Large-Deviation Principles of switching Markov processes via Hamilton-Jacobi Equations

Mark A. Peletier\textsuperscript{1,2} and Mikola C. Schlottke\textsuperscript{1}
\textsuperscript{1}Department of Mathematics and Computer Science, Eindhoven University of Technology
\textsuperscript{2}Institute for Complex Molecular Systems (ICMS), Eindhoven University of Technology

June 8, 2021

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 [FK06]. 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 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.

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 [JAP97, How01, KF07, Kol13].
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 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 [HKM08b, HKM08a]. Vorotnikov proved sufficient conditions for transport in deterministically switching [Vor11] and randomly switching systems [Vor14]. Perthame, Souganidis, and Mirrahimi developed a dynamic point of view on systems of molecular motors [PS09a, PS09b, MS13]. 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 [MS13] 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, Souganidis and Mirrahimi, which appears to be inspired by large-deviation theory, as evidenced by the title of [PS09a] 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 [Kol13] and Section 2, as well as [PS09a, PS09b, MS13, HKM08b, HKM08a]).

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 [KP17] 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 [KP17, Lemma 1], we formulate conditions that imply the comparison principle. Faggionato and Silvestri establish large-deviation principles for fully discrete, ‘pseudo-one-dimensional’ systems [FS17].

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. [FL96, HY14, HMS16, BDG18, KS20]). 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 [MS13, 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 [WPE03, 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 [MS13], 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).
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 \(T \times \{1, 2\}\), where \(T = \mathbb{R}/\mathbb{Z}\) is the one-dimensional flat torus. We fix the initial condition \((X^n(0), I^n(0)) = (x_0, i_0)\) for some \((x_0, i_0) \in T \times \{1, 2\}\). 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

\[
\frac{dX^n_t}{dt} = -\psi'(nX^n_t, I^n_t) dt + \frac{1}{\sqrt{n}} dB_t, \tag{1}
\]

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

\[
P\left( I^n(t+\Delta t) = j \mid I^n(t) = i, X^n(t) = x \right) = n \cdot r_{ij}(nx) \Delta t + O(\Delta t^2), \quad \text{as } \Delta t \to 0. \tag{2}
\]
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 [YZ10, 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 $\mathbb{R}$.

![Figure 2: A typical time evolution of $(X^n, I^n)$ satisfying (1) and (2). In the left diagram, the black bullet represents a particle that moves according to (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 $\mathbb{R}$.]

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

$$dX_t = -\psi'(X_t, I_t) \, dt + dB_t,$$

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

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

The large-scale behaviour of $(X_t, I_t)$ is studied by considering the rescaled process $(X^n_t, I^n_t)$ defined by $X^n_t := \frac{1}{n} X_{nt}$ and $I^n_t := I_{nt}$, and characterizing the dynamics of $(X^n_t, I^n_t)$ 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 $(X^n_t, I^n_t)$ satisfies (1) and (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 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 (1) and (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 $\mathcal{X} := C_T[0, \infty)$ the set of continuous trajectories in $T$, 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 $P(X^n \in \cdot) \in \mathcal{P}(\mathcal{X})$. We will show that there exists a rate function $I : \mathcal{X} \to [0, \infty]$ 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

$$P(X^n \approx x) \sim e^{-nI(x)}, \quad n \to \infty.$$  \hfill (3)

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

$$I(x) = I_0(x(0)) + \int_0^\infty L(\partial_t x(t)) \, dt.$$  \hfill (4)

Here $I_0 : T \to [0, \infty]$ 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 $I_0(x_0) = 0$ and $+\infty$ otherwise. The Lagrangian is the Legendre dual of a Hamiltonian $H : \mathbb{R} \to \mathbb{R}$, that is $L(v) = \sup_{p \in \mathbb{R}} [pv - H(p)]$, 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 (4) has the following properties:

(i) $I : \mathcal{X} \to [0, \infty]$ is nonnegative.
(ii) $I(x) = 0$ if and only if $\partial_t x(t) = v$, with $v = H'(0)$.
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 $I(x) > 0$ for a path $x \in \mathcal{X}$, then by (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 $H(p)$. For the Hamiltonian of this example, strict convexity can be proven as demonstrated in [MS13, 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 $v = H'(0)$. 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 $r_{12}e^{-\psi_1} = C r_{21}e^{-\psi_2}$ 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 [EK86, Section 3.5]. We work with the definition of a rate function as given in [BD19, Chapter 1].

**Definition 3.1 (Rate function).** We call a map $I : \mathcal{X} \to [0, \infty]$ a rate function if for every $C \geq 0$, the sub-level set $\{x \in \mathcal{X} : I(x) \leq C\}$ is compact. In particular, a rate function is lower semi-continuous. For a Borel subset $A \subseteq \mathcal{X}$, we write $\text{int}(A)$ and $\text{clos}(A)$ for its interior and closure.

**Definition 3.2 (Large-deviation principle).** For $n = 1, 2, \ldots$, let $P_n$ be a probability measure on $\mathcal{X}$, and let $I : \mathcal{X} \to [0, \infty]$ be a rate function. We say that the sequence $\{P_n\}_{n \in \mathbb{N}}$ satisfies a large-deviation principle with rate function $I$ if for every Borel subset $A \subseteq \mathcal{X}$,

$$- \inf_{x \in \text{int}(A)} I(x) \leq \liminf_{n \to \infty} \frac{1}{n} \log P_n(A) \leq \limsup_{n \to \infty} \frac{1}{n} \log P_n(A) \leq - \inf_{x \in \text{clos}(A)} I(x).$$

A large-deviation principle provides an estimate of the probabilities $P_n(A)$ on the logarithmic scale. At least intuitively,

$$P_n(A) \approx e^{-n \inf_{x \in A} I(x)}, \quad n \to \infty.$$ 

Illustrating examples of a large-deviation principle can be found for instance in Ellis’ note on Boltzmann’s discoveries [Ell99]. General introductions to the topic are also provided in [BD19, Chapter 1] and [FK06, 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 $AC([0, \infty); \mathbb{T}^d)$ 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 $I : D_{\mathbb{R}^d}[0, \infty) \to [0, \infty]$ is of action-integral form if there is a non-trivial convex map $L : \mathbb{R}^d \to [0, \infty]$ with which

$$I(x) = \begin{cases} I_0(x(0)) + \int_0^\infty L(\partial_t x(t)) \, dt & \text{if } x \in AC([0, \infty); \mathbb{T}^d), \\ +\infty & \text{otherwise,} \end{cases}$$

where $I_0 : \mathbb{T}^d \to [0, \infty]$ is a rate function. We refer to the map $L$ as the Lagrangian. \hfill \Box

*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 $f \in \mathbb{N}$. For $n \in \mathbb{N}$, the state space $E_n$ is a product space $E_n := E_n^X \times \{1, \ldots, f\}$, where $E_n^X$ be a compact Polish space satisfying the following: there are continuous maps $I_n : E_n^X \to \mathbb{T}^d$ such that for all $x \in \mathbb{T}^d$ there exists $x_n \in E_n^X$ with which $I_n(x_n) \to x$ as $n \to \infty$. \hfill \Box

This condition means that the $E_n^X$ are asymptotically dense in the torus $\mathbb{T}^d$. The typical example is the periodic lattice $E_n = (n^{-1})^d / \mathbb{Z}^d$, 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 $I_n$ in the notation.

Let $X_n := D_{E_n}[0, \infty)$. For a distribution $\mu \in \mathcal{P}(E_n)$, we define an $E_n$-valued two-component process $(X^n, I^n)$ with initial condition $\mu$ by defining its path distribution $\mathbb{P}_\mu^n \in \mathcal{P}(X_n)$. In order to define a path distribution, we shall specify a linear map $L_n : \mathcal{D}(L_n) \subseteq \mathcal{C}(E_n) \to \mathcal{C}(E_n)$ on a domain $\mathcal{D}(L_n)$ and assume well-posedness of the martingale problem of the pair $(L_n, \mu)$; we refer to [EK86, Section 4.3] for a precise treatment of the martingale problem. We call a linear map $L_n$ as above a *generator* if it gives rise to a well-posed martingale problem. We specify the generators of $(X^n, I^n)$ from the following ingredients:

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

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

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

$$L_nf(x,i) := L_n^i f(\cdot, i)(x) + \sum_{j=1}^{f} r_{ij}^n(x) [f(x,j) - f(x,i)], \quad (5)$$

where the domain is $\mathcal{D}(L_n) = \{f \in \mathcal{C}(E_n) : f(\cdot, i) \in \mathcal{D}(L_n^i), i = 1, \ldots, f\}$. 

8
Condition 3.5 (Well-posedness). Let $\mu \in \mathcal{P}(E_n)$. Existence and uniqueness holds for the $\mathcal{X}_n$ martingale problem for $(L_n, \mu)$. Denote the solution to the martingale-problem solution of $L_n$ by $\mathbb{P}^n_\mu$. The map $E_n \ni z \mapsto \mathbb{P}^n_\delta_z \in \mathcal{P}(\mathcal{X}_n)$ is Borel measurable with respect to the weak topology on $\mathcal{P}(\mathcal{X}_n)$.

Condition 3.5 is the basic assumption on the processes in [FK06]. A sufficient condition for the measurability in there is given in [EK86, Theorem 4.4.6]. We do not give general conditions on a map $L_n$ that imply Condition 3.5. For examples regarding existence and regularity properties we refer to the book of Yin and Zhu about switching hybrid diffusions [YZ10, Part I].

Definition 3.6 (Switching Markov processes in a periodic setting). Let $(X^n, I^n)$ be a two-component Markov process taking values in $E_n = E^n_X \times \{1, \ldots, J\}$ satisfying Condition 3.4. We call $(X^n, I^n)$ a switching Markov process if its generator $L_n$ is given by (5) and satisfies Condition 3.5.

