Electron. J. Probab. **23** (2018), no. 112, 1–33.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP228

Large deviations for small noise diffusions in a fast markovian environment

Amarjit Budhiraja[*†] Paul Dupuis[‡§] Arnab Ganguly[¶‖]

Abstract

A large deviation principle is established for a two-scale stochastic system in which the slow component is a continuous process given by a small noise finite dimensional Itô stochastic differential equation, and the fast component is a finite state pure jump process. Previous works have considered settings where the coupling between the components is weak in a certain sense. In the current work we study a fully coupled system in which the drift and diffusion coefficient of the slow component and the jump intensity function and jump distribution of the fast process depend on the states of both components. In addition, the diffusion can be degenerate. Our proofs use certain stochastic control representations for expectations of exponential functionals of finite dimensional Brownian motions and Poisson random measures together with weak convergence arguments. A key challenge is in the proof of the large deviation lower bound where, due to the interplay between the degeneracy of the diffusion and the full dependence of the coefficients on the two components, the associated local rate function has poor regularity properties.

Keywords: large deviations; variational representations; stochastic averaging; averaging principle; small noise asymptotics; multi-scale analysis; switching diffusions; Markov modulated diffusions; Poisson random measures.
AMS MSC 2010: 60F10; 60J75; 60G35; 60K37.
Submitted to EJP on January 25, 2018, final version accepted on October 1, 2018.

[*]University of North Carolina Chapel Hill, Department of Statistics and Operations Research.
E-mail: budhiraj@email.unc.edu
[†]Research supported in part by the National Science Foundation (DMS-1305120), the Army Research Office (W911NF-14-1-0331) and DARPA (W911NF-15-2-0122).
[‡]Brown University, Division of Applied Mathematics. E-mail: Paul_Dupuis@brown.edu
[§]Research supported in part by the National Science Foundation (DMS-1317199), the Army Research Office (W911NF-12-1-0222), and the Defense Advanced Research Projects Agency (W911NF-15-2-0122).
[¶]Louisiana State University, Department of Mathematics. E-mail: aganguly@lsu.edu
[‖]Research supported in part by Louisiana Board of Regents through the Board of Regents Support Fund (contract number: LEQSF(2016-19)-RD-A-04).

1 Introduction

We study a stochastic system with two time scales where the slow scale evolution is described through a continuous stochastic process, given by a small noise finite dimensional Itô stochastic differential equation, and the fast component is given as a rapidly oscillating pure jump process. The two processes are fully coupled in that the drift and diffusion coefficient of the slow process and the jump intensity function and jump distribution of the fast process depend on the states of both components. Multiscale systems of the form considered in this work arise in many problems from systems biology, financial engineering, queuing systems, etc. For example, most cellular processes are inherently multiscale in nature with reactions occurring at varying speeds. This is especially true in many genetic networks, where protein concentration, usually modeled by a small-noise diffusion process, is controlled by different genes rapidly switching between their respective active and inactive states [9]. The key characterizing feature of such slow-fast systems is that the fast component reaches its equilibrium state at much shorter time scales at which the slow system effectively remains unchanged. This local equilibration phenomenon allows the approximation of the properties of the slow system by averaging out the coefficients over the local stationary distributions of the fast component. Such approximations yield a significant model simplification and are mathematically justified by establishing an appropriate *averaging principle*.

The averaging principle, which has its roots in the works of Laplace and Lagrange, has a long history of applications in celestial mechanics, oscillation theory, radiophysics, etc. For deterministic systems, the first rigorous results were obtained by Bogoliubov and Mitropolsky [3], and further developments and generalizations were subsequently carried out by Volosov, Anosov, Neishtadt, Arnold and others (for example, see [1, 24]). The stochastic version of the theory originated with the seminal paper of Khasminskii [14] and later advanced in the works of Freidlin, Lipster, Skorohod, Veretennikov, Wentzel and others (for example, see [13, 26, 27]). Stochastic averaging principles for various models arising from systems biology have been studied in [2, 18, 19]. As noted above, an averaging principle provides a model simplification in an appropriate scaling regime. In order to capture the approximation errors due to the use of such simplified models one needs a more precise asymptotic analysis. The goal of the current work is to study one such asymptotic result that gives a *large deviation principle* (LDP) for the slow process as the parameter governing the magnitude of the small noise in the diffusion component and the speed of the fast component approaches its limit. Such a result, in addition to providing estimates on the rate of convergence of the trajectories of the slow component to that of the averaged system, is a starting point for developing accelerated Monte-Carlo schemes for the estimation of probabilities of rare events (cf. [11]).

For a two-scale system where both components are continuous processes given through finite dimensional Itô stochastic differential equations, the problem has been studied in [23, 13, 28, 29]. In all these works the coupling between the two components is weak in a certain sense. By this we mean that either the slow component has no diffusion term [13, 28], or the dynamics of the fast component does not depend on the slow one [23], or at least the diffusion coefficient of the fast component does not depend on the slow term [29]. A recent paper by Puhalskii [25] studies a large deviation principle for a fully coupled two-scale diffusion system. Under various conditions on the coefficients of the two diffusions, including in particular certain non-degeneracy conditions on the diffusion coefficients, the paper uses the exponential tightness and limit characterization approach of [12] to establish a LDP for the slow component. Large deviation results for certain multiscale dynamical systems given through a system of ordinary differential equations have been studied in Kifer [20, 21].

For settings where the fast component is a jump process, there are only a few results. In [15, 17] the authors study a large deviation principle for a two-scale system in which the trajectories of the slow diffusion component is modulated by a fast moving Markov chain (whose evolution does not depend on the slow component). An earlier paper, [16], considered a simpler case with no diffusion term in the equation for the slow component. This simpler case, under a somewhat more restrictive condition, was also studied by Freidlin and Wentzell in [13]. However, in all of these works the dynamics of the Markov chain do not depend on that of the slow diffusion component. Large deviation problems for general two-scale jump diffusions have recently been considered in [22]. The authors prove a large deviation principle for each fixed time $t > 0$ using the nonlinear semigroup and viscosity solution based approach developed in [12]; however, a process level large deviation result is not considered. One of the critical assumptions in this work is the validity of a comparison principle for a certain nonlinear Cauchy problem (see Theorem 3 therein). Verification of the comparison principle is in general a challenging task which needs to be done on a case by case basis for different systems. Specifically, the Hamilton-Jacobi equations obtained in the current setting through the Perron-Frobenius theory for the associated eigenvalue problems will in general have poor regularity for classical comparison results to be applicable. We also note that [22] makes the assumption that the jump coefficients are Lipschitz continuous in an appropriate sense. Such a property fails to hold for systems considered in the current paper; specifically, the integrand in the second equation in (2.3) is not Lipschitz continuous (in fact not even continuous).

