that both $\eta_n(y_n,i_n) \to x$ and $\eta_n'(y_n,i_n) \to (y,i)$ as $n \to \infty$. For $d = 1$, the point $y_n := \frac{1}{n}(\lfloor nx \rfloor + y)$ satisfies $y_n \to x$ and $ny_n = y$ in $\mathbb{T}^d$ (i.e. modulo 1). For $d \geq 2$ this construction can be done for each coordinate. Therefore, (C1) holds with $y_n = \frac{1}{n}(\lfloor nx \rfloor + y)$ and $i_n = i$.

Regarding (C2), let $(f, H_{f,\varphi}) \in H$. The function f_n defined by $f_n(x,i) := f(x) + \frac{1}{n}\,\varphi\,(nx,i)$ satisfies

$$\|f \circ \eta_n - f_n\|_{L^\infty(E_n)} = \sup_{(x,i)\in E_n} |f(x) - f_n(x,i)| = \frac{1}{n}\cdot\|\varphi\|_{L^\infty(E_n)} \xrightarrow{n\to\infty} 0,$$

and

$$\begin{aligned}\|H_{f,\varphi}\circ(\eta_n,\eta_n') - H_n f_n\|_{L^\infty(E_n)} &= \sup_{(x,i)\in E_n} |H_{f,\varphi}(x,nx,i) - H_n f_n(x,i)| \\ &= \frac{1}{n}\frac{1}{2}\sup_{(x,i)\in E_n} |\,\Delta f(x)| \leq \frac{1}{n}\frac{1}{2}\sup|\Delta f| \xrightarrow{n\to\infty} 0.\end{aligned}$$

Regarding (C3), the fact that the images $H_{f,\varphi}$ depend on x only via the gradients of f, can be recognized in (7.3). □

Verification of (T2) of Theorem 4.2. Let f be a function in $\mathcal{D}(H) = C^2(\mathbb{T}^d)$ and $x \in \mathbb{T}^d$. We establish the existence of a vector function $\varphi = (\varphi_1,\dots,\varphi_J) \in (C^2(\mathbb{T}^d))^J$ such that for all $(y,i) \in E' = \mathbb{T}^d \times \{1,\dots,J\}$ and some constant $\mathcal{H}(\nabla f(x)) \in \mathbb{R}$, we have

$$H_\varphi(\nabla f(x), y, i) = \mathcal{H}(\nabla f(x)).$$

For the flat torus $E = \mathbb{T}^d$, this means that for fixed $\nabla f(x) = p \in \mathbb{R}^d$, we search for a vector function φ_p such that $\tilde{H}_{\varphi_p}(p, y, i) = \mathcal{H}(p)$ becomes independent of the variables $(y, i) \in E'$. We can find this vector function by solving a principal eigenvalue problem. We prove Item (T2) by the following lemma.

Lemma 7.1. *Let $E' = \mathbb{T}^d \times \{1, \ldots, J\}$ and H be the limit operator (7.2). Then:*

(a) For $f \in \mathcal{D}(H)$, the limiting images $H_\varphi(\nabla f(x), y, i)$ are of the form

$$H_\varphi(\nabla f(x), y, i) = e^{-\varphi(y,i)} \left[(B_p + V_p + R)e^\varphi\right](y, i),$$

with $p = \nabla f(x) \in \mathbb{R}^d$, and operators $B_p, V_p, R : C^2(E') \to C(E')$ defined as

$$\begin{aligned}
(B_p h)(y, i) &:= \frac{1}{2}\Delta_y h(y, i) + (p + b_i(y)) \cdot \nabla_y h(y, i)\\
(V_p h)(y, i) &:= \left(\frac{1}{2}p^2 + p \cdot b_i(y)\right) h(y, i),\\
(R\, h)(y, i) &:= \sum_{j=1}^{J} r_{ij}(y) \left[h(y, j) - h(y, i)\right].
\end{aligned}$$

(b) For any $p \in \mathbb{R}^d$, there exists an eigenfunction $g_p = (g_p^1, \ldots, g_p^J) \in (C^2(\mathbb{T}^d))^J$ with strictly positive component functions, $g_p^i > 0$ on $\mathbb{T}^d$ for $i = 1, \ldots, J$, and an eigenvalue $\mathcal{H}(p) \in \mathbb{R}$ such that

$$[B_p + V_p + R]\, g_p = \mathcal{H}(p)\, g_p. \tag{7.4}$$

Now (T2) follows by (a) and (b) in Lemma 7.1, since with $\varphi_p := \log g_p$,

$$\begin{aligned}
H_{\varphi_p}(p, y, i) &\overset{(a)}{=} e^{-\varphi_p(y,i)} \left[B_p + V_p + R\right] e^{\varphi_p(y,i)}\\
&= \frac{1}{g_p(y, i)} \left[B_p + V_p + R\right] g_p(y, i) \overset{(b)}{=} \mathcal{H}(p).
\end{aligned}$$

Proof of Lemma 7.1. Writing $p = \nabla f(x)$, Item (a) follows directly by regrouping the terms in (7.3). Regarding Item (b), $[B_p + V_p + R]\, g_p = \mathcal{H}(p) g_p$ is a system of weakly-coupled nonlinear elliptic PDEs on the flat torus. They are weakly coupled in the sense that the component functions g_p^i are only coupled in the lowest order terms by means of the operator R, while the operators B_p and V_p act solely on the diagonal. By Proposition B.2, there exists a $\lambda(p)$ and $g_p > 0$ such that $[-B_p - V_p - R]\, g_p = \lambda(p) g_p$. Thereby, $[B_p + V_p + R]\, g_p = \mathcal{H}(p) g_p$ follows with the same eigenfunction $g_p > 0$ and the principal eigenvalue $\mathcal{H}(p) = -\lambda(p)$. This finishes the verification of (T2). □

Verification of (T3) of Theorem 4.3. We prove that the principal eigenvalue $\mathcal{H}(p)$ of Lemma 7.1 is convex in $p \in \mathbb{R}^d$ and satisfies $\mathcal{H}(0) = 0$. By Proposition B.2, the eigenvalue $\mathcal{H}(p) = -\lambda(p)$ admits the representation

$$\begin{aligned}
\mathcal{H}(p) &= -\sup_{g>0} \inf_{z' \in E'} \left\{ \frac{1}{g(z')} \left[(-B_p - V_p - R)g\right](z') \right\}\\
&= \inf_{g>0} \sup_{z' \in E'} \left\{ \frac{1}{g(z')} \left[(B_p + V_p + R)g\right](z') \right\}\\
&= \inf_{\varphi} \sup_{z' \in E'} \left\{ e^{-\varphi(z')} \left[(B_p + V_p + R)e^\varphi\right](z') \right\} =: \inf_{\varphi} \sup_{z' \in E'} F(p, \varphi)(z'),
\end{aligned}$$

with a map F defined by

$$F(p,\varphi)(y,i) := \frac{1}{2}\Delta_y\varphi_i(y) + \frac{1}{2}|\nabla_y\varphi_i(y)+p|^2 + b_i(y)(\nabla_y\varphi_i(y)+p) + \sum_{j=1}^J r_{ij}(y)\left[e^{\varphi_j(y)-\varphi_i(y)}-1\right].$$

The map F is jointly convex in p and φ. For the eigenfunction $\varphi=\varphi_p$, equality holds in the sense that for any $z\in E'$, we have $\mathcal{H}(p)=F(p,\varphi_p)(z)$. Therefore, we obtain for $\tau\in[0,1]$ and any $p_1,p_2\in\mathbb{R}^d$ with corresponding eigenfunctions $g_1=e^{\varphi_1}$ and $g_2=e^{\varphi_2}$ that

$$\begin{aligned}\mathcal{H}(\tau p_1+(1-\tau)p_2) &= \inf_{\varphi}\sup_{E'} F\left(\tau p_1+(1-\tau)p_2,\varphi\right)\\ &\le \sup_{E'} F\left(\tau p_1+(1-\tau)p_2, \tau\varphi_1+(1-\tau)\varphi_2\right)\\ &\le \sup_{E'}\left[\tau F(p_1,\varphi_1)+(1-\tau)F(p_2,\varphi_2)\right]\\ &\le \tau\sup_{E'}F(p_1,\varphi_1)+(1-\tau)\sup_{E'}F(p_2,\varphi_2)\\ &= \tau\mathcal{H}(p_1)+(1-\tau)\mathcal{H}(p_2).\end{aligned}$$