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 (5). 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 (5). The nonlinear generator is the map $H_n : D(H_n) \subseteq C(E_n) \to C(E_n)$ defined by

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

(6)

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

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:

$$
\begin{array}{c}
\mathcal{T}^d \times E' \\
\downarrow \text{proj} \\
\mathcal{T}^d \\
\end{array}
\xleftarrow{\eta_n} \xrightarrow{\eta_n} \begin{array}{c}
E_n \\
\end{array}
$$

In the diagram, $\eta_n : E_n \to \mathcal{T}^d$ is the projection defined by $\eta_n(x, i) := \iota_n(x)$, where $\iota_n : E_n^* \to \mathcal{T}^d$ is the embedding of Condition 3.4. The map $\eta_n : E_n \to E'$ is assumed to be continuous. We shall assume that the $E_n$ are asymptotically dense:

(C1) For $(x, z') \in \mathcal{T}^d \times E'$ 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 $H \subseteq C(\mathcal{T}^d) \times C(\mathcal{T}^d \times E')$, a multi-valued operator. We shall assume the following convergence condition:

(C2) The domain $\mathcal{D}(H)$ satisfies $C^\infty(\mathcal{T}^d) \subseteq \mathcal{D}(H) \subseteq C^1(\mathcal{T}^d)$. 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 $\mathcal{C} \subseteq C(E')$ and functions $H_{f, \varphi} \in C(\mathcal{T}^d \times E')$ with which

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

**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 $H \subseteq C(\mathcal{T}^d) \times C(\mathcal{T}^d \times E')$ be a multivalued 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 $f \in \mathcal{D}(H)$,

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

(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 $\mathcal{T}^d$ with rate function $I_0 : \mathcal{T}^d \to [0, \infty]$. Then the family of processes $\{X^n\}_{n \in \mathbb{N}}$ satisfies a large-deviation principle in $\mathcal{D}^2(0, \infty)$ with a rate function $I : \mathcal{D}^2(0, \infty) \to [0, \infty]$, and there exists a semigroup $V(t)$ with which the rate function is given by (9).

We give the proof in Section 5. The formula for the rate function $I$ is not important here, which is why we report it only below in (9) in the proof section. Condition (T1) means that the images depend on the variable $x \in \mathcal{T}^d$ 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 (9) 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

\[
L(v) = \sup_{p \in \mathbb{R}^d} [ p \cdot v - \mathcal{H}(p) ].
\]

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

\[
L_n f(x, i) := b^i(nx) \cdot \nabla_x f(x, i)(x) + \frac{1}{n} \Delta_x f(x, i)(x) + \sum_{j \neq i} \frac{\gamma(n)}{n} r_{ij}(nx) [ f(x, j) - f(x, i) ], \tag{7}
\]

where \( \gamma(n) > 0, r_{ij}(\cdot) \in C^\infty(T^d; [0, \infty)), \) and \( b^i(\cdot) \in C^\infty(T^d) \).

This is an example of a switching Markov process with generators \( L^n_i \) defined on the core \( C^2(T^d) \) by

\[
L^n_i g(x) := b^i(nx) \cdot \nabla_x g(x) + \frac{1}{n} \Delta_x g(x),
\]

and rates \( r^n_{ij}(x) = \gamma(n) r_{ij}(nx) \). The domain \( \mathcal{D}(L^n_i) \) of the generators \( L^n_i \) contains the core, but is larger than \( C^2(T^d) \). 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 \( f(\cdot, i) \in C^2(T^d) \), the generator acts as defined in (7). 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 \( \gamma(n) = n \).

**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 \( \{1, \ldots, J\} \) into two disjoint sets \( J_1 \) and \( J_2 \) such that \( A_{ij} = 0 \) whenever \( i \in J_1 \) and \( j \in J_2 \).
Theorem 4.6 (Limit I). Let \((X^0_t, \mathcal{I}^0_t)\) 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 \(I_0 : \mathbb{T}^d \to [0, \infty]\).

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 [Swe92].

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 \((X^0_t, \mathcal{I}^0_t)\) be the Markov process of Definition 4.4, with parameter \(\gamma(n)\) such that \(n^{-1} \gamma(n) \to 0\) 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 \(I_0 : \mathbb{T}^d \to [0, \infty]\).

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 \(b'(y) = -\nabla_y \phi_i(y)\) and \(\gamma(n) = n\) in the generators defined in (7). We say that the set of potentials and rates \(\{r_{ij}, \psi_i\}\) satisfies detailed balance if for all \(i, j \in \{1, \ldots, J\}\) and \(y \in \mathbb{T}^d\), we have

\[
    r_{ij}(y)e^{-2\phi_i(y)} = r_{ji}(y)e^{-2\phi_j(y)}. \tag{8}
\]

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 (8) 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

\[
    \mathcal{H}(p) = \sup_{\mu \in \mathcal{P}} \left[ K_{\mathcal{P}}(\mu) - \mathcal{R}(\mu) \right],
\]
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 (31), and $K_p(\mu)$ is given by

$$K_p(\mu) = \inf_{\phi} \left\{ \sum_{i=1}^{j} \int_{\mathbb{T}^d} \left( \frac{1}{2} |\nabla \phi_i(x) + p|^2 - \sum_{j=1}^{j} r_{ij}(x) \right) d\mu_i(x) + \sum_{i,j=1}^{j} \int_{\mathbb{T}^d} \pi_{ij}(x) \sqrt{\pi_i(x) \pi_j(x)} e^{\phi_i(x) + \phi_j(x)} \cosh (\phi(x,j) - \phi(x,i)) \, dx \right\},$$

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

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(·) terms appear is symmetric in the sense that

$$C(\phi) := \sum_{i,j=1}^{j} \int_{\mathbb{T}^d} \pi_{ij}(x) \sqrt{\pi_i(x) \pi_j(x)} e^{\phi_i(x) + \phi_j(x)} \cosh (\phi(x,j) - \phi(x,i)) \, dx$$

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

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 $b'(y) = F - \psi'(y, i)$, where $F$ is a constant (modeling an external force) and $\psi \in C^\infty(\mathbb{T})$ is a smooth periodic potential. The process $(X^n, I^n)$ is $\mathbb{T} \times \{1, \ldots, j\}$-valued and satisfies

$$dX^n_t = (F - \psi'(nX^n_t, I^n_t)) \, dt + \frac{1}{\sqrt{n}} dB_t,$$

where $I^n_t$ a jump process on $\{1, \ldots, j\}$ with jump rates $nr_{ij}(nx)$. Under detailed balance, one can show with arguments similar as above that the Hamiltonian for this process is symmetric around $(-F)$. Since $\mathcal{H}(0) = 0$ and $\mathcal{H}(p)$ is strictly convex, this means that the model predicts a positive force-velocity feedback under detailed balance: $F > 0$ implies $\partial_p \mathcal{H}(0) > 0$, and $F < 0$ implies $\partial_p \mathcal{H}(0) < 0$.

**Theorem 4.9 (Separation of time scales implies a symmetric Hamiltonian).** Let the stochastic process $(X^n, I^n)$ of Definition 4.4, with $b'(y) = -\nabla \psi'$, 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

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

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

\[
\mathbf{P} = \{ \mu \in \mathcal{P}(\mathbb{T}^d) : \mu \ll dx \text{ and } d\mu = \nabla \log(\mu) \in L^2(\mathbb{T}^d) \}.
\]

The map \( \mathcal{R} \) is the relative Fisher information; with the stationary measure \( \nu \) of the jump process on \( \{1, \ldots, f\} \) with rates \( r_{ij} \),

\[
\mathcal{R}(\mu) = \frac{1}{8} \int_{\mathbb{T}^d} \left| \nabla \log \left( \frac{\mu}{e^{-2\varphi}} \right) \right|^2 d\mu, \quad \varphi(x) = \sum_i v_i \psi_i(x).
\]

\[\square\]

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 [FK06] to our setting.

5.1 Strategy of proof

Viscosity solutions and comparison principle. We adapt [FK06, 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 multivalued operators. For the following definition, \( E \) and \( E' \) are compact Polish spaces, \( B(E \times E') \) 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 multivalued 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

\[
(u_1 - f)(x) = \sup (u_1 - f) \quad \text{and} \quad u_1(x) - \tau g(x, z') - h(x) \leq 0.
\]

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

\[
(f - u_2)(x) = \sup (f - u_2) \quad \text{and} \quad u_2(x) - \tau g(x, z') - h(x) \geq 0.
\]
iii) A function \( u_1 : E \rightarrow \mathbb{R} \) 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 [FK06]. First, formulating viscosity solutions via sequences as in [FK06, 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 [FK06].

**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 \( E \).

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 \( \eta_n \) and \( \eta_n' \) by

\[\begin{array}{ccc}
E_n & \xrightarrow{(\eta_n, \eta_n')} & E \\
\downarrow & & \downarrow \text{proj}_1 \\
E & & \end{array}\]

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 [FK06, Theorem 7.18] to our setting. This adaptation is obtained by collecting in one place assumptions that are mentioned in several places in [FK06], and specializing them to the compact setting.

**Theorem 5.3.** Let \( L_n \) be the generator of an \( E_n \)-valued process \( Y^n \), and let \( H_n \) be the nonlinear generators defined by \( H_n f = \frac{1}{n} e^{-nf} L_n e^{nf} \). 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 [FK06] 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 $K_1^n \subseteq E$ and $K_2^n \subseteq E \times E'$ such that
\[
\sup_{y \in A_1^n} \inf_{x \in K_1^n} d_E(\eta_n(y), x) \xrightarrow{n \to \infty} 0,
\]
and
\[
\sup_{y \in A_1^n} \inf_{x, z \in K_2^n} [d_E(\eta_n(y), x) + d_E(\eta_n(y), z)] \xrightarrow{n \to \infty} 0.
\]

(d) For each compact $K \subseteq E$, there exists $q \in Q$ such that $K \subseteq \lim inf n \eta_n(A_1^n)$.