As noted previously, the current paper studies a setting where the two components are fully coupled. Specifically, for fixed $\varepsilon > 0$, we consider a two component Markov process $(X^\varepsilon, Y^\varepsilon)$, where X^ε is a d-dimensional continuous stochastic process given as the solution of a stochastic equation of the form

$$dX^\varepsilon(t) = b(X^\varepsilon(t), Y^\varepsilon(t))dt + \sqrt{\varepsilon}a(X^\varepsilon(t), Y^\varepsilon(t))dW(t),$$

where W is an m-dimensional Brownian motion, and Y^ε is a process with a finite state space described in terms of a jump intensity function $c(\cdot, \cdot)$ and a probability transition kernel $r(\cdot, \cdot, dy)$, both of which depend on the states of X^ε and Y^ε. We make standard Lipschitz assumptions on the coefficients of the diffusion, however we do not impose any non-degeneracy restrictions on the diffusion. In the setting we consider methods based on approximations, exponential tightness estimates and Girsanov change of measure appear to be quite hard to implement. One of the main challenges in the analysis is due to the interplay between the possible degeneracy of the diffusion coefficient and the dependence of the various coefficients (b, a, c and r) on both components. In our approach we bypass discretizations and approximations by using certain variational representations of expectations of positive functionals of Brownian motions and Poisson random measures together with weak convergence techniques. The variational representations for these noise processes that we use were developed in [4, 6] and have been previously used in proving large deviation principles for a variety of complex systems (see [5, 7, 8] and references therein). Using these representations, the proof of the upper bound reduces to proving the tightness and characterizations of weak limit points of certain controlled versions of the state process X^ε. We note that in the description of these controlled systems there are two types of controls – one that controls the drift of the Brownian noise and the other that controls the intensity of the underlying Poisson random measure through a random 'thinning' function. The presence of these two controls coupled with the strong dependence of the coefficients on both the components make the required asymptotic analysis challenging.

The main challenge in this work arises in the proof of the lower bound. When using the variational representations, the proof of the lower bound requires the construction of

controls which lead to a prescribed limit trajectory with a prescribed cost. In particular, when multiple times scales are present, one generally needs to establish the convergence of the empirical measure for the fast variables to an a priori identified measure (which could depend on the state of the slow variables). A natural technique is to first show that the velocities of the trajectory can be made piecewise smooth (e.g., piecewise constant), so that transition probabilities associated with the fast variables can be treated as essentially constant over each interval where the velocity is continuous. Unfortunately, this smoothing in time of the state requires establishing regularity properties of the local rate function, which is the function $L(x,\beta)$ when the rate function is written in the somewhat standard form

$$I(\xi) = \int_0^T L(\xi(t), \dot{\xi}(t))dt.$$

It is the need for these regularity properties which leads to undesirable assumptions that may not in fact be necessary (e.g., nondegeneracy of a diffusion coefficient).

We will use a different method to establish convergence that does not rely on any smoothing in the time variable, and which in particular will allow for degenerate diffusion coefficients. This alternative approach, which is one of the novel contributions of this work, instead slightly perturbs the controls used on the noise space (both the control of the Brownian term that directly impacts the slow variables and the control of the Poisson term determining evolution of the fast variables), in such a way that the resulting mapping from controls into the state trajectory is unique. This uniqueness result is the key to the construction of near optimal controls for the prelimit process for which the appropriate convergence properties can be proved and from which the lower bound follows readily. The perturbation argument and resulting uniqueness, which is given in Proposition 4.1, is described in detail at the beginning of Section 5. The strategy for the proof of Proposition 4.1 is explained in Remark 4.2.

The rest of the paper is organized as follows. In Section 2 we give a precise mathematical formulation of the model and the statement of our main result. The large deviation upper bound is proved in Section 3. Section 4 constructs suitable near optimal controls and controlled trajectories with appropriate uniqueness properties. The large deviation lower bound is proved in Section 5.

Notation: The following mathematical notation and conventions will be used in the paper. For a Polish space S, we denote by $\mathcal{P}(S)$ (resp. $\mathcal{M}_F(S)$) the space of probability measures (resp. finite measures) on S equipped with the topology of weak convergence. We denote by $C_b(S)$ the space of real continuous and bounded functions on S. The space of continuous functions from $[0,T]$ to S, equipped with the uniform topology, will be denoted as $C([0,T]:S)$. For a bounded $\mathbb{R}^d$ valued function f on S, we define $\|f\|_\infty = \sup_{x\in S}\|f(x)\|$. For a finite set $\mathbb{L}$, we denote by $\mathbb{M}(\mathbb{L})$ the space of real functions on $\mathbb{L}$. Cardinality of such a set will be denoted as $|\mathbb{L}|$. Given a probability function $r : \mathbb{L} \to [0,1]$ (i.e. $\sum_{x\in\mathbb{L}} r(x) = 1$), we denote, abusing notation, $\sum_{x\in A} r(x)$ by $r(A)$ for all $A \subset \mathbb{L}$ and $\sum_{x\in\mathbb{L}} f(x)r(x)$ by $\int_{\mathbb{L}} f(x)r(dx)$ for all $f \in \mathbb{M}(\mathbb{L})$. Space of Borel measurable maps from $[0,T]$ to a metric space S will be denoted as $\mathbb{M}([0,T]:S)$. Infimum over an empty set, by convention, is taken to be ∞. In the Appendix we give a list of other notation used frequently in this work.