Regarding the claim $\mathcal{H}(0)=0$, we choose the constant function $\varphi=(1,\dots,1)$ in the variational representation of $\mathcal{H}(p)$. Thereby, we obtain the estimate $\mathcal{H}(0)\le 0$. For the opposite inequality, we show that for any $\varphi\in C^2(E')$

$$\lambda(\varphi) := \sup_{z'\in E'}\left\{e^{-\varphi(z')}\left[(B_0+V_0+R)e^{\varphi}\right](z')\right\}\ge 0,$$

which then implies $\mathcal{H}(0)=\inf_\varphi\lambda(\varphi)\ge 0$. Let $\varphi\in C^2(E')$; the continuous function φ on the compact set E' admits a global minimum $z_m=(y_m,i_m)\in E'$. Thereby, noting that $V_0\equiv 0$, we find

$$\begin{aligned}\lambda(\varphi)\ge e^{-\varphi(z_m)}(B_0+R)e^{\varphi(z_m)} &= \underbrace{\frac{1}{2}\Delta_y\varphi(y_m,i_m)}_{\ge 0}+\frac{1}{2}|\underbrace{\nabla_y\varphi(y_m,i_m)}_{=0}|^2\\ &\quad + b^{i_m}(y_m)\cdot\underbrace{\nabla_y\varphi(y_m,i_m)}_{=0}+\sum_{j\neq i}r_{ij}(y_m)\underbrace{\left[e^{\varphi(y_m,j)-\varphi(y_m,i_m)}-1\right]}_{\ge 0}\ge 0.\end{aligned}$$

This finishes the verification of (T3), and thereby the proof of Theorem 4.6. □

7.2 Proof of Theorem 4.7

In this section, we consider the process (X^n,I^n) of Definition 4.4 in the limit regime $\frac{1}{n}\gamma(n)\to\infty$ as $n\to\infty$. As above in the proof of Theorem 4.6, we start with the nonlinear generator H_n given by (7.1), and verify conditions (T1), (T2) and (T3) of Theorems 4.2 and 4.3. Conditions (C1), (C2) and (C3) can be shown similarly as in Section 7.1.

Verification of (T1) of Theorem 4.2. We choose functions $f_n(x,i)$ of the form

$$f_n(x,i)=f(x)+\frac{1}{n}\varphi(nx)+\frac{1}{\gamma(n)}\xi(nx,i).$$

We abbreviate $y = nx$ in the following equation. Computing $H_n f_n$ results in

$$H_n f_n(x,i) = \frac{1}{n}\frac{1}{2}\Delta f(x) + \frac{1}{2}\left[\Delta_y \varphi(y) + \frac{n}{\gamma(n)}\Delta_y \xi_i(y)\right] + \frac{1}{2}\left|\nabla f(x) + \nabla_y \varphi(y) + \frac{n}{\gamma(n)}\nabla_y \xi_i(y)\right|^2 + b_i(y)\left(\nabla f(x) + \nabla_y\varphi(y) + \frac{n}{\gamma(n)}\nabla_y \xi_i(y)\right) + \frac{1}{n}\gamma(n)\sum_{j=1}^J r_{ij}(y)\left[e^{n(\xi(y,j)-\xi(y,i))/\gamma(n)} - 1\right].$$

The $n/\gamma(n)$ terms vanish as $n \to \infty$. The last term satisfies

$$\frac{1}{n}\gamma(n)\sum_{j=1}^J r_{ij}(y)\left[e^{n(\xi_j-\xi_i)/\gamma(n)} - 1\right] = \sum_{j=1}^J r_{ij}(y)\left[\xi_j(y) - \xi_i(y)\right] + o_{n\to\infty}(1).$$

Therefore, we choose again $E' := \mathbb{T}^d \times \{1,\dots,J\}$ as the state space of the macroscopic variables, and use the following limit operator H,

$$H := \left\{(f, H_{f,\varphi,\xi}) \,:\, f \in C^2(\mathbb{T}^d) \text{ and } H_{f,\varphi,\xi} \in C(\mathbb{T}^d \times E')\right\}, \tag{7.5}$$

with functions $\varphi \in C^2(\mathbb{T}^d)$ and $\xi = (\xi_1,\dots,\xi_j) \in C^2(E') \simeq (C^2(\mathbb{T}^d))^J$. In the notation of (C3), with $k=2$, the set $\mathcal{C} \subset C(E';\mathbb{R}^k)$ is

$$\mathcal{C} = \left\{\alpha \in C(E';\mathbb{R}^2) \,:\, \alpha(x,i) = (\varphi(x), \xi(x,i)),\, \varphi \in C^2(\mathbb{T}^d),\, \xi \in C^2(E')\right\}.$$

The image functions $H_{f,\varphi,\xi} : \mathbb{T}^d \times E' \to \mathbb{R}$ are

$$H_{f,\varphi,\xi}(x,y,i) := \frac{1}{2}\Delta_y\varphi(y) + \frac{1}{2}|\nabla f(x) + \nabla_y\varphi(y)|^2 + b_i(y)\left(\nabla f(x) + \nabla_y \varphi(y)\right) + \sum_{j=1}^J r_{ij}(y)\left[\xi(y,j) - \xi(y,i)\right]. \tag{7.6}$$

Then H satisfies (T1), which is shown by the same line of argument as above in the proof of Theorem 4.6, with the same maps η_n and η_n'. The image functions depend only on gradients, $H_{f,\varphi,\xi}(x,y,i) = H_{\varphi,\xi}(\nabla f(x), y, i)$. □

Verification of (T2) of Theorem 4.2. For any $p \in \mathbb{R}^d$, we establish the existence of functions $\varphi_p \in C^2(\mathbb{T}^d)$ and $\xi \in C^2(E')$ such that $H_{\varphi,\xi}(p,\cdot)$ becomes constant on $E' = \mathbb{T}^d \times \{1,\dots,J\}$. To that end, we find a constant $\mathcal{H}(p) \in \mathbb{R}$ and φ_p and ξ_p such that for all $(y,i) \in E'$, we have

$$H_{\varphi_p,\xi_p}(p,y,i) = \mathcal{H}(p).$$

We reduce the problem to finding a principal eigenvalue.

Lemma 7.2. *Let $E' = \mathbb{T}^d \times \{1,\dots,J\}$ and let H be the operator* (7.5). *Then:*

(a) For $f \in \mathcal{D}(H)$, the images $H_{\varphi,\xi}$ are given by

$$H_{\varphi,\xi}(p,y,i) = e^{-\varphi(y)}\left[(B_{p,i} + V_{p,i})e^{\varphi}\right](y) + \sum_{j=1}^J r_{ij}(y)\left[\xi(y,j) - \xi(y,i)\right],$$

where $p = \nabla f(x) \in \mathbb{R}^d$, $B_{p,i} = \frac{1}{2}\Delta_y + (p + b_i(y))\cdot\nabla_y$, and $V_{p,i}(y) = p^2/2 + p\cdot b_i(y)$ is a multiplication operator.