(ii) (Convergence Condition 7.11 of [FK06]) There exist multifunctions $H_t, H_t^q \subseteq C(E) \times C(E \times E')$ which are the limit of the $H_n$'s in the following sense:

(a) For each $(f, g) \in H_t$, there exist $f_n \in D(H_n)$ 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 $q \in Q$, $\lim_{n \to \infty} \sup_{y \in A_1^n} |f_n(y) - f(\eta_n(y))| = 0$. Furthermore, for each $q \in Q$ and every sequence $y_n \in A_1^n$ such that $\eta_n(y_n) \to x \in E$ and $\eta_n'(y_n) \to z' \in E'$, we have $\limsup_{n \to \infty} H_n f_n(y_n) \leq g(x, z')$.

(b) For each $(f, g) \in H_t^q$, there exist $f_n \in D(H_n)$ (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 $q \in Q$, $\lim_{n \to \infty} \sup_{y \in A_1^n} |f_n(y) - f(\eta_n(y))| = 0$. 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 $\liminf_{n \to \infty} H_n f_n(y_n) \geq g(x, z')$.

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

Let $X^n_t := \eta_n(Y^n_t)$ 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 $I_0 : E \to [0, \infty]$.

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

\[
I(x) = I_0(x(0)) + \sup_{k \in \mathbb{N}} \sup_{(t_1, \ldots, t_k) \in \mathbb{R}^k} \sum_{i=1}^k I_{t_i-t_{i-1}}(x(t_i)|x(t_{i-1})), \quad (9)
\]

where for $z, y \in E$,
\[
I_t(z|y) = \sup_{f \in C(E)} [f(z) - V(t)f(y)]. \quad (10)
\]

The semigroup $V(t)$ is defined via the Crandall-Liggett Theorem—for details we refer to [FK06, Chapter 5].

16
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.

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 multivalued 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 multivalued 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 \( H = H_1 = H_2 \). 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 \( H = H_1 = H_2 \). Therefore the large-deviation principle follows by Theorem 5.3.

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

\[
\begin{array}{ccc}
(\eta_n, \eta_n') & \xrightarrow{\text{proj}_1} & \mathbb{T}^d \\
E_n & \xrightarrow{\eta_n} & E' \\
\downarrow \iota_n & & \\
\mathbb{T}^d & & \\
\end{array}
\]

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 \( x_n \in E_n^q \) such that \( \iota_n(x_n) \to x \). Part (c) follows by taking the compact sets \( K_1^q := \mathbb{T}^d \) and \( K_2^q := \mathbb{T}^d \times E' \).

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

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

With these \( f_n \), both conditions (a) and (b) are simultaneously satisfied for the operator \( H = H_1 = H_2 \), where 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 \( \eta_n'(y_n) \to z' \). The required boundedness,

\[
\sup_{n \in \mathbb{N}} \left( \sup_{y \in E_n} |f_n(y)| + \sup_{y \in E_n} |H_n f_n(y)| \right) < \infty,
\]
follows from the uniform-convergence condition (C2).

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

\[
H_1 f(x) := \inf \sup_{\varphi} H_{f,\varphi}(x, z') \quad \text{and} \quad H_2 f(x) := \sup \inf_{\varphi} H_{f,\varphi}(x, z'),
\]

with equal domains \( D(H_1) = D(H_2) := D(H) \). 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 \( H_1 f(x) = H_1(\nabla f(x)) \) and \( H_2 f(x) = H_2(\nabla f(x)) \), with two maps \( H_1, H_2 : \mathbb{R}^d \to \mathbb{R} \). We prove Proposition 5.5 with the following Lemmas.

**Lemma 5.6** (Local operators admit strong solutions). Let \( H \subseteq C^1(T^d) \times C(T^d \times E') \) be a multivalued limit operator satisfying (T1) of Theorem 4.2. Then for any \( \tau > 0 \) and \( h \in C(T^d) \), 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 multivalued operator satisfying (T1) and (T2) of Theorem 4.2. For all \( h \in C(T^d) \) and \( \tau > 0 \), strong viscosity subsolutions \( u_1 \) of \( (1 - \tau H_1)u = h \) are strong viscosity subsolutions of \( (1 - \tau H_1)u = h \), and strong viscosity supersolutions \( u_2 \) of \( (1 - \tau H_1)u = h \) 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 multivalued operator satisfying (T1) and (T2) of Theorem 4.2. Then \( H_1(p) \leq H_2(p) \) for all \( p \in \mathbb{R}^d \).

**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

\[
\max_{T^d}(u_1 - u_2) \leq \tau [H_1(p_\delta) - H_2(p_\delta)] + h(x_\delta) - h(x'_\delta), \tag{11}
\]

with some \( x_\delta, x'_\delta \in T^d \) such that \( \text{dist}(x_\delta, x'_\delta) \to 0 \) as \( \delta \to 0 \), and certain \( p_\delta \in \mathbb{R}^d \). Then using that \( h \in C(T^d) \) is uniformly continuous since \( T^d \) is compact, and that \( H_1(p_\delta) \leq H_2(p_\delta) \) by Lemma 5.8, we can further estimate as

\[
\max_{T^d}(u_1 - u_2) \leq h(x_\delta) - h(x'_\delta) \leq \omega_h(\text{dist}(x_\delta, x'_\delta)),
\]

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 \( (u_1 - u_2) \leq 0 \) follows by taking the limit \( \delta \to 0 \).

We are left with proving (11). Define \( \Phi_\delta : T^d \times T^d \to \mathbb{R} \) by

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

where

\[
\Psi(x, x') := \sum_{j=1}^d \sin^2 \left( \pi(x_j - x'_j) \right), \quad \text{for all } x, x' \in T^d. \tag{12}
\]
Then $\Psi \geq 0$, and $\Psi(x, x') = 0$ holds if and only if $x = x'$, and
\[
\nabla_1 \left[ \Psi(\cdot, x') \right](x) = -\nabla_2 \left[ \Psi(x, \cdot) \right](x') \quad \text{for all } x, x' \in \mathbb{T}^d. \quad (13)
\]
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 $\Phi_\delta(x_\delta, x'_\delta) \leq \Phi(x_\delta, x'_\delta)$ and $u_2$ is bounded, we obtain
\[
\Psi(x_\delta, x'_\delta) \leq 2\delta (u_2(x_\delta) - u_2(x'_\delta)) \leq 4\delta \|u_2\|_{L^\infty(\mathbb{T}^d)} = O(\delta).
\]
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 $f_1^\delta$ and $f_2^\delta$ as
\[
f_1^\delta(x) := u_2(x'_\delta) + \frac{\Psi(x, x'_\delta)}{2\delta} \quad \text{and} \quad f_2^\delta(x') := u_1(x_\delta) - \frac{\Psi(x_\delta, x')}{2\delta},
\]
Then $f_1^\delta, f_2^\delta \in C^\infty(\mathbb{T}^d) \subseteq D(H)$ are both in the domain of $H$, and hence in the domain of $H_1$ and $H_2$, respectively. Furthermore, $(u_1 - f_1)$ has a maximum at $x = x_\delta$, and $(f_2 - u_2)$ has a maximum at $x' = x'_\delta$, by definition of $(x_\delta, x'_\delta)$ and $\Phi_\delta$. Since $u_1$ is a strong subsolution of $(1 - \tau H_1)u = h$,
\[
u_1(x_\delta) - \tau H_1 f_1^\delta(x_\delta) - h(x_\delta) \leq 0,
\]
and since $u_2$ is a strong supersolution of $(1 - \tau H_2)u = h$,
\[
u_2(x'_\delta) - \tau H_2 f_2^\delta(x'_\delta) - h(x'_\delta) \geq 0.
\]
Thereby, we can estimate $\max(u_1 - u_2)$ as
\[
\max_{\mathbb{T}^d}(u_1 - u_2) \leq \Phi_\delta(x_\delta, x'_\delta)
\]
\[
\leq u_1(x_\delta) - u_2(x'_\delta)
\]
\[
\leq \tau \left[ H_1 f_1^\delta(x_\delta) - H_2 f_2^\delta(x'_\delta) \right] + h(x_\delta) - h(x'_\delta)
\]
\[
= \tau \left[ H_1 (\nabla f_1^\delta(x_\delta)) - H_2 (\nabla f_2^\delta(x'_\delta)) \right] + h(x_\delta) - h(x'_\delta).
\]
By (13), $\nabla f_1^\delta(x_\delta) = \nabla f_2^\delta(x'_\delta) =: p_\delta \in \mathbb{R}^d$, which establishes (11), and thereby finishes the proof. $\square$

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 [FK06, 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 $(f, H_{f,p}) \in H$, and let $x \in \mathbb{T}^d$ be such that $(u_1 - f)(x) = \sup(u_1 - f)$. 

19
The function \( \tilde{f} \) defined by \( \tilde{f}(x') := \Psi(x', x) \), with \( \Psi(x', x) \) defined by (12), 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

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

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

\[
H_{f, \varphi}(x, z') = H_{\varphi}((\nabla f + \nabla \tilde{f})(x), z') = H_{\varphi}(\nabla f(x), z') = H_{f, \varphi}(x, z').
\]

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 \( (u_1 - f)(x) = \sup_{x \in \mathbb{T}^d} (u_1 - f) \) for a point \( x \in \mathbb{T}^d \), then there exists a point \( z' \in E' \) such that

\[
 u_1(x) - \tau H_{f, \varphi}(x, z') - h(x) \leq 0. \tag{14}
\]

Let \( f \in \mathcal{D}(H_1) = \mathcal{D}(H) \) and \( x \in \mathbb{T}^d \) be such that \( (u_1 - f)(x) = \sup_{z' \in E'} u_1(z') \). For any \( \varphi \) there exists a point \( z' \in E' \) such that the above subsolution inequality (14) 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 \( \varphi \), we obtain

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

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

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

20
Taking the infimum and supremum over $\varphi$, we find
\[
H_1(p) = \inf_{\varphi} \sup_{z'} H_{\varphi}(p, z') \\
\leq \sup_{z'} H_{\varphi}(p, z') = H(p) = \inf_{\varphi} H_{\varphi}(p, z') \\
\leq \sup_{\varphi} \inf_{z'} H_{\varphi}(p, z') = H_2(p),
\]
which finishes the proof.  \hfill $\Box$

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 [FK06, Chapter 8]. We first briefly summarize the strategy in Section 6.1, specialized to our setting.