2 Mathematical preliminaries and main result

For fixed $\varepsilon > 0$, we consider a two component Markov process $\{(X^\varepsilon(t), Y^\varepsilon(t))\}_{0\le t\le T}$ with values in $\mathbb{R}^d \times \mathbb{L}$, where $\mathbb{L} = \{1,\dots,|\mathbb{L}|\}$ is equipped with the usual operation of addition modulo $|\mathbb{L}|$. A precise stochastic evolution equation for the pair $(X^\varepsilon, Y^\varepsilon)$ will be given below in terms of an m-dimensional Brownian motion and suitable Poisson random

measure. However, roughly speaking, the pair $(X^\varepsilon, Y^\varepsilon)$ describes a jump-diffusion, where the diffusion component (namely X^ε) has "small noise" while the jump component (Y^ε) has jumps at rate $O(\varepsilon^{-1})$. The drift and diffusion coefficients of the continuous component are given by suitable functions $b : \mathbb{R}^d \times \mathbb{L} \to \mathbb{R}^d$ and $a : \mathbb{R}^d \times \mathbb{L} \to \mathbb{R}^{d\times m}$. The evolution of the pure-jump fast component is described through a jump intensity function $c : \mathbb{R}^d \times \mathbb{L} \to [0, \infty)$ and a transition probability function $r : \mathbb{R}^d \times \mathbb{L} \times \mathbb{L} \to [0, 1]$. Our main assumptions on these functions are as follows.

Assumption 2.1.

1. *There exists* $d_{\text{lip}} \in (0, \infty)$ *such that for all* $y, y' \in \mathbb{L}$ *and* $x, x' \in \mathbb{R}^d$,

$$|c(x,y) - c(x',y)| + \|a(x,y) - a(x',y)\| + \|b(x,y) - b(x',y)\| + |r(x,y,y') - r(x',y,y')| \\ \leq d_{\text{lip}}\|x - x'\|.$$

2. c *is a bounded function.*

3. *For all* $(x, y) \in \mathbb{R}^d \times \mathbb{L}$, $\sum_{y'\in\mathbb{L}} r(x, y, y') = 1$, $r(x, y, y) = 0$.

We will occasionally write $c(x, y), a(x, y), b(x, y), r(x, y, y')$ as $c_y(x), a_y(x), b_y(x), r_{yy'}(x)$, respectively.

Remark 2.2. Assumption 2.1(1) implies that, for some $\kappa_1 \in (0, \infty)$,

$$\|b(x,y)\| + \|a(x,y)\| \leq \kappa_1(1 + \|x\|), \text{ for all } (x,y) \in \mathbb{R}^d \times \mathbb{L}.$$

Let

$$\bar{\varsigma} \doteq \sup_{(x,y)\in\mathbb{R}^d\times\mathbb{L}} c_y(x),\ \varsigma \doteq \bar{\varsigma} + 1,$$

and let $\lambda = \lambda_\varsigma$ be the Lebesgue measure on $([0, \varsigma], \mathcal{B}([0, \varsigma]))$. For $(x, y, y') \in \mathbb{R}^d \times \mathbb{L} \times \mathbb{L}$, $y \neq y'$, let

$$E_{yy'}(x) \doteq [0, c_y(x)r_{yy'}(x)]. \tag{2.1}$$

From Assumption 2.1, for some $\kappa_2 \in (0, \infty)$,

$$\sup_{(y,y')\in\mathbb{L}\times\mathbb{L}, y\neq y'} \lambda[E_{yy'}(x) \triangle E_{yy'}(x')] \leq \kappa_2\|x - x'\| \tag{2.2}$$

for all $x, x' \in \mathbb{R}^d$, where Δ denotes the symmetric difference. For each fixed $x \in \mathbb{R}^d$, the operator Π_x acting on $\mathbb{M}(\mathbb{L})$ and defined by

$$\begin{aligned}\Pi_x\phi(y) &\doteq c_y(x) \sum_{y'\in\mathbb{L}} (\phi(y') - \phi(y))r_{yy'}(x) \\ &= c_y(x) \left(\sum_{y'\in\mathbb{L}} \phi(y')r_{yy'}(x) - \phi(y) \right)\end{aligned}$$

describes the generator of an $\mathbb{L}$-valued Markov process. Let

$$\hat{r}^n_{yz}(x) \doteq \sum_{y'\in\mathbb{L}} r_{y'z}(x)\hat{r}^{n-1}_{yy'}(x),\ n > 1;\ \hat{r}^1_{yz}(x) = r_{yz}(x)$$

be the n-step transition probability kernel of the corresponding embedded chain. Define

$$\alpha \doteq \inf_x \min_{y,z\in\mathbb{L}} \sum_{n=1}^{|\mathbb{L}|} \hat{r}^n_{yz}(x), \quad \underline{\varsigma} \doteq \inf_x \min_{y\in\mathbb{L}} c(x, y)$$

Let

$$\mathbb{T} \doteq \{(y, y') \in \mathbb{L} \times \mathbb{L} : r_{yy'}(x) > 0 \text{ for some } x \in \mathbb{R}^d\}$$

and let

$$\kappa_3 \doteq \inf_{x \in \mathbb{R}^d} \min_{(y,y') \in \mathbb{T}} r_{yy'}(x).$$

Recall that, from Assumption 2.1(2) $\bar{\varsigma} \doteq \sup_x \bar{\varsigma}_x < \infty$. We will make the following additional assumption.

Assumption 2.3. *$\alpha > 0$, $\underline{\varsigma} > 0$ and $\kappa_3 > 0$.*

Assumptions 2.1 and 2.3 will be taken to hold throughout this work and will not always be mentioned in the statement of various results. Assumption 2.3 in particular says that the $|\mathbb{L}| \times |\mathbb{L}|$ adjacency matrix whose (i, j) entry is 1 if $(i, j) \in \mathbb{T}$ and 0 otherwise is irreducible.

The evolution of Y^ε can be described through a stochastic differential equation driven by a finite collection of Poisson random measures which is constructed as follows. For $(i, j) \in \mathbb{T}$ let $\bar{N}_{ij}$ be a Poisson random measure (PRM) on $[0, \zeta] \times [0, T] \times \mathbb{R}_+$ with intensity measure $\lambda_\zeta \otimes \lambda_T \otimes \lambda_\infty$, where λ_T (resp. λ_∞) denotes the Lebesgue measure on $[0, T]$ (resp. $\mathbb{R}_+$), on some complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0 \le t \le T})$ such that for $t \in [0, T]$,

$$\bar{N}_{ij}(A \times [0, t] \times B) - t\lambda_\zeta(A)\lambda_\infty(B)$$

is a $\{\mathcal{F}_t\}$-martingale for all $A \in \mathcal{B}[0, \zeta]$ and $B \in \mathcal{B}(\mathbb{R}_+)$ with $\lambda_\infty(B) < \infty$. Then

$$N_{ij}^{\varepsilon^{-1}}(dr \times dt) \doteq \bar{N}_{ij}(dr \times dt \times [0, \varepsilon^{-1}])$$

is a PRM on $[0, \zeta] \times [0, T]$ with intensity measure $\varepsilon^{-1}\lambda_\zeta \otimes \lambda_T$, and can be regarded as a random variable with values in $\mathcal{M}_F([0, \zeta] \times [0, T])$, the space of finite measures on $[0, \zeta] \times [0, T]$ equipped with the weak topology. The processes $(\bar{N}_{ij})_{(i,j) \in \mathbb{T}}$ are taken to be mutually independent. We also suppose that on this filtered probability space there is an m-dimensional $\mathcal{F}_t$-Brownian motion $W = \{W(t)\}_{0 \le t \le T}$. We will assume that for $0 \le s \le t \le T$,

$$\{W(t) - W(s), \bar{N}_{ij}(A \times (s, t] \times B)\ A \in \mathcal{B}[0, \zeta], B \in \mathcal{B}(\mathbb{R}_+),\ (i, j) \in \mathbb{T}\}$$

is independent of $\mathcal{F}_s$.