(b) *For any* φ *and* $y \in \mathbb{T}^d$*, there exists a function* $\xi(y,\cdot)$ *on* $\{1,\dots,J\}$ *such that* $\xi \in C^2(E')$ *and for all* $i = 1,\dots,J$*,*

$$e^{-\varphi}\left[(B_{p,i}+V_{p,i})e^{\varphi}\right](y) + \sum_{j=1}^{J} r_{ij}(y)\left[\xi(y,j)-\xi(y,i)\right] = e^{-\varphi(y)}\left[B_p+V_p\right]e^{\varphi(y)},$$

where $B_p = \frac{1}{2}\Delta_y + (p+\bar{b}(y))\cdot\nabla_y$*,* $V_p(y) = \frac{p^2}{2} + p\cdot\bar{b}(y)$*. In the operators,* $\bar{b}(y) := \sum_{i=1}^{J}\mu_y(i)b_i(y)$ *is the average drift with respect to the stationary measure* $\mu_y \in \mathcal{P}(\{1,\dots,J\})$ *of the jump process with frozen jump rates* $r_{ij}(y)$*.*

(c) *There exists a strictly positive eigenfunction* g_p *and an eigenvalue* $\mathcal{H}(p) \in \mathbb{R}$ *such that*

$$\left[B_p+V_p\right]g_p = \mathcal{H}(p)g_p. \tag{7.7}$$

By (a), (b) and (c), taking $\varphi_p = \log g_p$ and the corresponding $\xi(y,i)$, we obtain (T2) via

$$\begin{aligned} H_{\varphi_p,\xi}(p,y,i) &\overset{(a)}{=} e^{-\varphi_p(y)}\left[B_{p,i}+V_{p,i}\right]e^{\varphi_p(y)} + \sum_{j\in\mathcal{J}} r_{ij}(y)\left[\xi(y,j)-\xi(y,i)\right] \\ &\overset{(b)}{=} e^{-\varphi_p(y)}\left[(B_p+V_p)e^{\varphi}\right](y) \overset{(c)}{=} \mathcal{H}(p). \end{aligned}$$

Proof of Lemma 7.2. Regarding (a), writing $\xi(y,i) = \xi_y(i)$ and $p = \nabla f(x) \in \mathbb{R}^d$, for all $(y,i) \in E'$ we find

$$H_{\varphi,\xi}(p,y,i) = \underbrace{\frac{1}{2}\Delta_y\varphi + \frac{1}{2}\left|p+\nabla_y\varphi\right|^2 + b_i(p+\nabla_y\varphi)}_{= e^{-\varphi}(B_{p,i}+V_{p,i})e^{\varphi}} + \underbrace{\sum_{j=1}^{J} r_{ij}(y)\left[\xi(y,j)-\xi(y,i)\right]}_{=: R_y\xi(y,\cdot)(i)},$$

with a generator R_y of a jump process with frozen jump rates $r_{ij}(y)$.

For (b), let $\varphi \in C^2(\mathbb{T}^d)$ and $y \in \mathbb{T}^d$. We wish to find a function $\xi_y(\cdot) = \xi(y,\cdot) \in C(\{1,\dots,J\})$ such that

$$e^{-\varphi}\left[B_{p,i}+V_{p,i}\right]e^{\varphi} + R_y\xi_y(i)$$

becomes constant in $i = 1,\dots,J$. By the Fredholm alternative, for any vector $v \in C(\{1,\dots,J\})$, the equation $R_y\xi_y = v$ has a solution $\xi_y(\cdot) \in C(\{1,\dots,J\})$ if and only if $v \perp \ker(R_y^*)$. Since R_y is the generator of a jump process on the finite discrete set $\{1,\dots,J\}$ with rates $r_{ij}(y)$, the null space $\ker(R_y^*)$ is one-dimensional and spanned by the unique stationary measure $\mu_y \in \mathcal{P}(\{1,\dots,J\})$, which exists by our irreducibility assumption of Theorem 4.7 (e.g. [22, Theorem 17.51]). Hence $e^{-\varphi}\left[B_{p,i}+V_{p,i}\right]e^{\varphi} + R_y\xi_y(i) = h(p,y)$ is independent of $i \in \{1,\dots,J\}$ if and only if

$$\sum_{i=1}^{J}\mu_y(i)\left[(h(p,y) - e^{-\varphi}\left[B_{p,i}+V_{p,i}\right]e^{\varphi}\right] = 0.$$

This solvability condition leads to

$$\sum_{i=1}^{J}\mu_y(i)\left[(h(p,y) - e^{-\varphi}\left(B_{p,i}+V_{p,i}\right)e^{\varphi}\right] = h(p,y) - e^{-\varphi(y)}\left(B_p+V_p\right)e^{\varphi(y)} = 0.$$

Hence for $h(p,y) := e^{-\varphi(y)}\left[B_p+V_p\right]e^{\varphi(y)}$, there exists $\xi(y,i)$ solving the equation $R_y\xi(y,\cdot) = h$. Furthermore, since the stationary measure is an eigenvector of a one-dimensional eigenspace, and the rates $r_{ij}(\cdot)$ are smooth by assumption, the eigenfunctions ξ_y depend smoothly on y as well, and (b) follows.

For proving (c) in Lemma 7.2, we note that Equation (7.7) corresponds to a principal-eigenvalue problem for a second-order uniformly elliptic operator. By Proposition B.1, the principal eigenvalue problem $[-B_p - V_p]\, g_p = \lambda(p) g_p$ has a solution $g_p > 0$, with eigenvalue $\lambda(p) \in \mathbb{R}$. The same function g_p and the eigenvalue $\mathcal{H}(p) = -\lambda(p)$ solve (7.7). □

Verification of (T3) of Theorem 4.3. The principal eigenvalue $\mathcal{H}(p)$ is of the form

$$\mathcal{H}(p) = \inf_{\varphi} \sup_{y \in \mathbb{T}^d} F\left(p, \varphi\right)(y),$$

with F jointly convex in p and φ. Convexity of $\mathcal{H}(p)$ and $\mathcal{H}(0) = 0$ follow as above in the proof of Theorem 4.6. □

8 Proof of symmetry of Hamiltonians

Theorem 4.8 shows that detailed balance implies symmetric Hamiltonians. The proof was based on a suitable variational representation of the Hamiltonian. In this section, we show in Proposition 8.1 how to obtain this representation.

Before giving the rigorous proof, we sketch the argument. To that end, we recall the setting. We work with $E' = \mathbb{T}^d \times \{1, \dots, J\}$ and denote by $\mathcal{P}(E')$ the set of probability measures on E'. The Hamiltonian $\mathcal{H}(p)$ is the principal eigenvalue of the cell problem (7.4) described in Lemma 7.1, and satisfies

$$\mathcal{H}(p) = \sup_{\mu \in \mathcal{P}(E')} \left[\int_{E'} V_p(z)\, \mathrm{d}\mu(z) - I_p(\mu) \right]. \tag{8.1}$$

In this formula, we have the continuous map V_p given by

$$V_p(x,i) := \frac{1}{2} p^2 - p \cdot \nabla \psi_i(x), \tag{8.2}$$

and the Donsker-Varadhan functional

$$I_p(\mu) := - \inf_{u > 0} \int_{E'} \frac{L_p u}{u}\, \mathrm{d}\mu, \tag{8.3}$$

where the infimum is over strictly positive $u \in C^2(E')$ and the operator L_p is

$$L_p u(x,i) := \frac{1}{2} \Delta_x u(x,i) + (p - \nabla \psi_i(x)) \cdot \nabla_x u(x,i) + \sum_{j=1}^{J} r_{ij}(x) \left[u(x,j) - u(x,i) \right]. \tag{8.4}$$

The variational representation (8.1) is a special case of Donsker's and Varadhan's representation theorem on principal eigenvalues [6]. Under their general conditions, the infimum is taken over functions that are in the domain of the infinitesimal generator of the semigroup generated by L_p. Pinsky showed that the infimum can be taken over C^2 functions if the coefficients appearing in the operator L_p are sufficiently regular (Theorem 1.4 in [33], Equation (3.1) in [34]).