6.1 Strategy of proof

Let $H = H(p)$ be the Hamiltonian of Theorem 4.3 and let $L = L(v)$ be the associated Lagrangian defined by
\[
L(v) := \sup_{p \in \mathbb{R}^d} [p \cdot v - H(p)] \tag{15}
\]
Define $V_{NS}(t) : C(\mathbb{T}^d) \rightarrow C(\mathbb{T}^d)$ by
\[
V_{NS}(t)f(x) = \sup_{\gamma \in AC_{\mathbb{T}^d}(0, \infty)} \left[ f(\gamma(t)) - \int_0^t L(\partial_s \gamma(s)) \, ds \right], \tag{16}
\]
where $AC_{\mathbb{T}^d}(0, \infty)$ is the set of absolutely continuous paths in the torus. The map $V_{NS}(t)$ is the Nisio semigroup with cost function $L$. In Definition 8.1 and Equation (8.10) in [FK06], the Nisio semigroup is defined by means of relaxed controls in order to cover a general class of possible cost functions. Since the Lagrangian $L(v)$ is convex, the semigroup $V_{NS}(t)$ equals the semigroup given in (8.10) of [FK06], which can be seen by using that $\lambda = \delta_{x(s)}$ is an admissible control and by applying Jensen’s inequality. Such an argument is given for example in Theorem 10.22 in [FK06].

The rate function $I$ of Theorem 4.2 is given in terms of a limiting semigroup $V(t)$ as shown in equations (9) and (10). The desired action-integral representation follows if the semigroup $V(t)$ of Theorem 4.2 is equal to the Nisio semigroup $V_{NS}(t)$ defined by (16). In [FK06, Chapter 8], the equality of semigroups is traced back to conditions on their generators. In our case, the generator of the limiting semigroup is the limiting multivalued operator $H$ of Theorem 4.2, and the generator of the Nisio semigroup is an operator $H$ defined by the Hamiltonian $H(p)$. We summarize in Proposition 6.1 below that the generators satisfy the required conditions of [FK06, 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 $H : D(H) \subseteq C^1(T^d) \to C(T^d)$ on the domain $D(H) = D(H)$ by setting $Hf(x) := \mathcal{H}(\nabla f(x))$. Let $\tau > 0$ and $h \in C(T^d)$. Then:

(i) The Lagrangian (15) and the operator $H$ satisfy Conditions 8.9, 8.10 and 8.11 of [FK06], with the set of controls $U = \mathbb{R}^d$, operator $Af(x, u) = \nabla f(x) \cdot u$, cost function $L(x, u) = L(u)$, and $H = H = H$.

(ii) The comparison principle (Definition 5.2) holds for viscosity sub- and supersolutions of $(1 - \tau H)u = h$.

(iii) Every viscosity solution $u$ of $(1 - \tau H)u = h$ is also a viscosity solution of $(1 - \tau H)u = h$.

Proof of Theorem 4.3. Let $V(t)$ be the semigroup obtained in Theorem 4.2 and let $V_{\text{NS}}(t)$ be the Nisio semigroup (16). We shall verify that $V(t) = V_{\text{NS}}(t)$. Then by [FK06, Theorem 8.14], the rate function of Theorem 4.2 (given by (9)) satisfies the control representation (8.18) of [FK06]. The action-integral representation follows from this control representation by applying Jensen’s inequality.

By [FK06, Theorem 8.27], we obtain $V_{\text{NS}}(t) = V(t)$, where the semigroup $V(t)$ is defined by

$$V(t) = \lim_{m \to \infty} \left[ (1 - \frac{t}{m} H)^{-1} \right]^m. \quad (17)$$

The conditions of Theorem 8.27 are satisfied since Conditions 8.9, 8.10 and 8.11 of [FK06] are satisfied by Item (i), and since the comparison principle holds by Item (ii).

By [FK06, Corollary 8.29], we obtain $V(t) = V(t)$. The conditions of Corollary 8.29 are satisfied: Item (iii) above corresponds to Item a) of Corollary 8.29, the conditions of [FK06, Theorem 6.14] are satisfied under the assumptions of our Theorem 4.2, the conditions of [FK06, Theorem 8.27] are satisfied for the same reasons as mentioned above, and $D_a = D(H)$. \hfill \square

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 [FK06], which are formulated in order to cover a more general and non-compact setting.

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

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

$$|\nabla f(x_0) \cdot v| \leq \psi_f(L(v)) \quad \text{and} \quad \lim_{r \to \infty} \frac{\psi_f(r)}{r} = 0.$$
Hence by monotonicity of $\psi$ of $\phi$

Let $\psi$ Regarding Items (1)-(5) of [FK06, Condition 8.9], the operator Item (c) under the milder assumption of convexity of a differential equation with a globally bounded vectorfield. Here, we verify $r$ and consequently is superlinear, because $H$

3.2.3 under the assumption of continuous differentiability of $\partial H$. For $p_0 \in \mathbb{R}^d$, define the subdifferential $\partial H(p_0)$ at $p_0$ as the set

\[ \partial H(p_0) := \{ \xi \in \mathbb{R}^d \mid \forall p \in \mathbb{R}^d : H(p) \geq H(p_0) + \langle \xi, p - p_0 \rangle \}. \]
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 (18) follows by noting that \( \mathcal{H}(\nabla f(y)) \geq \nabla f(y) \cdot \nu - \mathcal{L}(\nu) \) for all \( y \in \mathbb{T}^d \) and \( \nu \in \mathbb{R} \), 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 (18). Regarding the other inequality, since \( x \in \partial \mathcal{H}(\nabla f(x)) \), 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 \( [0, \infty) \),

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

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 [Dei92] 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: \( \sup_{x \in \mathbb{R}^d} \|F(x)\|_{\sup} \) 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

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

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. [Roc66, 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 [Dei92]: 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 \( \xi_n \) is bounded, and passing to a subsequence if necessary, it converges to some \( \xi \). 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. \hfill \Box

Proof of (ii) in Proposition 6.1. The comparison principle for the operator $\mathcal{H}$ follows from the fact that $\mathcal{H}f = \mathcal{H}(\nabla f)$ depends on $x$ only via gradients. Indeed, for subsolutions $u_1$ and supersolutions $u_2$ of $(1 - \tau \mathcal{H})u = h$, 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 $f_1, f_2 \in \mathcal{D}(\mathcal{H})$ satisfying $\nabla f_1(x_\delta) = \nabla f_2(x_\delta')$, and 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 $\delta \to 0$. \hfill \Box

Proof of (iii) in Proposition 6.1. Let $u \in C(\mathbb{T}^d)$ be a viscosity solution of the equation $(1 - \tau \mathcal{H})u = h$. By Lemmas 5.6 and 5.7, $u$ is a strong viscosity subsolution of $(1 - \tau \mathcal{H}_1)u = h$ and a strong viscosity supersolution of $(1 - \tau \mathcal{H}_2)u = h$. 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 $-\mathcal{H}_1 \geq -\mathcal{H} \geq -\mathcal{H}_2$. With that, we find that $u$ is both a strong viscosity sub- and supersolution of $(1 - \tau \mathcal{H})u = h$. \hfill \Box

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
\[
L_n f(x,i) = \frac{1}{n} \left[ \frac{1}{2} \Delta_x f(x,i)(x) + b^i(nx) \cdot \nabla_x f(x,i)(x) \right.
+ \sum_{j=1}^f \gamma(n) r_{ij}(nx) \left[ f(x,j) - f(x,i) \right],
\]
with state space $E_n = \mathbb{T}^d \times \{1, \ldots, f\} = \{(x,i)\}$, 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 by
\[
H_n f(x,i) = \frac{1}{n} \left[ \frac{1}{2} \Delta_x f^i(x)(x) + \frac{1}{2} |\nabla_x f^i(x)|^2 + b^i(nx) \nabla_x f^i(x) \right.
+ \left. \frac{1}{n} \gamma(n) \sum_{j=1}^f r_{ij}(nx) \left[ e^{n(f(x,j)-f(x,i))} - 1 \right]. \right. \] (19)
7.1 Proof of Theorem 4.6

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

\[
H_n f_n(x, i) = \frac{1}{n^2} \Delta f(x) + \frac{1}{2} \Delta y \varphi^i(nx) + \frac{1}{2} \| \nabla f(x) + \nabla y \varphi^i(nx) \|^2_{\mathbb{R}^J} + b^i(nx) \left( \nabla f(x) + \nabla y \varphi^i(nx) \right) + \sum_{j=1}^J r_{ij}(nx) \left[ e^{\varphi^i(nx, j)} - \varphi(nx, i) - 1 \right],
\]

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

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

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

\[
H_{f, \varphi}(x, y, i) := \frac{1}{2} \Delta y \varphi^i(y) + \frac{1}{2} \| \nabla f(x) + \nabla y \varphi^i(y) \|^2 + b^i(y)(\nabla f(x) + \nabla y \varphi^i(y)) + \sum_{j=1}^J r_{ij}(y) \left[ e^{\varphi^i(y, j)} - \varphi(y, i) - 1 \right],
\]

where \( \varphi = (\varphi^1, \ldots, \varphi^d) \) via the identification \( C^2(E') \simeq (C^2(\mathbb{T}^d))^J \).

We now verify (C1), (C2) and (C3) of (T1). For (C1), define the maps \( \eta_n^i : E_n \to E' \) by \( \eta_n^i(x, i) := (nx, i) \), 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 \( (y_n, i_n) \in \mathbb{T}^d \times \{1, \ldots, J\} \) such that both \( \eta_n(y_n, i_n) \to x \) and \( \eta_n^i(y_n, i_n) \to (y, i) \) as \( n \to \infty \). For \( d = 1 \), the point \( y_n := \frac{1}{n}(nx + y) \) satisfies \( y_n \to x \) and \( \eta_n^i = y \) in \( \mathbb{T}^d \) (i.e. modulo 1). For \( d \geq 2 \) this construction can be done for each coordinate. Therefore, (C1) holds with \( y_n = \frac{1}{n}(nx + y) \) and \( i_n = i \).