In terms of W and $N^{\varepsilon^{-1}}$, the Markov process $(X^\varepsilon, Y^\varepsilon) \equiv \{(X^\varepsilon(t), Y^\varepsilon(t))\}_{0 \le t \le T}$ with initial condition $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{L}$ is defined as the unique pathwise solution of the following system of equations:

$$\left.\begin{aligned} dX^\varepsilon(t) &= b(X^\varepsilon(t), Y^\varepsilon(t))dt + \sqrt{\varepsilon}a(X^\varepsilon(t), Y^\varepsilon(t))dW(t), && X^\varepsilon(0) = x_0 \\ dY^\varepsilon(t) &= \sum_{(i,j) \in \mathbb{T}} \int_{r \in [0,\zeta]} (j - i)1_{\{Y^\varepsilon(t-) = i\}} 1_{E_{ij}(X^\varepsilon(t))}(r) N_{ij}^{\varepsilon^{-1}}(dr \times dt), && Y^\varepsilon(0) = y_0. \end{aligned}\right\} \tag{2.3}$$

The unique solvability of (2.3) can be established by constructing the processes $(X^\varepsilon, Y^\varepsilon)$ from one jump to the next. For example, starting with $(X^\varepsilon(0), Y^\varepsilon(0)) = (x_0, y_0)$ we solve first the SDE

$$dX(t) = b(X(t), y_0)dt + \sqrt{\varepsilon}a(X(t), y_0)dW(t). \tag{2.4}$$

Using the solution $X(t)$ we determine the first jump time and jump location, (τ_1, y_1), of the fast process by the relation

$$\tau_1 \doteq \inf\left\{ t \ge 0 : \sum_{(y_0,j) \in \mathbb{T}} \int_0^t \int_{r \in [0,\zeta]} 1_{E_{y_0 j}(X(s))}(r) N_{y_0 j}^{\varepsilon^{-1}}(dr \times ds) \ne 0 \right\}$$

and $y_1 = j$ for which $\int_0^{\tau_1} \int_{r\in[0,\zeta]} 1_{E_{y_0 j}(X(s))}(r) N^{\varepsilon^{-1}}_{y_0 j}(dr \times ds) \neq 0$.

We now solve the SDE (2.4) replacing (x_0, y_0) with (x_1, y_1), where $x_1 = X(\tau_1)$. Using the new solution path $X(t)$ of the SDE the next jump-time and jump-location (τ_2, y_2) of the fast process is determined. This gives a recursive construction of a solution which is unique due to the Lipschitz properties of the coefficients a, b.

From unique pathwise solvability it follows that for every $\varepsilon > 0$, there exists a measurable map $\mathcal{G}^\varepsilon : C([0,T] : \mathbb{R}^m) \times (\mathcal{M}_F([0,\zeta] \times [0,T]))^{|\mathbb{T}|} \to C([0,T] : \mathbb{R}^d)$ such that $X^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon}W, \{\varepsilon N^{\varepsilon^{-1}}_{ij}\}_{ij})$.

The following is an immediate consequence of our assumptions.

Theorem 2.4. *For each $x \in \mathbb{R}^d$, there is a unique invariant probability measure, $\nu(x)$ for the $\mathbb{L}$-valued Markov process with generator Π_x.*

The proofs of the following two elementary lemmas are given in the Appendix.

Lemma 2.5. *The mapping $x \in \mathbb{R}^d \to \nu(x) \in \mathcal{P}(\mathbb{L})$ is Lipschitz continuous with the Lipschitz constant L^ν_{lip} (with respect to the total variation metric) depending only on $\alpha, \underline{\varsigma}, \bar{\varsigma}, \kappa_2$ and κ_3. Furthermore, $\inf_{x\in\mathbb{R}^d} \min_{y\in\mathbb{L}} \nu_y(x) \doteq \underline{\nu} > 0$.*

Lemma 2.6. *Let $f : \mathbb{R}^d \times \mathbb{L} \to \mathbb{R}$ satisfy*

$$|f(x,y) - f(x',y)| \leq L^f_{\text{lip}} \|x - x'\|, \quad x, x' \in \mathbb{R}^d, y \in \mathbb{L}$$

for some $L^f_{\text{lip}} \in (0,\infty)$. Define $\hat{f}(x) = \sum_{y\in\mathbb{L}} f(x,y)\nu_y(x)$. Then $\hat{f}$ is a locally Lipschitz function on $\mathbb{R}^d$ with linear growth.

Let $\hat{b}(x) = \sum_{y\in\mathbb{L}} b(x,y)\nu_y(x)$, and note that by Lemma 2.6 $\hat{b}$ is a locally Lipschitz function with linear growth. The proof of the following theorem follows along the lines of [26, Chapter 2, Theorem 8]. We omit the details since a similar result in a controlled setting will be shown in Proposition 3.4.

Theorem 2.7. *Fix $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{L}$. Let $(X^\varepsilon, Y^\varepsilon)$ be the solution of (2.3). Then as $\varepsilon \to 0$, X^ε converges uniformly on compacts in probability to the unique solution of*

$$\frac{d\xi(t)}{dt} = \hat{b}(\xi(t)), \quad \xi(0) = x_0. \tag{2.5}$$

The unique solvability of (2.5) is a consequence of the properties of $\hat{b}$ stated before the theorem.

The solution X^ε of the system (2.3) can be regarded as a $C([0,T] : \mathbb{R}^d)$-valued random variable. The main result of this work establishes a large deviation principle (LDP) for X^ε in $C([0,T] : \mathbb{R}^d)$ as $\varepsilon \to 0$. In rest of this section we formulate the rate function for $\{X^\varepsilon\}$ and present our main result.