Since it is not clear from (8.1) that $\mathcal{H}(p)$ is symmetric under the detailed-balance condition, we shall perform a suitable shift in the infimum of the functional (8.3) to obtain a suitable representation. Rewriting in (8.3) the strictly positive functions as $u = \exp(\varphi)$ (with $\varphi = \log u$, so φ has the same regularity as u), we find

$$I_p(\mu) = - \inf_{\varphi \in C^2} \sum_i \int \left(\frac{1}{2} \Delta \varphi_i + \frac{1}{2} |\nabla \varphi_i|^2 + (p - \nabla \psi_i) \nabla \varphi_i + \sum_j r_{ij} \left(e^{\varphi_j - \varphi_i} - 1 \right) \right) \mathrm{d}\mu_i. \tag{8.5}$$

Suppose that $\mathrm{d}\mu_i = \overline{\mu}_i\,\mathrm{d}x$ with a function $\overline{\mu}_i$ that is bounded away from zero, where $\mathrm{d}x$ is the Lebesgue measure on the torus. Then shifting in the infimum by setting $\varphi_i = \phi_i + \psi_i + \frac{1}{2}\log\overline{\mu}_i$, we find by calculation that

$$I_p(\mu) = \mathcal{R}(\mu) + \int_{E'} V_p\,\mathrm{d}\mu - K_p(\mu), \tag{8.6}$$

where $\mathcal{R}(\mu)$ is the Fisher information given by

$$\mathcal{R}(\mu) := \frac{1}{8}\sum_i \int_{\mathbb{T}^d} \left|\nabla\left(\log\frac{\overline{\mu}_i}{e^{-2\psi_i}}\right)\right|^2 \mathrm{d}\mu_i, \tag{8.7}$$

and $K_p(\mu)$ is given by

$$\begin{aligned} K_p(\mu) = \inf_{\phi}\Bigg\{ &\sum_{i=1}^J \int_{\mathbb{T}^d}\left(\frac{1}{2}|\nabla\phi_i(x)+p|^2 - \sum_{j=1}^J r_{ij}(x)\right)\mathrm{d}\mu_i(x) \\ &+ \sum_{i,j=1}^J \int_{\mathbb{T}^d} r_{ij}(x)e^{-2\psi_i(x)}\sqrt{\overline{\mu}_i(x)\overline{\mu}_j(x)}e^{\psi_j(x)+\psi_i(x)}e^{\phi(x,j)-\phi(x,i)}\,\mathrm{d}x\Bigg\}. \end{aligned} \tag{8.8}$$

Plugging formula (8.6) into the variational representation (8.1) leads to the desired representation of the Hamiltonian as used at the beginning of the proof of Theorem 4.8 in Section 4.4. The transformation we used is equivalent to shifting by $(1/2)\log(\overline{\mu}_i/\pi_i)$, where $\pi_i = e^{-2\psi_i}$ is the stationary measure up to multiplicative constant. This transformation is reminiscent of a *symmetrization* discussed in Touchette's notes [39, Eq. (36)]. Also when formulating the detailed-balance condition with additional constants in (4.3), that is when not shifting the potentials by constants to renormalize, one can include these constants in the shift to arrive at the same conclusions.

In order to make the strategy as outlined above rigorous, we prove that we can restrict to measures μ having the required regularity properties. The central step is to exploit the fact that $I_p(\mu)$ is finite since $\mathcal{H}(p)$ is finite. By a result of Stroock [37, Theorem 7.44], finiteness of the Donsker-Varadhan functional implies certain regularity properties in case the generator is reversible. Since the generator L_p is not reversible, we further bound I_p by a suitable Donsker-Varadhan functional I_{rev} corresponding to a reversible process in order to be able to apply [37, Theorem 7.44].

Proposition 8.1. *The Hamiltonian $\mathcal{H}(p)$ given by* (8.1) *satisfies the following:*

(a) The supremum in (8.1) *can be taken over a smaller set $\mathbf{P}$ of measures, that is*

$$\mathcal{H}(p) = \sup_{\mu\in\mathbf{P}}\left[\int_{E'} V_p\,\mathrm{d}\mu - I_p(\mu)\right],$$

where $\mathbf{P}\subset\mathcal{P}(E')$ are the probability measures $\mu = (\mu_1,\dots,\mu_J)$ such that:

(P1) Each μ_i is absolutely continuous with respect to the uniform measure on $\mathbb{T}^d$.

(P2) For each i, we have $\nabla(\log\overline{\mu}_i)\in L^2_{\mu_i}(\mathbb{T}^d)$, where $\mathrm{d}\mu_i(x) = \overline{\mu}_i(x)\mathrm{d}x$.

(b) We have

$$\mathcal{H}(p) = \sup_{\mu\in\mathbf{P}}\left[K_p(\mu) - \mathcal{R}(\mu)\right], \tag{8.9}$$

with the maps $\mathcal{R}$ and K_p given by (8.7) *and* (8.8) *above. In $K_p(\mu)$, the infimum should be taken over vectors of functions $\phi_i = \phi(\cdot,i)$ in $C^2(\mathbb{T}^d)$.*

(c) Under the detailed balance condition (4.3)*, for any* $\mu \in \mathbf{P}$*,*

$$K_p(\mu) = \inf_{\phi} \Bigg\{ \sum_{i=1}^J \int_{\mathbb{T}^d} \left(\frac{1}{2}|\nabla\phi_i(x) + p|^2 - \sum_{j=1}^J r_{ij}(x) \right) \mathrm{d}\mu_i(x) \\ + \sum_{i,j=1}^J \int_{\mathbb{T}^d} r_{ij}(x) e^{-2\psi_i(x)} \sqrt{\overline{\mu}_i(x)\overline{\mu}_j(x)} e^{\psi_j(x)+\psi_i(x)} \cosh\left(\phi(x,j) - \phi(x,i)\right) \mathrm{d}x \Bigg\}. \tag{8.10}$$

The representation (8.10) follows from (8.8) by rewriting the sums appearing therein as $\sum_{ij} a_{ij} = \frac{1}{2}\sum_{ij}(a_{ij} + a_{ji})$, where

$$a_{ij} = \int_{\mathbb{T}^d} r_{ij} e^{-2\psi_i} \sqrt{\overline{\mu}^i(x)\overline{\mu}^j(x)} e^{\psi^j(x)+\psi_i(x)} e^{\phi(x,j)-\phi(x,i)} \, \mathrm{d}x.$$

This leads to the $\cosh(\cdot)$ terms in (8.10), and proves (c). We now give the proof of (a) and (b) of Proposition 8.1.

Proof of (a) in Proposition 8.1. Let $p \in \mathbb{R}^d$. The supremum in (8.1) can be taken over measures μ such that $I_p(\mu)$ is finite, because $\mathcal{H}(p)$ is finite and $V_p(\cdot)$ is bounded. We show that finiteness of $I_p(\mu)$ implies that μ must satisfy (P1) and (P2). To that end, define the map $L_{\mathrm{rev}} : \mathcal{D}(L_{\mathrm{rev}}) \subseteq C(E') \to C(E')$ by setting $\mathcal{D}(L_{\mathrm{rev}}) := C^2(E')$ and

$$L_{\mathrm{rev}} f(x,i) := \frac{1}{2}\Delta_x f(x,i) - \nabla\psi_i(x) \cdot \nabla_x f(x,i) + \overline{\gamma} \sum_{j \neq i} s_{ij}(x) \left[f(x,j) - f(x,i) \right],$$

with jump rates s_{ij} defined as $s_{ij} \equiv 1$ and $s_{ji} \equiv e^{2\psi_j - 2\psi_i}$, for $i \leq j$, and with $\overline{\gamma} := \sup_{\mathbb{T}^d}(r_{ij}/s_{ij}) < \infty$, where $r_{ij}(\cdot)$ are the jump rates appearing in L_p. Furthermore, define $I_{L_{\mathrm{rev}}} : \mathcal{P}(E') \to [0,\infty]$ by

$$I_{L_{\mathrm{rev}}}(\mu) := - \inf_{\varphi \in C^2(E')} \int_{E'} e^{-\varphi} L_{\mathrm{rev}}(e^{\varphi}) \, \mathrm{d}\mu. \tag{8.11}$$

We shall prove two statements:

(I) If $I_{L_{\mathrm{rev}}}(\mu)$ is finite, then the measure μ satisfies (P1) and (P2).

(II) If $I_p(\mu)$ is finite, then $I_{L_{\mathrm{rev}}}(\mu)$ is finite.

The two statements combined finish the proof.