Regarding Item (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

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

and

\[
\| H_{f, \varphi} \circ \eta_n^i - H_n f_n \|_{L^\infty(E_n)} = \sup_{(x, i) \in E_n} | H_{f, \varphi}(x, nx, i) - H_n f_n(x, i) | = \frac{1}{2} \sup_{(x, i) \in E_n} \| \Delta f(x) \| \leq \frac{1}{n^2} \| \varphi \|_{L^\infty(E_n)} \to 0.
\]

Item (C3), the fact that the images \( H_{f, \varphi} \) depend on \( x \) only via the gradients of \( f \), can be recognized in (21).
Verification of (T2) of Theorem 4.2. Let \( f \) be a function in \( D(H) = C^2(\mathbb{T}^d) \) and \( x \in \mathbb{T}^d \). We establish the existence of a vector function \( \varphi = (\varphi^1, \ldots, \varphi^d) \in (C^2(\mathbb{T}^d))^d \) such that for all \((y, i) \in E = \mathbb{T}^d \times \{1, \ldots, J\}\) and some constant \( \mathcal{H}(\nabla f(x)) \in \mathbb{R} \), we have

\[
H_p(\nabla f(x), y, i) = \mathcal{H}(\nabla f(x)).
\]

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 \( \varphi_p \) such that \( H_{\varphi_p}(p, y, i) = \mathcal{H}(p) \) 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) with the following Lemma.

**Lemma 7.1.** Let \( E' = \mathbb{T}^d \times \{1, \ldots, J\} \) and \( H \) be the limit operator (20). Then:

(a) For \( f \in D(H) \), the limiting images \( H_p(\nabla f(x), y, i) \) are of the form

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

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

\[
(B_p h)(y, i) := \frac{1}{2} \Delta_h y, i + \left( p + b'(y) \right) \cdot \nabla y, i
\]

\[
(V_p h)(y, i) := \left( \frac{1}{2} p^2 + p \cdot b'(y) \right) h(y, i),
\]

\[
(R h)(y, i) := \sum_{j=1}^J r_{ij}(y) \left[ h(y, j) - h(y, i) \right].
\]

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

\[
(B_p + V_p + R) g_p = \mathcal{H}(p) g_p. \tag{22}
\]

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

\[
H_{\varphi_p}(p, y, i) \overset{(a)}{=} e^{-\varphi_p(y, i)} \left[ (B_p + V_p + R)e^{\varphi_p} \right](y, i)
\]

\[
= \frac{1}{g_p(y, i)} \left[ (B_p + V_p + R) g_p(y, i) \right] \overset{(b)}{=} \mathcal{H}(p).
\]

**Proof of Lemma 7.1.** Writing \( p = \nabla f(x) \), Item (a) follows directly by regrouping the terms in (21). 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 \( g^i_p \) 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 \( [-B_p - V_p - R] g_p = \lambda(p) g_p \). Therefore, \( (B_p + V_p + R) g_p = \mathcal{H}(p) g_p \) follows with the same eigenfunction \( g_p > 0 \) and the principal eigenvalue \( \mathcal{H}(p) = -\lambda(p) \). This finishes the verification of (T2). \( \square \)
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

\[
\mathcal{H}(p) = -\sup_{g > 0} \inf_{z' \in E'} \left\{ \frac{1}{g(z')} \left[ (-B_p - V_p - R)g \right] (z') \right\}
\]

\[
= \inf_{g > 0} \sup_{z' \in E'} \left\{ \frac{1}{g(z')} \left[ (B_p + V_p + R)g \right] (z') \right\}
\]

\[
= \inf_{\varphi, z' \in E'} \left\{ e^{-\varphi(z')} \left[ (B_p + V_p + R)e^\varphi \right] (z') \right\} =: \inf_{\varphi, z' \in E'} F(p, \varphi)(z'),
\]

with a map \( F \) defined by

\[
F(p, \varphi)(y, i) := \frac{1}{2} \Delta \varphi^2(y) + \frac{1}{2} |\nabla \varphi^2(y) + p|^2 + b^j(y)(\nabla \varphi^j(y) + p)
\]

\[
+ \sum_{j=1}^r r_{ij}(y) \left[ e^{\varphi^j(y)} - \varphi^j(y) - 1 \right],
\]

The map \( F \) is jointly convex in \( p \) and \( \varphi \). 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

\[
\mathcal{H}(\tau p_1 + (1 - \tau)p_2) = \inf_{\varphi \in E'} F(\tau p_1 + (1 - \tau)p_2, \varphi)
\]

\[
\leq \sup_{E'} F(\tau p_1 + (1 - \tau)p_2, \tau \varphi_1 + (1 - \tau)\varphi_2)
\]

\[
\leq \sup_{E'} [\tau F(p_1, \varphi_1) + (1 - \tau)F(p_2, \varphi_2)]
\]

\[
\leq \tau \sup_{E'} F(p_1, \varphi_1) + (1 - \tau) \sup_{E'} F(p_2, \varphi_2)
\]

\[
= \tau \mathcal{H}(p_1) + (1 - \tau)\mathcal{H}(p_2).
\]

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

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

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

\[
\lambda(\varphi) \geq e^{-\varphi(z_m)} (B_0 + R)e^{\varphi(z_m)} = \frac{1}{2} \Delta y \varphi(y_m, i_m) + \frac{1}{2} |\nabla_y \varphi(y_m, i_m)|^2 \geq 0
\]

\[
+ b^j(y_m) \cdot \nabla_y \varphi(y_m, i_m) + \sum_{j \neq i} r_{ij}(y_m) \left[ e^{\varphi(y_m, j)} - \varphi(y_m, j) - 1 \right] \geq 0.
\]

This finishes the verification of (T3), and thereby the proof of Theorem 4.6. \( \square \)
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 \(H_0\) given by (19), and verify Conditions (T1), (T2) and (T3) of Theorems 4.2 and 4.3.

**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

\[
H_n f_n(x, i) = \frac{1}{n} \Delta f(x) + \frac{1}{2} \left[ \Delta \varphi(y) + \frac{n}{\gamma(n)} \Delta \xi(y) \right] + \frac{1}{2} \nabla f(x) + \nabla \varphi(y) + \frac{n}{\gamma(n)} \nabla \xi(y) + b'(y) \left( \nabla f(x) + \nabla \varphi(y) + \frac{n}{\gamma(n)} \nabla \xi(y) \right) + \frac{1}{n} \gamma(n) \sum_{j=1}^f r_{ij}(y) \left[ \varphi^i(y) - \xi^i(y) \right].
\]

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

\[
\frac{1}{n} \gamma \sum_{j=1}^f r_{ij}(y) \left[ \varphi^i(y) - \xi^i(y) \right] = \sum_{j=1}^f r_{ij}(y) \left[ \varphi^i(y) - \xi^i(y) \right] + o_{n \to \infty}(1).
\]

Therefore, we choose again \(E' := I^d \times \{1, \ldots, f\}\) as the state space of the macroscopic variables, and use the following limit operator \(H\),

\[
H := \left\{ f, H_{f, \varphi, \xi} : f \in C^2(I^d) \text{ and } H_{f, \varphi, \xi} \in C(I^d \times E') \right\},
\]

with functions \(\varphi\) and \(\xi\) in the sets \(\varphi \in C^2(I^d)\) and \(\xi = (\xi^1, \ldots, \xi^f) \in C^2(E') \simeq (C^2(I^d))^f\). The image functions \(H_{f, \varphi, \xi} : I^d \times I^d \times \{1, \ldots, f\} \to \mathbb{R}\) are

\[
H_{f, \varphi, \xi}(x, y, i) := \frac{1}{2} \Delta \varphi(y) + \frac{1}{2} \nabla f(x) + \nabla \varphi(y) + b'(y) \left( \nabla f(x) + \nabla \varphi(y) \right) + \sum_{j=1}^f r_{ij}(y) \left[ \varphi^i(y) - \xi^i(y, i) \right].
\]

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 \(\eta_n\) and \(\eta'_n\). 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 \(\varphi_p \in C^2(I^d)\) and \(\xi_p \in C^2(E')\) such that \(H_{\varphi_p, \xi_p}(p, \cdot)\) becomes constant on \(E' = I^d \times \{1, \ldots, f\}\). To that end, we find a constant \(\mathcal{H}(p) \in \mathbb{R}\) and \(\varphi_p\) and \(\xi_p\) 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 \( E' = \mathbb{T}^d \times \{1, \ldots, J\} \) and let \( H \) be the operator (23). Then:

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

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

where \( p = \nabla f(x) \in \mathbb{R}^d \), \( B_p = \frac{1}{2} \Delta + (p + b^i(y)) \cdot \nabla \) and multiplication operator \( V_p(y) = p^2 / 2 + p \cdot b^i(y) \).

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

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

where \( B_p = \frac{1}{2} \Delta + (p + \bar{b}(y)) \cdot \nabla \), \( V_p(y) = \frac{p^2}{2} + p \cdot \bar{b}(y) \). In the operators, \( \bar{b}(y) := \sum_{i=1}^J \mu_i(i) b^i(y) \) is the average drift with respect to the stationary measure \( \mu_0 \in \mathcal{P}(\{1, \ldots, J\}) \) of the jump process with frozen jump rates \( r_i(y) \).