2.1 Rate function

For $\psi = (\psi_j)_{j\in\mathbb{L}}$, with $\psi_j : [0,\zeta] \to \mathbb{R}_+$ a measurable map for every j, let

$$\Gamma^\psi_{ij}(x) = \begin{cases} \int_{E_{ij}(x)} \psi_j(z)\lambda_\zeta(dz), & i \neq j \\ -\sum_{y:y\neq j} \Gamma^\psi_{jy}(x) & i = j. \end{cases} \tag{2.6}$$

Note that any ψ as above, such that ψ_j is integrable for each j, can be identified with $\eta = (\eta_j)_{j\in\mathbb{L}}$ such that for each j, $\eta_j \in \mathcal{M}_F([0,\zeta])$ on setting $\eta_j(dz) \doteq \psi_j(z)dz$. More generally, for $x \in \mathbb{R}^d$, and any $\eta = (\eta_j)_{j\in\mathbb{L}}$ such that each $\eta_j \in \mathcal{M}_F([0,\zeta])$, we define $\Gamma^\eta(x)$ as

$$\Gamma^\eta_{ij}(x) = \begin{cases} \eta_j(E_{ij}(x)), & i \neq j \\ -\sum_{y:y\neq j} \eta_y(E_{jy}(x)) & i = j. \end{cases} \tag{2.7}$$

We will make use of such Γ^η in the next section. Although the introduction of a second notation for the controlled intensities is regrettable, the measure formulation is more natural when discussing topologies.

Define $\ell : [0,\infty) \to [0,\infty)$ by $\ell(x) \doteq x\log x - x + 1$ and let

$$\mathcal{R} \doteq \{\varphi = (\varphi_{ij})_{(i,j)\in\mathbb{T}} : \varphi_{ij} : [0,T]\times[0,\zeta] \to \mathbb{R}_+ \text{ is a measurable map}, (i,j) \in \mathbb{T}\}.$$

Recall that $\mathbb{M}([0,T] : \mathcal{P}(\mathbb{L}))$, $\mathbb{M}([0,T] : \mathbb{R}^d)$ denote the space of measurable maps from $[0,T]$ to $\mathcal{P}(\mathbb{L})$ and from $[0,T]$ to $\mathbb{R}^d$ respectively.

For $\xi \in C([0,T] : \mathbb{R}^d)$, define

$$I(\xi) \doteq \inf_{(u,\varphi,\pi)\in\mathcal{V}(\xi)} \left\{ \sum_{i\in\mathbb{L}} \frac{1}{2}\int_0^T \|u_i(s)\|^2 \pi_i(s)ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds \right\}, \tag{2.8}$$

where $\mathcal{V}(\xi)$ is the collection of all

$$(u = (u_i), \varphi = (\varphi_{ij}), \pi = (\pi_i)) \in \mathbb{M}([0,T] : \mathbb{R}^m)^{|\mathbb{L}|} \times \mathcal{R} \times \mathbb{M}([0,T] : \mathcal{P}(\mathbb{L}))$$

such that $\int_0^T \|u_i(s)\|^2\pi_i(s)ds < \infty$ for each $i \in \mathbb{L}$,

$$\xi(t) = x_0 + \sum_{j\in\mathbb{L}} \int_0^t b_j(\xi(s))\pi_j(s)ds + \sum_{j\in\mathbb{L}} \int_0^t a_j(\xi(s))u_j(s)\pi_j(s)ds,\ t \in [0,T], \tag{2.9}$$

and

$$\sum_{i\in\mathbb{L}} \pi_i(s)\Gamma_{ij}^{\varphi_{i,\cdot}(s,\cdot)}(\xi(s)) = 0, \text{ for a.e. } s \in [0,T] \text{ and } j \in \mathbb{L}, \tag{2.10}$$

where $\varphi_{i,\cdot} = (\varphi_{i,j})_{j\in\mathbb{L}}$ (with the convention $\varphi_{i,j} = 1$ if $(i,j) \notin \mathbb{T}$). Equation (2.10) characterizes the invariant distributions that would be associated with controlled PRMs with controls φ_{ij}, which influence the rate of transition from i to j through (2.6). This form of the rate function is very much analogous to the control formulation of a small noise diffusion as in [4]. In (2.10) we follow the convention that $0\cdot\infty = 0$ and $\infty - \infty = \infty$.

The following is the main result of this work. Recall that Assumptions 2.1 and 2.3 are taken to hold throughout the paper. A function $I : C([0,T] : \mathbb{R}^d) \to [0,\infty]$ is called a rate function on $C([0,T] : \mathbb{R}^d)$ if it has compact sub-level sets, namely for every $\alpha \in (0,\infty)$, the set $\{\xi \in C([0,T] : \mathbb{R}^d) : I(\xi) \le \alpha\}$ is a compact subset of $C([0,T] : \mathbb{R}^d)$.

Theorem 2.8. *The map I in* (2.8) *is a rate function on $C([0,T] : \mathbb{R}^d)$ and $\{X^\varepsilon\}_{\varepsilon>0}$ satisfies the Laplace principle on $C([0,T] : \mathbb{R}^d)$, as $\varepsilon \to 0$, with rate function I: for all $F \in C_b(C([0,T] : \mathbb{R}^d))$,*

$$\lim_{\varepsilon\to 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] = \inf_{\xi\in C([0,T]:\mathbb{R}^d)} \{F(\xi) + I(\xi)\}. \tag{2.11}$$

The proof of the Laplace upper bound

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \le -\inf_{\xi\in C([0,T]:\mathbb{R}^d)} \{F(\xi) + I(\xi)\}, \tag{2.12}$$

which corresponds to a variational lower bound, is given in Section 3. The corresponding lower bound

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \ge -\inf_{\xi\in C([0,T]:\mathbb{R}^d)} \{F(\xi) + I(\xi)\}, \tag{2.13}$$

which is a variational upper bound, is proven in Section 5. The fact that I is a rate function is shown in Section 2.3 (Proposition 2.14).

Remark 2.9.

(a) Note that the rate function I depends on the initial value of $(X^\varepsilon, Y^\varepsilon)$, namely (x_0, y_0). To emphasize this dependence denote I by I_{x_0,y_0}. Using a straightforward argument by contradiction, one can show that the Laplace limit (2.11), with I replaced by I_{x_0,y_0} on the right side, holds uniformly for $y_0 \in \mathbb{L}$ and for x_0 in any compact subset of $\mathbb{R}^d$.