Regarding (I), suppose $I_{L_{\mathrm{rev}}}(\mu)$ is finite. Since $s_{ij}e^{-2\psi_i} = s_{ji}e^{-2\psi_j}$, the operator L_{rev} admits a reversible measure ν_{rev} in $\mathcal{P}(E')$ given by

$$\nu_{\mathrm{rev}}(A_1, \dots, A_J) = \frac{1}{\mathcal{Z}} \sum_{i=1}^J \nu^i_{\mathrm{rev}}(A_i), \quad \text{where } \mathrm{d}\nu^i_{\mathrm{rev}} = e^{-2\psi_i}\mathrm{d}x \text{ and } \mathcal{Z} = \sum_i \nu^i_{\mathrm{rev}}(\mathbb{T}^d).$$

The measure ν_{rev} is reversible for L_{rev} in the sense that for all $f, g \in \mathcal{D}(L_{\mathrm{rev}})$,

$$\langle L_{\mathrm{rev}} f, g \rangle_{\nu_{\mathrm{rev}}} = \langle f, L_{\mathrm{rev}} g \rangle_{\nu_{\mathrm{rev}}}, \quad \text{where } \langle f, h \rangle_{\nu_{\mathrm{rev}}} = \frac{1}{\mathcal{Z}} \sum_i \int_{\mathbb{T}^d} f_i(x) h_i(x) \, \mathrm{d}\nu^i_{\mathrm{rev}}(x).$$

By Stroock's result [37, Theorem 7.44],

$$I_{L_{\mathrm{rev}}}(\mu) = \begin{cases} -\langle f_\mu, L_{\mathrm{rev}} f_\mu \rangle_{\nu_{\mathrm{rev}}}, & f_\mu = \sqrt{g_\mu} \in D^{1/2} := \mathcal{D}\left(\sqrt{-L_{\mathrm{rev}}}\right) \text{ and } g_\mu = \frac{\mathrm{d}\mu}{\mathrm{d}\nu_{\mathrm{rev}}}, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\mathrm{d}\mu/\mathrm{d}\nu_{\text{rev}}$ is the Radon-Nikodym derivative. In particular, since $I_{L_{\text{rev}}}(\mu)$ is finite, we find that $\mu \ll \nu_{\text{rev}}$ and that $I_{L_{\text{rev}}}(\mu)$ is explicitly given by

$$\begin{aligned} I_{L_{\text{rev}}}(\mu) &= -\langle f, L_{\text{rev}} f\rangle_{\nu_{\text{rev}}} \\ &= \frac{1}{\mathcal{Z}}\sum_{i=1}^{J}\left[\int_{\mathbb{T}^d}|\nabla f_i(x)|^2\, d\nu^i_{\text{rev}}(x) + \overline{\gamma}\sum_{j=1}^{J}\int_{\mathbb{T}^d} s_{ij}(x)|f_j(x)-f_i(x)|^2\, d\nu^i_{\text{rev}}(x)\right], \end{aligned} \tag{8.12}$$

where we write $f_i = (\mathrm{d}\mu^i/\mathrm{d}\nu^i_{\text{rev}})^{1/2}$. Furthermore, μ^i is absolutely continuous with respect to $\nu^i = e^{-2\psi_i}\mathrm{d}x$. Since $e^{-2\psi_i}\mathrm{d}x \ll \mathrm{d}x$, we find that μ^i is absolutely continuous with respect to the volume measure on $\mathbb{T}^d$. Hence (P1) holds true.

We verify (P2) by showing that the integral $\int_{\mathbb{T}^d}|\nabla(\log\overline{\mu}^i)|^2\, d\mu^i$ is finite. Let $g^i_\mu := \mathrm{d}\mu^i/\mathrm{d}\nu^i_{\text{rev}}$ be the density of μ^i with respect to ν^i_{rev}. Then the densities $\overline{\mu}^i = \mathrm{d}\mu^i/\mathrm{d}x$ satisfy $g^i_\mu = \overline{\mu}^i e^{2\psi_i}$, because

$$\overline{\mu}^i = \frac{\mathrm{d}\mu^i}{\mathrm{d}\nu^i_{\text{rev}}}\frac{\mathrm{d}\nu^i_{\text{rev}}}{\mathrm{d}x} = \frac{\mathrm{d}\mu^i}{\mathrm{d}\nu^i_{\text{rev}}}e^{-2\psi_i}.$$

Let $f^i_\mu := \sqrt{g^i_\mu}$. By (8.12), $\int_{\mathbb{T}^d}|\nabla f^i_\mu|^2\mathrm{d}\nu^i_{\text{rev}}$ is finite for every $i = 1,\dots,J$. Hence with the estimate

$$\begin{aligned} \int_{\mathbb{T}^d}|\nabla f^i_\mu|^2\mathrm{d}\nu^i_{\text{rev}} &\geq \int_{\mathbb{T}^d}|\nabla f^i_\mu|^2\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}\nu^i_{\text{rev}} = \frac{1}{4}\int_{\mathbb{T}^d}\frac{|\nabla g^i_\mu|^2}{g^i_\mu}\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}\nu^i_{\text{rev}} \\ &= \frac{1}{4}\int_{\mathbb{T}^d}\frac{|e^{2\psi_i}\nabla\overline{\mu}^i + 2\overline{\mu}^i\nabla\psi_i e^{2\psi_i}|^2}{\overline{\mu}^i}e^{-4\psi_i}\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}x \\ &= \frac{1}{4}\int_{\mathbb{T}^d}|\nabla(\log\overline{\mu}^i) + 2\nabla\psi_i|^2\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}\mu^i \\ &\geq \frac{1}{8}\int_{\mathbb{T}^d}|\nabla(\log\overline{\mu}^i)|^2\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}\mu^i - \int_{\mathbb{T}^d}|\nabla\psi_i|^2\mathbf{1}_{\{\overline{\mu}^i>0\}}\mathrm{d}\mu^i, \end{aligned}$$

we find $\nabla(\log\overline{\mu}^i) \in L^2_{\mu^i}(\mathbb{T}^d)$.

Regarding (II), we start from $I_p(\mu)$ as given in (8.5). Suppose that $I_p(\mu)$ is finite. We estimate r_{ij}/s_{ij} from above by $\overline{\gamma} = \sup_{\mathbb{T}^d}(r_{ij}/s_{ij})$ to find

$$\begin{aligned} I_p(\mu) \geq \sup_{\varphi}\sum_i\int_{\mathbb{T}^d} -\Bigg[\frac{1}{2}\Delta\varphi_i(x) + \frac{1}{2}|\nabla\varphi_i(x)|^2 + (p - \nabla\psi_i(x))\nabla\varphi_i(x) \\ + \overline{\gamma}\sum_{j\neq i} s_{ij}(x)(e^{\varphi(x,j)-\varphi(x,i)} - 1)\Bigg]\,\mathrm{d}\mu^i - s_0(\mu), \end{aligned}$$

where $s_0(\mu) = \sum_{ij}\int_{\mathbb{T}^d}[\overline{\gamma}\, s_{ij}(x) - r_{ij}(x)]\,\mathrm{d}\mu^i$ is finite.

For $p = 0$, this means that $I_0(\mu) \geq I_{L_{\text{rev}}}(\mu) - s_0(\mu)$, which follows from writing out the definition of $I_{L_{\text{rev}}}(\mu)$ given in (8.11). Hence, $I_{L_{\text{rev}}}(\mu)$ is finite.

For $p \neq 0$, the additional p-term can be dealt with by Young's inequality applied as

$-p \cdot \nabla \phi^i \geq -p^2/(2\varepsilon) - \frac{\varepsilon}{2}|\nabla \phi^i|^2$, with $\varepsilon > 0$. Thereby,

$$\begin{aligned} I_p(\mu) &\geq \sup_\varphi \sum_i \int_{\mathbb{T}^d} -\left[\frac{1}{2}\Delta\varphi_i(x) + \frac{1+\varepsilon}{2}|\nabla\varphi_i(x)|^2 - \nabla\psi_i(x)\nabla\varphi_i(x)\right. \\ &\qquad \left. + \overline{\gamma}\sum_{j\neq i} s_{ij}(x)(e^{\varphi(x,j)-\varphi(x,i)} - 1)\right] \mathrm{d}\mu^i - \frac{p^2}{2\varepsilon} - s_0(\mu) \\ &= \frac{1}{\lambda}\sup_\varphi \sum_i \int_{\mathbb{T}^d} -\left[\frac{1}{2}\Delta\varphi_i(x) + \frac{1}{2}|\nabla\varphi_i(x)|^2 + -\nabla\psi_i(x)\nabla\varphi_i(x)\right. \\ &\qquad \left. + \lambda\overline{\gamma}\sum_{j\neq i} s_{ij}(x)(e^{(\varphi(x,j)-\varphi(x,i))/\lambda} - 1)\right] \mathrm{d}\mu^i - \frac{p^2}{2\varepsilon} - s_0(\mu), \end{aligned}$$

where the last equality follows by rescaling $\varphi \to \varphi/\lambda$, with $\lambda = 1 + \varepsilon > 1$.