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

\[ [B_p + V_p] g_p = \mathcal{H}(p) g_p. \]  

(25)

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

\[ H_{\varphi_p, \xi}(p, y, i) \overset{(a)}{=} e^{-\varphi_p(y)} \left( (B_p + V_p) e^{\varphi_p} \right)(y) + \sum_{j \in J} r_{ij}(y) [\xi(y, j) - \xi(y, i)] \]

\[ = \overset{(b)}{=} e^{-\varphi_p(y)} \left( (B_p + V_p) e^\varphi \right)(y) \overset{(c)}{=} \mathcal{H}(p). \]

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

\[ H_{\varphi, \xi}(p, y, i) = \frac{1}{2} \Delta \varphi + \frac{1}{2} [p + \nabla \varphi]^2 + b^i(p + \nabla \varphi) + \sum_{j=1}^J r_{ij}(y) [\xi(y, j) - \xi(y, i)], \]

\[ = e^{-\varphi} (B_{p,j} + V_{p,j}) e^\varphi =: R_y \xi_y(i) \]

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 \( \xi_y(\cdot) = \xi_y(y, \cdot) \in C(\{1, \ldots, J\}) \) such that

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

becomes constant in \( i = 1, \ldots, J \). By the Fredholm alternative, for any vector \( h \in C(\{1, \ldots, J\}) \), the equation \( R_y \xi_y = h \) has a solution \( \xi_y(\cdot) \in C(\{1, \ldots, J\}) \) if and only if \( h \perp \ker(R_y^*) \). Since \( R_y \) is the generator of a jump process on
the finite discrete set \( \{1, \ldots, J \} \) with rates \( r_{ij}(y) \), the null space \( \ker(R_y) \) is one-dimensional and spanned by the unique stationary measure \( \mu_y \in \mathcal{P}(\{1, \ldots, J\}) \), which exists by our irreducibility assumption of Theorem 4.7 (e.g. [Kle13, Theorem 17.51]). Hence \( e^{-\varphi} [B_{p,i} + V_{p,i}] e^\varphi + R_y \xi_y(i) = h(p, y) \) is independent of \( i \in \{1, \ldots, J\} \) if and only if

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

This solvability condition leads to

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

Hence for \( h(p, y) := e^{-\varphi(y)} (B_p + V_p) e^\varphi(y) \), 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 \( \xi_y \) depend smoothly on \( y \) as well, and (b) follows.

For proving (c) in Lemma 7.2, we note that Equation (25) 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 (25).

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

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

with \( F \) jointly convex in \( p \) and \( \varphi \). 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

In Theorem 4.8, we proved 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 \( E' = \mathbb{T}^d \times \{1, \ldots, J\} \) 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 (22) described in Lemma 7.1, and satisfies

\[
\mathcal{H}(p) = \sup_{\mu \in \mathcal{P}(E')} \left[ \int_{E'} V_p(z) \, d\mu(z) - I_p(\mu) \right]. \tag{26}
\]

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

\[
V_p(x, i) := \frac{1}{2} p^2 - p \cdot \nabla \psi^i(x), \tag{27}
\]
and the Donsker-Varadhan functional

\[ I_p(\mu) := - \inf_{u > 0} \int_{\mathcal{E}'} \frac{L_p u}{u} \, d\mu, \]  

(28)

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

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

(29)

The variational representation (26) is a special case of Donsker’s and Varadhan’s representation theorem on principal eigenvalues [DV75]. 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 [Pin85], Equation (3.1) in [Pin07]).

Since it is not clear from (26) that \( \mathcal{H}(p) \) is symmetric under the detailed-balance condition, we shall perform a suitable shift in the infimum of the functional (28) to obtain a suitable representation. Rewriting in (28) the strictly positive functions as \( u = \exp(\varphi) \), we find

\[ I_p(\mu) = - \inf_{\varphi} \sum_i \int_{\mathcal{E}'} \left[ \frac{1}{2} \Delta \varphi_i + \frac{1}{2} \nabla \varphi_i^2 + (p - \nabla \psi_i) \nabla \varphi_i + \sum_j r_{ij}(e^{\varphi_i - \varphi_j} - 1) \right] \, d\mu_i. \]

Suppose that \( d\mu_i = \pi_i \, dx \) with strictly positive \( \pi_i \), where \( dx \) is the Lebesgue measure on the torus. Then shifting in the infimum as \( \varphi_i \to \varphi_i + \psi_i + \frac{1}{2} \log \pi_i \), we find by calculation that

\[ I_p(\mu) = \mathcal{R}(\mu) + \int_{\mathcal{E}'} V_p \, d\mu - K_p(\mu), \]

(30)

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

\[ \mathcal{R}(\mu) := \frac{1}{8} \sum_i \int_{\mathbb{T}^d} \left| \nabla \left( \log \frac{\pi_i}{e^{-2\psi_i}} \right) \right|^2 \, d\mu_i, \]

(31)

and \( K_p(\mu) \) is given by

\[ K_p(\mu) = \inf_{\varphi} \left\{ \sum_{i=1}^I \int_{\mathbb{T}^d} \left( \frac{1}{2} |\nabla \varphi_i(x) + p|^2 - \sum_{j=1}^I r_{ij}(x) \right) \, d\mu_i(x) \right. \]

\[ \left. + \sum_{i,j=1}^I \int_{\mathbb{T}^d} r_{ij}(x)e^{-2\psi_i(x)} \sqrt{\pi_i(x)\pi_j(x)}e^{\varphi_i(x) + \varphi_j(x) - \varphi_i(x)} \, dx \right\}. \]

(32)

Plugging formula (30) into the variational representation (26) leads to the desired representation of the Hamiltonian. The transformation we used is equivalent to shifting by \( (1/2) \log(\pi_i/\pi_j) \), where \( \pi_j = e^{-2\psi_i} \) is the stationary measure up to multiplicative constant. This transformation is reminiscent of a symmetrization discussed in Touchette’s notes [Tou18, Eq. (36)]. Also when formulating the detailed-balance condition with additional constants in (8), 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 $\mu$ 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 [Str12, 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 $I_{\text{rev}}$ corresponding to a reversible process in order to be able to apply [Str12, Theorem 7.44].

**Proposition 8.1.** The Hamiltonian $\mathcal{H}(p)$ given by (26) satisfies the following:

(a) The supremum in (26) can be taken over a smaller set $\mathcal{P}$ of measures, that is

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

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

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

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

(b) We have

$$\mathcal{H}(p) = \sup_{\mu \in \mathcal{P}} \left[ K_p(\mu) - \mathcal{R}(\mu) \right], \quad (33)$$

with the maps $\mathcal{R}$ and $K_p$ given by (31) and (32) above. In $K_p(\mu)$, the infimum can be taken over vectors of functions $\phi_i = \phi(\cdot, i)$ such that $\nabla \phi_i \in L^2_{\mu_i}(\mathbb{T}^d)$.

(c) Under the detailed balance condition,

$$K_p(\mu) = \inf_{\phi} \left\{ \sum_{i=1}^J \int_{\mathbb{T}^d} \left( \frac{1}{2} |\nabla \phi_i(x)|^2 + p - \sum_{j=1}^J r_{ij}(x) \right) d\mu_i(x) \right.\right.$$ \n
$$\left. + \sum_{i,j=1}^J \int_{\mathbb{T}^d} r_{ij}(x)e^{-2\phi_i(x)}(\sqrt{\mu_i(x)}E\sqrt{\mu_j(x)})e^{\phi_j(x)+\phi_i(x)} \cosh(\phi(x,j) - \phi(x,i)) \, dx \right\}. \quad (34)$$

The representation (34) follows from (32) by rewriting the sums appearing therein as $\sum_{i} a_{ij} = \frac{1}{2} \sum_{i} (a_{ij} + a_{ji})$, where

$$a_{ij} = \int_{\mathbb{T}^d} r_{ij}(x)e^{-2\phi_i(x)}(\sqrt{\mu_i(x)}E\sqrt{\mu_j(x)})e^{\phi_j(x)+\phi_i(x)} \cosh(\phi(x,j) - \phi(x,i)) \, dx.$$  

This leads to the $\cosh(\cdot)$ terms in (34), 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 (26) can be taken over measures $\mu$ 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 $\mu$ must satisfy (P1)
and (P2). To that end, define the map $L_{\text{rev}} : \mathcal{D}(L_{\text{rev}}) \subseteq C(E') \to C(E')$ by setting $\mathcal{D}(L_{\text{rev}}) := C^2(E')$ and

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

with jump rates $s_{ij}$ defined as $s_{ij} \equiv 1$ and $s_{ij} \equiv e^{2\psi_j - 2\psi_i}$, for $i \leq j$, and with $\tau := \sup_{x \in \mathbb{R}^d} (r_{ij}/s_{ij}) < \infty$, where $r_{ij}(\cdot)$ are the jump rates appearing in $L_{\mu}$.

Furthermore, define $I_{\text{rev}} : \mathcal{P}(E') \to [0, \infty]$ by

$$I_{\text{rev}}(\mu) := - \inf_{\vartheta \in C(E')} \int_{E'} e^{-\vartheta} L_{\text{rev}}(e^{\vartheta}) \, d\mu.$$

We shall prove two statements:

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

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

The two statements combined finish the proof.

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

$$\nu_{\text{rev}}(A_1, \ldots, A_l) = \frac{1}{Z} \sum_{i=1}^l v_i^{\text{rev}}(A_i), \quad \text{where } d\nu_{\text{rev}} = e^{-2\psi_i} \, dx$$

and $Z = \sum_i v_i^{\text{rev}}(E^d)$.

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

$$(L_{\text{rev}} f, g)_{\nu_{\text{rev}}} = (f, L_{\text{rev}} g)_{\nu_{\text{rev}}}, \quad \text{where } (f, h)_{\nu_{\text{rev}}} = \frac{1}{Z} \sum_i \int_{E^d} f_i(x) h_i(x) \, d\nu_i^{\text{rev}}(x).$$

By Stroock’s result [Str12, Theorem 7.44],

$$I_{\text{rev}}(\mu) = \left\{ \begin{array}{ll} - (f, L_{\text{rev}} f)_{\nu_{\text{rev}}} & f_{\mu} = \sqrt{\mu'} \in D^{1/2} := \mathcal{D}(\sqrt{-L_{\text{rev}}}) \quad \text{and } g_{\mu} = \frac{d\mu}{d\nu_{\text{rev}}}, \\ + \infty, & \text{otherwise,} \end{array} \right.$$  

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

$$I_{\text{rev}}(\mu) = - (f, L_{\text{rev}} f)_{\nu_{\text{rev}}}$$

$$= \frac{1}{Z} \sum_{i=1}^l \left[ \int_{E^d} |\nabla f_i(x)|^2 \, d\nu_i^{\text{rev}}(x) + \tau \sum_{j=1}^l \int_{E^d} s_{ij}(x) |f_j(x) - f_i(x)|^2 \, d\nu_i^{\text{rev}}(x) \right],$$

(35)

where we write $f^i = (d\mu^i / d\nu_i^{\text{rev}})^{1/2}$. Furthermore, $\mu^i$ is absolutely continuous with respect to $\nu^i = e^{-2\psi^i} \, dx$. Since $e^{-2\psi^i} \, dx \ll dx$, we find that $\mu^i$ is absolutely continuous with respect to the volume measure on $E^d$. Hence (P1) holds true.