(b) For $\varepsilon > 0$, define the $C([0,T] : \mathcal{M}_F(\mathbb{L}))$ valued random variable β^ε by

$$\beta^\varepsilon(t, A) \doteq \int_0^t 1_A(Y^\varepsilon(s))ds, \ t \in [0,T], \ A \subset \mathbb{L}.$$

Then Theorem 2.8 can be generalized as follows: The pair $(X^\varepsilon, \beta^\varepsilon)$ satisfies a large deviation principle on $C([0,T] : \mathbb{R}^d \times \mathcal{M}_F(\mathbb{L}))$ with rate function $\bar{I}$, where for $(\xi, \vartheta) \in C([0,T] : \mathbb{R}^d \times \mathcal{M}_F(\mathbb{L}))$, $\bar{I}(\xi, \vartheta)$ is defined by the right side of (2.8) by replacing $\mathcal{V}(\xi)$ with $\bar{\mathcal{V}}(\xi, \vartheta)$ which is the collection of all (u, φ, π) that satisfy in addition to (2.9) and (2.10) the equality

$$\vartheta(t, A) = \sum_{j \in A} \int_0^t \pi_j(s)ds, \ t \in [0,T], \ A \subset \mathbb{L}.$$

2.2 An equivalent representation for the rate function

In this section we present a different representation for the rate function that will be more convenient to work with in some instances. Recall that λ_ζ denotes Lebesgue measure on $[0, \zeta]$.

Let $\hat{\ell} : \mathcal{M}_F[0,\zeta] \to [0,\infty]$ be defined by

$$\hat{\ell}(\eta) \doteq \begin{cases} \int_{[0,\zeta]} \ell\left(\frac{d\eta}{d\lambda_\zeta}(z)\right) \lambda_\zeta(dz), & \text{if } \eta \ll \lambda_\zeta, \\ \infty, & \text{otherwise.} \end{cases}$$

For $\eta = (\eta_i)_{i \in \mathbb{L}}$, where each $\eta_i \in \mathcal{M}_F[0,\zeta]$, with an abuse of notation we define

$$\hat{\ell}(\eta) = \sum_{i \in \mathbb{L}} \hat{\ell}(\eta_i).$$

For $0 \le a < b \le T$ let $\mathbb{H}_{[a,b]}$ denote the space

$$\mathbb{H}_{[a,b]} \doteq [a,b] \times \mathbb{L} \times (\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|} \times \mathbb{R}^m. \tag{2.14}$$

For convenience, when $[a,b] = [0,t]$, we use the notation $\mathbb{H}_t$ instead of $\mathbb{H}_{[0,t]}$. Let $\mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ denote the space of finite measures Q on $\mathbb{H}_T$ such that

$$Q(\mathbb{H}_{[a,b]}) = b - a, \ \text{ for all } 0 \le a \le b \le T.$$

In other words, denoting the marginal on the i^{th} coordinate of $\mathbb{H}_T$ by $[Q]_i$, $Q \in \mathcal{M}_F(\mathbb{H}_T)$ is in $\mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ if and only if $[Q]_1 = \lambda_T$, where λ_T is Lebesgue measure on $[0,T]$. For notational simplicity, we will denote a typical $(s, y, \eta, z) \in \mathbb{H}_T$ as $\mathbf{v}$. Q encodes time (s), the state of the controlled fast process (y), the measures controlling the jump rates (η), and the control (z) applied to perturb the mean of the Brownian motion. Recall that $b(x,y), a(x,y)$ are the same as $b_y(x), a_y(x)$. For $\xi \in C([0,T] : \mathbb{R}^d)$, let $\mathcal{P}_s(\xi)$ be the family of all $Q \in \mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ such that

$$\int_{\mathbb{H}_T} \|z\|^2 Q(d\mathbf{v}) < \infty, \tag{2.15}$$

$$\xi(t) = x_0 + \int_{\mathbb{H}_t} b(\xi(s), y) Q(d\mathbf{v}) + \int_{\mathbb{H}_t} a(\xi(s), y) z Q(d\mathbf{v}), \tag{2.16}$$

and

$$\int_{\mathbb{H}_t} \Gamma^{\eta}_{y,j}(\xi(s))Q(d\mathbf{v}) = 0 \quad \text{for all } j \in \mathbb{L} \text{ and a.e. } t \in [0,T]. \tag{2.17}$$

Equation (2.16) gives the controlled dynamics, and (2.17) guarantees that the conditional distribution of Q, in the y-variable (i.e., the second coordinate) given the time instant $s \in [0,T]$, the state $\xi(s)$ of the dynamics, and that the rate control measure η are used, is the stationary distribution associated with the generator $\Gamma^{\eta}(\xi(s))$. Define the function $\hat{I} : C([0,T] : \mathbb{R}^d) \to [0,\infty]$ by

$$\hat{I}(\xi) = \inf_{Q \in \mathcal{P}_s(\xi)} \left\{ \int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta) \right] Q(d\mathbf{v}) \right\}. \tag{2.18}$$

In the expression for $\hat{I}$, all statistical relations between the controls (z,η) and the empirical measure for the fast variables are determined by the joint distribution. It is a natural object for purposes of weak convergence analysis, and these relations can be determined by the use of suitable test functions. The following result shows that $\hat{I}$ and I are the same.

Proposition 2.10. *For every $\xi \in C([0,T] : \mathbb{R}^d)$, $\hat{I}(\xi) = I(\xi)$.*

Proof. Fix $\xi \in C([0,T] : \mathbb{R}^d)$. We first show that $\hat{I}(\xi) \le I(\xi)$. Without loss of generality we assume that $I(\xi) < \infty$. Fix $\varepsilon > 0$ and let $(u,\varphi,\pi) \in \mathcal{V}(\xi)$ be such that

$$\sum_{i \in \mathbb{L}} \frac{1}{2} \int_0^T \|u_i(s)\|^2 \pi_i(s) ds + \sum_{(i,j) \in \mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds \le I(\xi) + \varepsilon. \tag{2.19}$$