Therefore, apart from the factor $1/\lambda$ in the exponential term and the multiplicative factor $\lambda\overline{\gamma}$, we obtain the same estimate as above in the $p = 0$ case. Denoting the supremum term in the last line by $I^\lambda_{L_{\text{rev}}}$, we found the estimate

$$I_p(\mu) \geq \frac{1}{\lambda} I^\lambda_{L_{\text{rev}}}(\mu) - s_p(\mu), \tag{8.13}$$

where $s_p(\mu) = (2\varepsilon)^{-1}p^2 + s_0(\mu)$. Hence $I^\lambda_{L_{\text{rev}}}(\mu)$ is finite.

We show that this enforces finiteness of $I_{L_{\text{rev}}}(\mu)$, by proving that $I_{L_{\text{rev}}}(\mu) = \infty$ implies $I^\lambda_{L_{\text{rev}}}(\mu) = \infty$.

So, suppose that $I_{L_{\text{rev}}}(\mu) = \infty$. Then by definition of $I_{L_{\text{rev}}}$ in (8.11), there exist functions φ_n such that

$$a(\varphi_n) := -\sum_{i=1}^J \int_{\mathbb{T}^d} \left[\frac{1}{2}\Delta\varphi_n^i + \frac{1}{2}|\nabla\varphi_n^i|^2 - \nabla\psi_i\nabla\varphi_n^i + \overline{\gamma}\sum_{j\neq i} s_{ij}\left(e^{\varphi_n(x,j)-\varphi_n(x,i)} - 1\right)\right] \mathrm{d}\mu^i(x)$$

diverges, that is $a(\varphi_n) \to \infty$ as $n \to \infty$. Write

$$\begin{aligned} a^\lambda(\varphi_n) := -\sum_i \int_{\mathbb{T}^d} &\left[\frac{1}{2}\Delta\varphi_n^i + \frac{1}{2}|\nabla\varphi_n^i|^2 - \nabla\psi_i\nabla\varphi_n^i\right. \\ &\qquad \left. + \lambda\overline{\gamma}\sum_{j\neq i} s_{ij}(e^{(\varphi_n(x,j)-\varphi_n(x,i))/\lambda} - 1)\right] \mathrm{d}\mu^i(x) \end{aligned}$$

for the according evaluation of φ_n in $I^\lambda_{L_{\text{rev}}}(\mu)$. By definition, $I^\lambda_{L_{\text{rev}}}(\mu) \geq a^\lambda(\varphi_n)$. We shall prove that with a finite constant $C = C(\mu)$ and sequences $\overline{a}_n$ and $\overline{a}^\lambda_n$, we have the estimates

$$a(\varphi_n) \overset{(1)}{\leq} \overline{a}_n \overset{(2)}{\leq} \overline{a}^\lambda_n \overset{(3)}{\leq} a^\lambda(\varphi_n) + C.$$

That finishes the proof, since then

$$I^\lambda_{L_{\text{rev}}}(\mu) \geq a^\lambda(\varphi_n) \geq [a(\varphi_n) - C] \xrightarrow{n\to\infty} +\infty.$$

Define the sequences $\overline{a}_n$ and $\overline{a}^\lambda_n$ by

$$\begin{aligned} \overline{a}_n := -\sum_{i=1}^J \int_{\mathbb{T}^d} &\left[\frac{1}{2}\Delta\varphi_n^i + \frac{1}{2}|\nabla\varphi_n^i|^2 - \nabla\psi_i\nabla\varphi_n^i\right. \\ &\qquad \left. + \overline{\gamma}\sum_{j\neq i} s_{ij}\left(e^{\varphi_n(x,j)-\varphi_n(x,i)}\mathbf{1}_{\{\varphi_n(x,j)-\varphi_n(x,i)\geq 0\}} - 1\right)\right] \mathrm{d}\mu^i(x), \end{aligned}$$

and

$$\overline{a}_n^\lambda := -\sum_i \int_{\mathbb{T}^d} \left[\frac{1}{2}\Delta\varphi_n^i + \frac{1}{2}|\nabla\varphi_n^i|^2 - \nabla\psi_i\nabla\varphi_n^i \right. \\ \left. + \lambda\overline{\gamma}\sum_{j\neq i} s_{ij}\left(e^{(\varphi_n(x,j)-\varphi_n(x,i))/\lambda}\mathbf{1}_{\{\varphi_n(x,j)-\varphi_n(x,i)\geq 0\}} - 1\right)\right] \mathrm{d}\mu^i(x).$$

Define the constant C by

$$C := \lambda\overline{\gamma}\sum_{ij;j\neq i}\int_E s_{ij}(x)\,\mathrm{d}\mu^i(x),$$

Regarding inequality (1),

$$\overline{a}_n - a(\varphi_n) = \overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}e^{\varphi_n(x,j)-\varphi_n(x,i)}\left[1-\mathbf{1}_{\{\varphi_n(x,j)-\varphi_n(x,i)\}}\right]\,\mathrm{d}\mu^i(x) \geq 0.$$

Regarding inequality (2), writing $\delta\varphi_n(x,j,i) := \varphi_n(x,j)-\varphi_n(x,i)$, we find

$$\begin{aligned}\overline{a}_n^\lambda - \overline{a}_n &= \overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}\left[e^{\delta\varphi_n(x,j,i)}\mathbf{1}_{\delta\varphi_n(x,j,i)\geq 0\}} - 1\right]\,\mathrm{d}\mu^i(x) \\ &\quad - \lambda\overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}\left[e^{\delta\varphi_n(x,j,i)/\lambda}\mathbf{1}_{\{\delta\varphi_n(x,j,i)\geq 0\}} - 1\right]\,\mathrm{d}\mu^i(x) \\ &= \overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}\left[e^{\delta\varphi_n(x,j,i)} - \lambda e^{\delta\varphi_n(x,j,i)/\lambda}\right]\mathbf{1}_{\{\delta\varphi_n(x,j,i)\geq 0\}}\,\mathrm{d}\mu^i(x) \\ &\quad + \overline{\gamma}(\lambda-1)\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}\,\mathrm{d}\mu^i(x) \\ &\geq 0,\end{aligned}$$

using in the last estimate $\lambda = 1+\varepsilon > 1$ and $e^z - \lambda e^{z/\lambda} \geq (1-\lambda)$ for $z \geq 0$.