We verify (P2) by showing that the integral $\int_{E^d} |\nabla (\log \theta^i)|^2 \, d\mu^i$ is finite. Let $s_{\theta^i} := d\mu^i / d\nu_i^{\text{rev}}$ be the density of $\mu^i$ with respect to $\nu_i^{\text{rev}}$. Then the densities $\theta^i = d\mu^i / dx$ satisfy $s_{\theta^i} = \theta^i e^{2\psi^i}$, because

$$\theta^i = d\mu^i / d\nu_i^{\text{rev}} \frac{d\nu_i^{\text{rev}}}{dx} \frac{d\mu^i}{d\nu_i^{\text{rev}}} e^{-2\psi^i}.$$
Let \( f^i_i := \sqrt{\gamma_i} \). By (35), \( \int_{\mathbb{T}^d} |\nabla f^i_i|^2 \mathrm{d}\nu_{\text{rev}} \) is finite for every \( i = 1, \ldots, J \). Hence with the estimate

\[
\int_{\mathbb{T}^d} |\nabla f^i_i|^2 \mathrm{d}\nu_{\text{rev}} \geq \int_{\mathbb{T}^d} |\nabla f^i_i|^2 1_{\{\nu > 0\}} \mathrm{d}\nu_{\text{rev}} = \frac{1}{4} \int_{\mathbb{T}^d} \frac{|\nabla g_i|^2}{\gamma_i} - 1_{\{\nu > 0\}} \mathrm{d}\nu_{\text{rev}},
\]

\[
= \frac{1}{4} \int_{\mathbb{T}^d} \frac{|\nabla f_i|^2 + 2\mu_i \nabla \varphi_i \nabla \varphi_i|^2}{\mu_i} e^{-4\nu_i} 1_{\{\nu > 0\}} \mathrm{d}x
\]

\[
= \frac{1}{4} \int_{\mathbb{T}^d} |\nabla (\log \mu_i)|^2 + 2|\nabla \varphi_i|^2 1_{\{\nu > 0\}} \mathrm{d}\mu_i
\]

\[
\geq \frac{1}{8} \int_{\mathbb{T}^d} |\nabla (\log \mu_i)|^2 1_{\{\nu > 0\}} \mathrm{d}\mu_i - \int_{\mathbb{T}^d} |\nabla \varphi_i|^2 1_{\{\nu > 0\}} \mathrm{d}\mu_i,
\]

we find \( |\nabla (\log \mu_i)| \in L^2(\mathbb{T}^d) \).

Regarding (II), suppose that \( I_p(\mu) \) is finite. We estimate \( r_{ij}/s_{ij} \) from above by \( \mathcal{T} = \sup_{\mathbb{T}^d} (r_{ij}/s_{ij}) \) to find

\[
I_p(\mu) \geq \sup_{\varphi} \sum_i \int_{\mathbb{T}^d} \left[ \frac{1}{2} \Delta \varphi(x) + \frac{1}{2} |\nabla \varphi(x)|^2 + \left( \mu - \nabla \varphi(x) \nabla \varphi(x) \right) + \mathcal{T} \sum_{j \neq i} s_{ij}(x) (e^{\varphi(x,j) - \varphi(x,i)} - 1) \right] \mathrm{d}\mu_i - s_0(\mu),
\]

where \( s_0(\mu) = \sum_{ij} \int_{\mathbb{T}^d} \left( \mathcal{T} s_{ij}(x) - r_{ij}(x) \right) \mathrm{d}\mu_i \) is finite. For \( p = 0 \), this means that \( I_0(\mu) \geq I_{\text{rev}}(\mu) - s_0(\mu), \) and hence that \( I_{\text{rev}}(\mu) \) is finite.

For \( p \neq 0 \), the additional \( p \)-term can be dealt with by Young’s inequality applied as \(-p \cdot \nabla \varphi \geq -p \epsilon^2/(2\epsilon) - \frac{p^2}{2\epsilon} \nabla \varphi_i|^2 \). Thereby,

\[
I_p(\mu) \geq \sup_{\varphi} \sum_i \int_{\mathbb{T}^d} \left[ \frac{1}{2} \Delta \varphi(x) + \frac{1}{2} |\nabla \varphi(x)|^2 + \left( \mu - \nabla \varphi(x) \nabla \varphi(x) \right) + \mathcal{T} \sum_{j \neq i} s_{ij}(x) (e^{\varphi(x,j) - \varphi(x,i)} - 1) \right] \mathrm{d}\mu_i - \frac{p^2}{2\epsilon} - s_0(\mu)
\]

\[
= \frac{1}{\lambda} \sup_{\varphi} \sum_i \int_{\mathbb{T}^d} \left[ \frac{1}{2} \Delta \varphi(x) + \frac{1}{2} |\nabla \varphi(x)|^2 + \left( \mu - \nabla \varphi(x) \nabla \varphi(x) \right) + \mathcal{T} \sum_{j \neq i} s_{ij}(x) (e^{\varphi(x,j) - \varphi(x,i)} - 1) \right] \mathrm{d}\mu_i - s_0(\mu),
\]

where the last equality follows by rescaling \( \varphi \to \varphi/\lambda \), with \( \lambda = 1 + \epsilon > 1 \). Therefore, apart from the factor \( 1/\lambda \) in the exponential term and the multiplicative factor \( \lambda \mathcal{T} \), we obtain the same estimate as above in the \( p = 0 \) case. Denoting the supremum term in the last line by \( I^\lambda_{\text{rev}} \), we found the estimate

\[
I_p(\mu) \geq \frac{1}{\lambda} I^\lambda_{\text{rev}}(\mu) - s_p(\mu),
\]

where \( s_p(\mu) = (2\epsilon)^{-1} p^2 + s_0(\mu) \). Hence \( I^\lambda_{\text{rev}}(\mu) \) is finite. To show that this enforces finiteness of \( I_{\text{rev}}(\mu) \), we prove that \( I_{\text{rev}}(\mu) = \infty \) implies \( I^\lambda_{\text{rev}}(\mu) = \infty \).
If $I_{\text{inv}}(\mu) = \infty$, then by definition there exist functions $\varphi_n$ such that

$$a(\varphi_n) := -\sum_{i=1}^l \int_{\mathcal{O}} \left[ \frac{1}{2} \Delta \varphi_n^i + \frac{1}{2} |\nabla \varphi_n^i|^2 - \nabla \varphi_n^i \cdot \nabla \varphi_n^i \right. $$

$$+ \frac{\gamma}{n} \sum_{j \neq i} s_{ij} \left( e^{\varphi_n(x_j)} - \varphi_n(x_j) - 1 \right) \right] d\mu^i(x)$$

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

$$a^\lambda(\varphi_n) := -\sum_{i=1}^l \int_{\mathcal{O}} \left[ \frac{1}{2} \Delta \varphi_n^i + \frac{1}{2} |\nabla \varphi_n^i|^2 - \nabla \varphi_n^i \cdot \nabla \varphi_n^i \right. $$

$$+ \lambda \frac{\gamma}{n} \sum_{j \neq i} s_{ij} \left( e^{(\varphi_n(x_j) - \varphi_n(x_i))} - 1 \right) \right] d\mu^i(x)$$

for the according evaluation of $\varphi_n$ in $I_{\text{inv}}^\lambda(\mu)$. By definition, $I_{\text{inv}}^\lambda(\mu) \geq a^\lambda(\varphi_n)$.

We show that with $C = C(\mu) > 0$ defined by

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

we have

$$a^\lambda(\varphi_n) \geq a(\varphi_n) - C. \quad (37)$$

To that end, define the sequences $\overline{a}_n$ and $\overline{a}^\lambda_n$ by

$$\overline{a}_n := -\sum_{i=1}^l \int_{\mathcal{O}} \left[ \frac{1}{2} \Delta \varphi_n^i + \frac{1}{2} |\nabla \varphi_n^i|^2 - \nabla \varphi_n^i \cdot \nabla \varphi_n^i \right. $$

$$+ \frac{\gamma}{n} \sum_{j \neq i} s_{ij} \left( e^{\varphi_n(x_j)} - \varphi_n(x_j) - 1 \right) \right] d\mu^i(x),$$

and

$$\overline{a}^\lambda_n := -\sum_{i=1}^l \int_{\mathcal{O}} \left[ \frac{1}{2} \Delta \varphi_n^i + \frac{1}{2} |\nabla \varphi_n^i|^2 - \nabla \varphi_n^i \cdot \nabla \varphi_n^i \right. $$

$$+ \lambda \frac{\gamma}{n} \sum_{j \neq i} s_{ij} \left( e^{(\varphi_n(x_j) - \varphi_n(x_i))} - 1 \right) \right] d\mu^i(x).$$

By the elementary estimates $e^x \geq e^x 1_{\{x \geq 0\}}$ and $(e^x 1_{\{x \geq 0\}} - e^x) \geq -1$, we obtain the inequalities

$$\overline{a}_n \geq a(\varphi_n) \quad \text{and} \quad a^\lambda(\varphi_n) \geq \overline{a}^\lambda_n - C.$$

Furthermore, the bound $\overline{a}_n \leq \overline{a}^\lambda_n$ is obtained by noting that

$$\overline{a}_n - \overline{a}^\lambda_n = -\sum_{i=1}^l \int_{\mathcal{O}} \frac{\gamma}{n} \sum_{j \neq i} s_{ij} \left( e^{\varphi_n(x_j)} - \varphi_n(x_j) - 1 \right) d\mu^i$$

$$+ \lambda \frac{\gamma}{n} \sum_{i=1}^l \int_{\mathcal{O}} s_{ij} \left( e^{\varphi_n(x_j)} - 1 \right) d\mu^i$$

$$= \gamma(1 - \lambda) \sum_{ij} s_{ij} d\mu^i + \lambda \sum_{ij} s_{ij} \left( e^{\varphi_n(x_j)} - 1 \right) \int_{\mathcal{O}} e^{\varphi_n(x_j)} d\mu^i,$$
which is bounded above by zero since \( \lambda = 1 + \varepsilon > 1 \) and \( e^{x/\lambda} \leq e^x \) for \( x \geq 0 \).