Define $\bar{\eta}_{ij}(s,dz) \in \mathcal{M}_F[0,\zeta]$ for $i,j \in \mathbb{L}$ and $s \in [0,T]$ by

$$\bar{\eta}_{ij}(s,dz) \doteq \begin{cases} \varphi_{ij}(s,z)\lambda_\zeta(dz), & \text{if } (i,j) \in \mathbb{T} \text{ and } z \mapsto \varphi_{ij}(s,z) \text{ is integrable,} \\ \lambda_\zeta(dz), & \text{otherwise.} \end{cases}$$

Let $\bar{\eta}_i(s) \doteq (\bar{\eta}_{ij}(s,\cdot))_{j \in \mathbb{L}}$. Define $Q \in \mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ by

$$Q([a,b] \times \{i\} \times A \times B) \doteq \int_a^b \pi_i(s)\delta_{\bar{\eta}_i(s)}(A)\delta_{u_i(s)}(B)\, ds,$$

for $A \in \mathcal{B}((\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|})$, $B \in \mathcal{B}(\mathbb{R}^m)$, $0 \le a \le b \le T$, and $i \in \mathbb{L}$. Then it is easy to verify that $Q \in \mathcal{P}_s(\xi)$ and

$$\int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta) \right] Q(d\mathbf{v})$$

equals the left side of (2.19). This proves that $\hat{I}(\xi) \le I(\xi) + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we have $\hat{I}(\xi) \le I(\xi)$. We now consider the reverse inequality, namely $I(\xi) \le \hat{I}(\xi)$. We assume without loss of generality that $\hat{I}(\xi) < \infty$. Let $Q \in \mathcal{P}_s(\xi)$ be such that

$$\int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta) \right] Q(d\mathbf{v}) \le \hat{I}(\xi) + \varepsilon.$$

Let $[Q]_{34|12}(d\eta \times dz|y,s)$ denote the conditional distribution on the third and fourth coordinates given the first and second. Disintegrate the measure Q as

$$Q(ds \times \{y\} \times d\eta \times dz) = ds\, \pi_y(s)\, [Q]_{34|12}(d\eta \times dz|y,s).$$

Define

$$u_y(s) \doteq \int_{(\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}\times\mathbb{R}^m} z[Q]_{34|12}(d\eta \times dz|y,s),\ y \in \mathbb{L}, s \in [0,T].$$

Also, for $(y,s) \in \mathbb{L} \times [0,T]$, let

$$\bar{\eta}_y(s) \doteq \int_{(\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}\times\mathbb{R}^m} \eta[Q]_{34|12}(d\eta \times dz|y,s) \tag{2.20}$$

and write $\bar{\eta}_y = (\bar{\eta}_{yy'})_{y'\in\mathbb{L}}$. By convexity

$$\hat{\ell}(\bar{\eta}_y(s)) \le \int_{(\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}\times\mathbb{R}^m} \hat{\ell}(\eta)[Q]_{34|12}(d\eta \times dz|y,s),$$

and therefore

$$\sum_{y\in\mathbb{L}} \int_{[0,T]} \pi_y(s)\hat{\ell}(\bar{\eta}_y(s))ds \le \int_{\mathbb{H}_T} \hat{\ell}(\eta)Q(d\mathbf{v}).$$

Define

$$\varphi_{yy'}(s,z) \doteq \begin{cases} \frac{d\bar{\eta}_{yy'}(s,\cdot)}{d\lambda_\zeta(\cdot)}(z), & \text{if } \bar{\eta}_{yy'}(s,\cdot) \ll \lambda_\zeta(\cdot), \\ 1, & \text{otherwise.} \end{cases} \tag{2.21}$$

Then note that

$$\sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds \le \int_{\mathbb{H}_T} \hat{\ell}(\eta)Q(d\mathbf{v}).$$

A similar convexity argument shows that

$$\sum_{i\in\mathbb{L}} \frac{1}{2}\int_0^T \|u_i(s)\|^2\pi_i(s)ds \le \frac{1}{2}\int_{\mathbb{H}_T} \|z\|^2 Q(d\mathbf{v}).$$

To complete the proof it suffices to show that

$$(u=(u_i), \varphi=(\varphi_{ij}), \pi=(\pi_i)) \in \mathcal{V}(\xi).$$

First, it is easily checked that ξ satisfies (2.9). Thus it remains to verify (2.10). Since $Q \in \mathcal{P}_s(\xi)$, from (2.17) we have that for all $j \in \mathbb{L}$ and a.e. $s \in [0,T]$

$$\sum_{y\in\mathbb{L}} \pi_y(s) \int_{(\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}\times\mathbb{R}^m} \Gamma^\eta_{y,j}(\xi(s))[Q]_{34|12}(d\eta \times dz|y,s) = 0.$$

This equality can be rewritten as

$$\sum_{y:y\neq j} \pi_y(s) \int \Gamma^\eta_{y,j}(\xi(s))[Q]_{34|12}(d\eta \times dz|y,s) = \pi_j(s) \sum_{i:i\neq j} \int \Gamma^\eta_{j,i}(\xi(s))[Q]_{34|12}(d\eta \times dz|j,s).$$

Using the definition of Γ^η in (2.7), the last display becomes

$$\begin{aligned} &\sum_{y:y\neq j} \pi_y(s) \int \eta_j(E_{yj}(\xi(s)))[Q]_{34|12}(d\eta \times dz|y,s) \\ &\quad = \pi_j(s) \sum_{i:i\neq j} \int \eta_i(E_{ji}(\xi(s)))[Q]_{34|12}(d\eta \times dz|j,s), \end{aligned}$$

which owing to the definition of $\bar{\eta}_y$ in (2.20) is the same as

$$\sum_{y:y\neq j} \pi_y(s)\bar{\eta}_{yj}(s, E_{yj}(\xi(s))) = \pi_j(s) \sum_{i:i\neq j} \bar{\eta}_{ji}(s, E_{ji}(\xi(s))).$$

From the definition of (φ_{ij}) in (2.21) it is now immediate that this is same as (2.10). □

2.3 Compact level sets

We first prove the following lemmas, which will be used in the proof of the main result of this section.