Regarding inequality (3), using $e^{z/\lambda}(\mathbf{1}_{z\geq 0} - 1) \geq -1$ for $z \in \mathbb{R}$, we find

$$\begin{aligned}a^\lambda(\varphi_n) - \overline{a}_n^\lambda &= \lambda\overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}e^{(\varphi_n(x,j)-\varphi_n(x,i))/\lambda}\left[\mathbf{1}_{\{\varphi_n(x,j)-\varphi_n(x,i)\}} - 1\right]\,\mathrm{d}\mu^i \\ &\geq -\lambda\overline{\gamma}\sum_{i,j;j\neq i}\int_{\mathbb{T}^d} s_{ij}\,\mathrm{d}\mu^i = -C.\end{aligned}$$

□

Proof of (b) of Proposition 8.1. It is sufficient to show that for any $\mu \in \mathbf{P}$ the Donsker-Varadhan functional $I_p(\mu)$ satisfies (8.6). Integrating by parts in (8.5) gives

$$I_p(\mu) = -\inf_{\varphi\in C^2}\sum_i\int_{\mathbb{T}^d}\left[-\frac{1}{2}\nabla\varphi_i\nabla(\log\overline{\mu}^i) + \frac{1}{2}|\nabla\varphi_i|^2 + (p-\nabla\psi_i)\nabla\varphi_i \right. \\ \left. + \sum_j r_{ij}\left(e^{\varphi_j-\varphi_i} - 1\right)\right]\mathrm{d}\mu^i,$$

where we write $\mathrm{d}\mu^i = \overline{\mu}^i\mathrm{d}x$. Now shifting in the infimum as $\varphi_i = \phi_i + \frac{1}{2}\log(\overline{\mu}^i) + \psi_i$, we find after some algebra that

$$I_p(\mu) = - \inf_{\phi \in -\frac{1}{2}\log \overline{\mu} + C^2} \sum_i \int_{\mathbb{T}^d} \left[\frac{1}{2}|\nabla \phi_i + p|^2 - \frac{1}{2}\left|(p - \nabla \psi_i) - \frac{1}{2}\nabla \log \overline{\mu}^i\right|^2 \right.$$
$$\left. + \sum_j r_{ij} \left(\sqrt{\frac{\overline{\mu}^j}{\overline{\mu}^i}} e^{\psi_j - \psi_i} e^{\phi_j - \phi_i} - 1 \right) \right] \mathrm{d}\mu^i.$$

The term containing the square roots is not singular since it is integrated against $\mathrm{d}\mu^i$.

By using the usual methods of approximation by truncation in Sobolev spaces (see e.g. [1, Prop. 9.5]) the infimum above can be replaced by an infimum over all $\phi \in -\frac{1}{2}\log \overline{\mu} + C^2$ that are bounded, and by subsequent regularization by convolution the infimum can be taken over C^2.

Now writing out the terms and reorganizing them leads to the claimed equality. □

A Large-deviation principle implies almost-sure convergence

It is a well-known fact that a large-deviation principle implies a strong type of convergence of random variables [5, at the end of Section I.2]. We provide a proof here since we know of no reference in the literature providing an explicit proof of the well-known fact. Let $\mathcal{I} : \mathcal{X} \to [0, \infty]$ be a rate function. We denote by $\{\mathcal{I} = 0\}$ the set of its global minimizers.

Theorem A.1. *For $n = 1, 2 \ldots$, let X^n be a random variable taking values in a Polish space $(\mathcal{X}, d)$. Suppose that $\{X^n\}_{n\in\mathbb{N}}$ satisfies a large-deviation principle with rate function $\mathcal{I}$. Then $d(X^n, \{\mathcal{I} = 0\}) \to 0$ almost surely as $n \to \infty$.*

We point out that as specified in Definition 3.1, the rate function in Theorem A.1 is assumed to have compact sub-level sets.

Proof of Theorem A.1. For $k, n \in \mathbb{N}$, let A_k^n be the event

$$A_k^n := \{d(X^n, \{\mathcal{I} = 0\}) \geq 1/k\},$$

and write

$$A_k^n \text{ i.o.} := \bigcap_{N\geq 1} \bigcup_{n \geq N} A_k^n.$$

Let $k \in \mathbb{N}$. By the large-deviation upper bound, there exists a $\delta > 0$ such that for all n sufficiently large,

$$\mathbb{P}\left(A_k^n\right) \leq e^{-n\delta}.$$

Therefore $\sum_{n=1}^\infty \mathbb{P}\left(A_k^n\right)$ is finite, and by the Borel-Cantelli Lemma,

$$\mathbb{P}\left(A_k^n \text{ i.o.}\right) = 0.$$

With that, almost-sure convergence follows by noting that

$$\mathbb{P}\left(\{d(X^n, \{\mathcal{I} = 0\}) \xrightarrow{n\to\infty} 0\} \text{ is not true}\right) \leq \sum_{k=1}^\infty \mathbb{P}\left(A_k^n \text{ i. o.}\right) = 0.$$

□

B Principal eigenvalues

In this section we collect results on principal-eigenvalue problems that we encounter in the proofs of the molecular-motor models.

Proposition B.1. *Let P be a second-order uniformly elliptic operator given by*

$$P = -\sum_{k,\ell=1}^{d} a_{k\ell}(\cdot)\frac{\partial^2}{\partial x^k \partial x^\ell} + \sum_{k=1}^{d} b_k(\cdot)\frac{\partial}{\partial x^k} + c(\cdot), \tag{B.1}$$

with smooth coefficients $a_{k\ell}, b_k, c \in C^\infty(\mathbb{T}^d)$. Then there exists a strictly positive function $u \in C^\infty(\mathbb{T}^d)$ and a unique $\lambda \in \mathbb{R}$ such that $Pu = \lambda u$, and λ is given by

$$\lambda = \sup_{g>0} \inf_{x\in\mathbb{T}^d} \left[\frac{Pg(x)}{g(x)}\right] = \inf_{\mu\in\mathcal{P}(\mathbb{T}^d)} \sup_{g>0} \left[\int_{\mathbb{T}^d} \frac{Pg}{g}\, d\mu\right]. \tag{B.2}$$

Here μ is the unique stationary measure.

The existence of the principal eigenvalue λ and the associated positive eigenfunction u for second-order elliptic operators on closed manifolds, such as the torus $\mathbb{T}^d$, is given for instance by Padilla [30, Th. 6.1]; note that we do not assume a sign on c, and therefore we do not obtain a sign on λ either. The characterization (B.2) is given by Donsker and Varadhan [6, 7].

Proposition B.2. *Let $L : C^2(\mathbb{T}^d)^J \to C(\mathbb{T}^d)^J$ be a $J \times J$ diagonal matrix of uniformly elliptic operators,*

$$L = \begin{pmatrix} L^{(1)} & & 0 \\ & \ddots & \\ 0 & & L^{(J)} \end{pmatrix},\ L^{(i)} = -\sum_{k,\ell=1}^{d} a_{k\ell}^{(i)}(\cdot)\frac{\partial^2}{\partial x^k \partial x^\ell} + \sum_{k=1}^{d} b_k^{(i)}(\cdot)\frac{\partial}{\partial x^k} + c^{(i)}(\cdot), \tag{B.3}$$

with $a_{k\ell}^{(i)}(\cdot), b_k^{(i)}(\cdot), c^{(i)}(\cdot) \in C^\infty(\mathbb{T}^d)$, and let $R = R(y)$ be a $J\times J$-matrix-valued function with non-negative off-diagonal elements,

$$R = \begin{pmatrix} R_{11} & & \geq 0 \\ & \ddots & \\ \geq 0 & & R_{JJ} \end{pmatrix}, \qquad R_{ij} \geq 0 \text{ for all } i \neq j.$$

Suppose that the matrix $\overline{R}$ with entries $\overline{R}_{ij} := \sup_{y\in\mathbb{T}^d} R_{ij}(y)$ is irreducible. Then for the operator $P := L - R$, there exists a unique $\lambda \in \mathbb{R}$ and a vector-valued function $u \in \left(C^\infty(\mathbb{T}^d)\right)^J$, $u^i(\cdot) > 0$ for all $i = 1, \dots, J$, such that $Pu = \lambda u$. Furthermore, λ is given by

$$\lambda = \sup_{g>0} \inf_{z\in E'} \left[\frac{Pg(z)}{g(z)}\right] = \inf_{\mu\in\mathcal{P}(E')} \sup_{g>0} \left[\int_{E'} \frac{Pg}{g}\, d\mu\right].$$

(recall that $E' = \mathbb{T}^d \times \{1, \dots, J\}$).

Kifer proves this result in [21, Lemma 2.1 and Proposition 2.2], for the case in which all off-diagonal elements R_{ij} are strictly positive. Sweers [38, Th. 1.1] proves a similar result, relaxing $R_{ij} > 0$ to the combination of $R_{ij} \geq 0$ with irreducibility, but for domains with Dirichlet boundary conditions. The method of Sweers also applies to the case of periodic boundary conditions as for $\mathbb{T}^d$. We omit the details.