In conclusion, we have

\[
\alpha \phi_n - C \geq \alpha_n - C \geq \alpha(\phi_n) - C,
\]

finishing the proof. \( \square \)

**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 (30). Integration by parts gives

\[
I_p(\mu) = \inf_{\phi} \sum_i \int_{\mathbb{T}^d} \left[ -\frac{1}{2} \nabla \phi_i \nabla (\log \mu_i) + \frac{1}{2} |\nabla \phi_i|^2 + (p - \nabla \psi_i) \nabla \phi_i 
+ \sum_j r_{ij} \left( e^{\psi_j - \phi_i} - 1 \right) \right] d\mu_i,
\]

where \( d\mu_i = \mu_i dx \). By a density argument, the infimum can be taken over functions \( \phi \) such that \( \nabla \phi_i \in L^2_{\mu_i}(\mathbb{T}^d) \). Now shifting in the infimum as \( \phi_i \rightarrow \phi_i + \frac{1}{2} \log(\mu_i) + \psi_i \), we find after some algebra that

\[
I_p(\mu) = \inf_{\phi} \sum_i \int_{\mathbb{T}^d} \left[ \frac{1}{2} |\nabla \phi_i + \rho_i|^2 - \frac{1}{2} (p - \nabla \psi_i) - \frac{1}{2} \nabla \log \mu_i|^2 
+ \sum_j r_{ij} \left( \sqrt{\frac{\mu_i}{\mu_j}} e^{\psi_j - \phi_i} - 1 \right) \right] d\mu_i.
\]

The term containing the square roots and logarithms are not singular since they are integrated against \( d\mu_i \), so that the integration is over the set \( \{ \mu_i > 0 \} \). Now writing out the terms and reorganizing them leads to the claimed equality. \( \square \)

### 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. We provide a sketch of proof here since we know of no reference in the literature. Let \( I : \mathcal{X} \rightarrow [0, \infty] \) be a rate function. We denote by \( \{ I = 0 \} \) 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 \( I \). Then \( d(X^n, \{ I = 0 \}) \rightarrow 0 \) almost surely as \( n \rightarrow \infty \).

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 \( A^n_k \) be the event
\[
A^n_k := \{ d(X^n, \{ I = 0 \}) \geq 1/k \},
\]
and write
\[
A^n_k \text{ i.o.} := \bigcap_{N \geq 1} \bigcup_{n \geq N} A^n_k.
\]
Let \( k \in \mathbb{N} \). By the large-deviation upper bound, there exists a \( \delta > 0 \) such that for all \( n \) sufficiently large,
\[
P(A^n_k) \leq e^{-n\delta}.
\]
Therefore \( \sum_{n=1}^{\infty} P(A^n_k) \) is finite, and by the Borel-Cantelli Lemma,
\[
P(A^n_k \text{ i.o.}) = 0.
\]
With that, almost-sure convergence follows by noting that
\[
P \left( \{ d(X^n, \{ I = 0 \}) \nrightarrow 0 \} \right) \leq \sum_{k=1}^{\infty} P(A^n_k \text{ i.o.}) = 0.
\]

\[\square\]

B Principal eigenvalues

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

Proposition B.1. Let \( P \) be a second-order uniformly elliptic operator given by
\[
P = -\sum_{k\ell} a_{k\ell}(\cdot) \frac{\partial^2}{\partial x^k \partial x^\ell} + \sum_k b_k(\cdot) \frac{\partial}{\partial x^k} + c(\cdot),
\]
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 \( \lambda \) is given by
\[
\lambda = \sup_{g > 0} \inf_{x \in \mathbb{T}^d} \left[ \frac{Pg(x)}{g(x)} \right] = \inf_{\mu \in \mathcal{P}(\mathbb{T}^d)} \sup_{g > 0} \left[ \int_{\mathbb{T}^d} \frac{Pg}{g} \, d\mu \right].
\]

Proposition B.2. Let \( L : C^2(\mathbb{T}^d) \to C(\mathbb{T}^d) \) be a \( J \times J \) diagonal matrix of uniformly elliptic operators,
\[
L = \begin{pmatrix} L^{(1)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & L^{(J)} \end{pmatrix}, \quad L^{(i)} = -\sum_{k \ell} a_{k\ell}^{(i)}(\cdot) \frac{\partial^2}{\partial x^k \partial x^\ell} + \sum_k b_k^{(i)}(\cdot) \frac{\partial}{\partial x^k} + c^{(i)}(\cdot),
\]

(39)
with \( a^{(i)}(\cdot), b^{(i)}(\cdot), c^{(i)}(\cdot) \in C^\infty(T^d)\), and let \( R \) be a \( J \times J \) matrix with non-negative functions on the off-diagonal,

\[
R = \begin{pmatrix}
R_{11} & \geq 0 \\
\vdots & \ddots & \geq 0 \\
0 & \cdots & R_{jj}
\end{pmatrix}, \quad R_{ij} \geq 0 \text{ for all } i \neq j.
\]

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

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

In the above propositions, the eigenvalue \( \lambda \) is referred to as the principal eigenvalue. The principal-eigenvalue problem on closed manifolds, such as the torus \( T^d \), is solved for instance by Padilla [Pad97]. Donsker and Varadhan’s variational representations for principal eigenvalues, [DV75, DV76], apply to the case of compact metric spaces without boundary. A proof of how to obtain the principal eigenvalue for coupled systems of equations is given by Sweers [Swe92] and Kifer [Kif92]. Sweers considers a Dirichlet boundary problem, but his results transfer to the compact setting without boundary. Kifer gives an independent proof for the case of a compact manifold, in Lemma 2.1 and Proposition 2.2 in [Kif92].

**Acknowledgement**

The authors thank Frank Redig, Francesca Collet and Federico Sau for their remarks and suggestions during a couple of meetings. 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.

**References**

[BD19] A. Budhiraja and P. Dupuis. *Analysis and Approximation of Rare Events: Representations and Weak Convergence Methods*, volume 94 of Probability Theory and Stochastic Modelling. Springer, 2019.

[BDG18] A. Budhiraja, P. Dupuis, and A. Ganguly. Large deviations for small noise diffusions in a fast markovian environment. *Electronic Journal of Probability*, 23, 2018.

[Dei92] K. Deimling. *Multivalued Differential Equations*, volume 1 of De Gruyter Series in Nonlinear Analysis and Applications. Walter de Gruyter & Co., Berlin, 1992.
[Kle13] A. Klenke. *Probability Theory: A Comprehensive Course*. Springer Science & Business Media, 2013.

[Kol13] A. B. Kolomeisky. Motor proteins and molecular motors: How to operate machines at the nanoscale. *Journal of Physics: Condensed Matter*, 25(46):463101, 2013.

[KP17] R. Kumar and L. Popovic. Large deviations for multi-scale jump-diffusion processes. *Stochastic Processes and their Applications*, 127(4):1297–1320, 2017.

[Kra16] 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(2):321–345, 2016.

[KS20] R. C. Kraaij and M. C. Schlottke. A large deviation principle for Markovian slow-fast systems. *arXiv preprint arXiv:2011.05686*, 2020.

[MS13] S. Mirrahimi and P. E. Souganidis. A homogenization approach for the motion of motor proteins. *Nonlinear Differential Equations and Applications NoDEA*, 20(1):129–147, 2013.

[Pad97] P. Padilla. The principal eigenvalue and maximum principle for second order elliptic operators on Riemannian manifolds. *Journal of Mathematical Analysis and Applications*, 205(2):285–312, 1997.

[Pin85] R. G. Pinsky. On evaluating the Donsker-Varadhan I-function. *The Annals of Probability*, pages 342–362, 1985.

[Pin07] R. G. Pinsky. Regularity properties of the Donsker–Varadhan rate functional for non-reversible diffusions and random evolutions. *Stochastics and Dynamics*, 7(02):123–140, 2007.

[PS09a] B. Perthame and P. E. Souganidis. Asymmetric potentials and motor effect: A homogenization approach. *Annales de l’Institut Henri Poincare (C) Non Linear Analysis*, 26(6):2055–2071, 2009.

[PS09b] B. Perthame and P. E. Souganidis. Asymmetric potentials and motor effect: A large deviation approach. *Archive for Rational Mechanics and Analysis*, 193(1):153–169, 2009.

[Roc66] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510, 1966.

[Str12] D. W. Stroock. *An Introduction to the Theory of Large Deviations*. Springer Science & Business Media, 2012.

[Swe92] G. Sweers. Strong positivity in $C(\bar{\Omega})$ for elliptic systems. *Mathematische Zeitschrift*, 209(1):251, 1992.

[Tou18] H. Touchette. Introduction to dynamical large deviations of Markov processes. *Physica A: Statistical Mechanics and its Applications*, 504:5–19, 2018.

[Vor11] D. Vorotnikov. The flashing ratchet and unidirectional transport of matter. *Discrete & Continuous Dynamical Systems-B*, 16(3):963, 2011.
| Reference | Author(s) | Title | Journal | Volume/Issue/Range |
|-----------|-----------|-------|---------|--------------------|
| [Vor14]   | D. Vorotnikov | Analytical aspects of the Brownian motor effect in randomly flashing ratchets. | Journal of mathematical biology | 68(7):1677–1705, 2014. |
| [WPE03]   | H. Wang, C. S. Peskin, and T. C. Elston | A robust numerical algorithm for studying biomolecular transport processes. | Journal of theoretical biology | 221(4):491–511, 2003. |
| [YZ10]    | G. Yin and C. Zhu | Hybrid Switching Diffusions: Properties and Applications. | Springer New York | 2010. |