References

[1] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011. MR2759829

[2] A. Budhiraja and P. Dupuis, *Analysis and Approximation of Rare Events: Representations and Weak Convergence Methods*, Probability Theory and Stochastic Modelling, vol. 94, Springer, 2019. MR3967100

[3] A. Budhiraja, P. Dupuis, and A. Ganguly, *Large deviations for small noise diffusions in a fast Markovian environment*, Electronic Journal of Probability **23** (2018). MR3878137

[4] K. Deimling, *Multivalued Differential Equations*, De Gruyter Series in Nonlinear Analysis and Applications, vol. 1, Walter de Gruyter & Co., Berlin, 1992. MR1189795

[5] F. den Hollander, *Large Deviations*, Fields Institue Monographs, vol. 14, American Mathematical Soc., 2000. MR1739680

[6] M. D. Donsker and S. R. S. Varadhan, *On a variational formula for the principal eigenvalue for operators with maximum principle*, Proceedings of the National Academy of Sciences **72** (1975), no. 3, 780–783. MR0361998

[7] M. D. Donsker and S. R. S. Varadhan, *On the principal eigenvalue of second-order elliptic differential operators*, Communications on Pure and Applied Mathematics **29** (1976), no. 6, 595–621. MR0425380

[8] R. S. Ellis, *The theory of large deviations: from Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence*, Physica D: Nonlinear Phenomena **133** (1999), no. 1-4, 106–136. MR1721142

[9] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, vol. 282, Wiley, New York, 1986. MR0838085

[10] A. Faggionato, D. Gabrielli, and M. Ribezzi Crivellari, *Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors*, arXiv preprint arXiv:0808.1910 (2008). MR2759771

[11] A. Faggionato, D. Gabrielli, and M. Ribezzi Crivellari, *Non-equilibrium thermodynamics of piecewise deterministic Markov processes*, Journal of Statistical Physics **137** (2009), 259–304. MR2559431

[12] A. Faggionato and V. Silvestri, *Random walks on quasi one dimensional lattices: Large deviations and fluctuation theorems*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 53, 2017, pp. 46–78. MR3606734

[13] J. Feng and T. G. Kurtz, *Large Deviations for Stochastic Processes*, Mathematical surveys and monographs, vol. 131, American Mathematical Society, 2006. MR2260560

[14] M. Freidlin and T.-Y. Lee, *Wave front propagation and large deviations for diffusion–transmutation process*, Probability theory and related fields **106** (1996), no. 1, 39–70. MR1408416

[15] S. Hastings, D. Kinderlehrer, and J. B. McLeod, *Diffusion mediated transport in multiple state systems*, SIAM Journal on Mathematical Analysis **39** (2008), no. 4, 1208–1230. MR2368900

[16] S. Hastings, D. Kinderlehrer, and J. B. McLeod, *Diffusion mediated transport with a look at motor proteins*, Recent advances in nonlinear analysis (M. Chipot, C. S. Lin, and D. H. Tsai, eds.), World Scientific, 2008, pp. 95–112. MR2410740

[17] Q. He and G. Yin, *Large deviations for multi-scale Markovian switching systems with a small diffusion*, Asymptotic Analysis **87** (2014), no. 3-4, 123–145. MR3195726

[18] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton*, Sinauer Associates Sunderland, 2001.

[19] G. Huang, M. Mandjes, and P. Spreij, *Large deviations for Markov-modulated diffusion processes with rapid switching*, Stochastic Processes and their Applications **126** (2016), no. 6, 1785–1818. MR3483737

[20] F. Jülicher, A. Ajdari, and J. Prost, *Modeling Molecular Motors*, Reviews of Modern Physics **69** (1997), no. 4, 1269.

[21] Y. Kifer, *Principal eigenvalues and equilibrium states corresponding to weakly coupled parabolic systems of PDE*, Journal d'Analyse Mathématique **59** (1992), no. 1, 89–102. MR1226953

[22] A. Klenke, *Probability theory: A comprehensive course*, Springer Science & Business Media, 2013. MR3112259

[23] A. B. Kolomeisky, *Motor proteins and molecular motors: How to operate machines at the nanoscale*, Journal of Physics: Condensed Matter **25** (2013), no. 46, 463101.

[24] A. B. Kolomeisky and M. E. Fisher, *Molecular motors: A theorist's perspective*, Annual Review of Physical Chemistry **58** (2007), no. 1, 675–695.

[25] R. C. Kraaij, *Large deviations for finite state Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton-Jacobi equation*, Journal of Statistical Physics **164** (2016), no. 2, 321–345. MR3513255

[26] R. C. Kraaij and M. C. Schlottke, *A large deviation principle for Markovian slow-fast systems*, arXiv preprint arXiv:2011.05686 (2020).

[27] R. Kumar and L. Popovic, *Large deviations for multi-scale jump-diffusion processes*, Stochastic Processes and their Applications **127** (2017), no. 4, 1297–1320. MR3619272

[28] O. A. Ladyzhenskaia, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and Quasi-linear Equations of Parabolic Type*, vol. 23, American Mathematical Soc., 1968. MR0244627

[29] S. Mirrahimi and P. E. Souganidis, *A homogenization approach for the motion of motor proteins*, Nonlinear Differential Equations and Applications NoDEA **20** (2013), no. 1, 129–147. MR3011315

[30] P. Padilla, *The principal eigenvalue and maximum principle for second order elliptic operators on Riemannian manifolds*, Journal of Mathematical Analysis and Applications **205** (1997), no. 2, 285–312. MR1428349

[31] B. Perthame and P. E. Souganidis, *Asymmetric potentials and motor effect: A homogenization approach*, Annales de l'Institut Henri Poincaré (C) Non Linear Analysis **26** (2009), no. 6, 2055–2071. MR2569885

[32] B. Perthame and P. E. Souganidis, *Asymmetric potentials and motor effect: A large deviation approach*, Archive for Rational Mechanics and Analysis **193** (2009), no. 1, 153–169. MR2506073

[33] R. G. Pinsky, *On evaluating the Donsker-Varadhan I-function*, The Annals of Probability (1985), 342–362. MR0781409

[34] R. G. Pinsky, *Regularity properties of the Donsker–Varadhan rate functional for non-reversible diffusions and random evolutions*, Stochastics and Dynamics **7** (2007), no. 02, 123–140. MR2339689

[35] P. Reimann, *Brownian motors: Noisy transport far from equilibrium*, Physics Reports **361** (2002), no. 2-4, 57–265. MR1895137

[36] R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pacific Journal of Mathematics **17** (1966), no. 3, 497–510. MR0193549

[37] D. W. Stroock, *An introduction to the theory of large deviations*, Springer Science & Business Media, 2012. MR0755154

[38] G. Sweers, *Strong positivity in $C(\overline{\Omega})$ for elliptic systems*, Mathematische Zeitschrift **209** (1992), no. 1, 251–271. MR1147817

[39] H. Touchette, *Introduction to dynamical large deviations of Markov processes*, Physica A: Statistical Mechanics and its Applications **504** (2018), 5–19. MR3805496

[40] D. Vorotnikov, *The flashing ratchet and unidirectional transport of matter*, Discrete & Continuous Dynamical Systems-B **16** (2011), no. 3, 963. MR2806331

[41] D. Vorotnikov, *Analytical aspects of the Brownian motor effect in randomly flashing ratchets*, Journal of mathematical biology **68** (2014), no. 7, 1677–1705. MR3201910

[42] H. Wang, C. S. Peskin, and T. C. Elston, *A robust numerical algorithm for studying biomolecular transport processes*, Journal of Theoretical Biology **221** (2003), no. 4, 491–511. MR2043095

[43] G. Yin and C. Zhu, *Hybrid switching diffusions: Properties and applications*, Springer New York, 2010. MR2559912

Acknowledgments. The authors thank Frank Redig, Francesca Collet and Federico Sau for their remarks and suggestions. The authors also thank the reviewers for their very detailed feedback that significantly contributed to improving the quality of the paper. MS also thanks Georg Prokert, Jim Portegies and Richard Kraaij for answering various questions about principal eigenvalues, measure theory and large deviations. The authors acknowledge financial support through NWO grant 613.001.552.