Lemma 2.11. *For $\eta \in (\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}$ and $x \in \mathbb{R}^d$, let $\Gamma^\eta(x)$ be as defined in* (2.7)*. Then there is a $c_1 \in (0,\infty)$ such that for any $M \in [1,\infty)$, $x, x' \in \mathbb{R}^d$, $\eta \in (\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}$,*

$$\sup_{i,j\in\mathbb{L}, i\neq j} \left|\Gamma^\eta_{ij}(x) - \Gamma^\eta_{ij}(x')\right| \leq c_1 \left(e^M \|x - x'\| + \frac{\hat{\ell}(\eta)}{M} \right).$$

Proof. Let $\eta = (\eta_i)_{i=1}^{|\mathbb{L}|}$. Note that the result is automatic if $\eta_i \not\ll \lambda_\zeta$ for some $i \in \mathbb{L}$. Assume now that $\eta_i \ll \lambda_\zeta$ for each i. The following inequality will be used in the proof: for $u, v \in (0,\infty)$ and $\sigma \in [1,\infty)$,

$$uv \leq e^{\sigma u} + \frac{1}{\sigma}(v \log v - v + 1) = e^{\sigma u} + \frac{1}{\sigma}\ell(v). \tag{2.22}$$

Fix $x, x' \in \mathbb{R}^d$ and $i \neq j$ in $\mathbb{L}$. Denoting the set $E_{ij}(x) \Delta E_{ij}(x')$ by $\bar{E}_{ij}$, for $M \in [1,\infty)$

$$\begin{aligned}
\left|\Gamma^\eta_{ij}(x) - \Gamma^\eta_{ij}(x')\right| &= |\eta_j(E_{ij}(x)) - \eta_j(E_{ij}(x'))| \leq \eta_j(\bar{E}_{ij}) \\
&= \int_{\bar{E}_{ij}} 1 \cdot \frac{d\eta_j}{d\lambda_\zeta}(r)\lambda_\zeta(dr) \leq \int_{\bar{E}_{ij}} e^M \lambda_\zeta(dr) + \frac{1}{M}\int_{\bar{E}_{ij}} \ell\left(\frac{d\eta_j}{d\lambda_\zeta}(r)\right)\lambda_\zeta(dr) \\
&\leq e^M \kappa_2 \|x - x'\| + \frac{1}{M}\hat{\ell}(\eta_j),
\end{aligned}$$

where the inequality on the second line follows from (2.22) and the last inequality follows from (2.2). The result follows. □

Remark 2.12. Using the inequality in (2.22) one can similarly show that there exists a $C_1 \in (0,\infty)$ such that for any $M \in [1,\infty)$ and $\eta \in (\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}$,

$$\sup_{x\in\mathbb{R}^d} \max_{i,j\in\mathbb{L}, i\neq j} \Gamma^\eta_{ij}(x) \leq C_1 \left(e^M + \frac{\hat{\ell}(\eta)}{M} \right).$$

The following lemma will be used at several places in weak convergence arguments.

Lemma 2.13. *Let (η^n, Z^n, Y^n) be a sequence of $(\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|} \times \mathbb{R}^d \times \mathbb{L}$ valued random variables given on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, which converges in probability to $(\bar{\eta}, \bar{Z}, \bar{Y})$. Further suppose that, for some $\bar{C} \in (0,\infty)$, $\sup_{n\in\mathbb{N}} \bar{\mathbb{E}}\hat{\ell}(\eta^n) \leq \bar{C}$. Then for all $j \in \mathbb{L}$, $\Gamma^{\eta^n}_{Y^n,j}(Z^n)$ converges in L^1 to $\Gamma^{\bar{\eta}}_{\bar{Y},j}(\bar{Z})$, as $n \to \infty$.*

Proof. Fix $j \in \mathbb{L}$. Using Lemma 2.11 we see that

$$|\Gamma^{\eta^n}_{Y^n,j}(Z^n) - \Gamma^{\eta^n}_{Y^n,j}(\bar{Z})| \to 0, \quad \text{in probability.} \tag{2.23}$$

From Fatou's lemma and the lower semicontinuity of $\hat{\ell}$ it follows that $\bar{\mathbb{E}}\hat{\ell}(\bar{\eta}) \leq \bar{C}$. Next, assume without loss of generality by using a subsequential argument that the convergence of (η^n, Z^n, Y^n) holds a.s. Let $\Omega_0 \in \bar{\mathcal{F}}$ be such that $\bar{\mathbb{P}}(\Omega_0) = 1$ and $\forall \omega \in \Omega_0$,

$$(\eta^n(\omega), Z^n(\omega), Y^n(\omega)) \to (\bar{\eta}(\omega), \bar{Z}(\omega), \bar{Y}(\omega)),$$

and $\hat{\ell}(\bar{\eta}(\omega)) < \infty$. Fix $\omega \in \Omega_0$. We will suppress ω from the notation at some places below. Since $\mathbb{L}$ is a finite set, there exists an $N \equiv N(\omega)$ such that for $n \geq N$, $Y^n(\omega) = \bar{Y}(\omega)$, and consequently, $\Gamma^{\eta^n}_{Y^n,j}(\bar{Z}) = \Gamma^{\eta^n}_{\bar{Y},j}(\bar{Z})$. Since $\eta^n \to \bar{\eta}$ and $\bar{\eta}_j$ is absolutely continuous with respect to λ_ζ for every j, we conclude that

$$\eta^n_j(E_{ij}(x)) \to \bar{\eta}_j(E_{ij}(x)), \quad \forall i, j \in \mathbb{L},\ i \neq j,\ x \in \mathbb{R}^d.$$

Using this in (2.7) we now have that

$$\Gamma^{\eta^n}_{Y^n,j}(\bar{Z}) \to \Gamma^{\bar{\eta}}_{\bar{Y},j}(\bar{Z}), \quad \text{a.s.}$$

Combining this with (2.23) we have that

$$\Gamma^{\eta^n}_{Y^n,j}(Z^n) \to \Gamma^{\bar{\eta}}_{\bar{Y},j}(\bar{Z}) \quad \text{in probability.}$$

Finally to show the L^1-convergence, it suffices to argue that $\eta^n_j[0,\zeta]$, or equivalently $\eta^n_{*,j}[0,\zeta] \doteq \eta^n_j[0,\zeta]/\lambda_\zeta[0,\zeta]$, is uniformly integrable. For this note that

$$\bar{\mathbb{E}}(\ell(\eta^n_{*,j}[0,\zeta])) = \bar{\mathbb{E}}\left(\ell\left(\frac{1}{\zeta}\int_{[0,\zeta]}\frac{d\eta^n_j}{d\lambda_\zeta}(r)\lambda_\zeta(dr)\right)\right) \le \frac{1}{\zeta}\bar{\mathbb{E}}(\hat{\ell}(\eta^n_j)) \le \frac{\bar{C}}{\zeta},$$

where the first inequality follows from the convexity of ℓ. The desired uniform integrability is now an immediate consequence of the superlinearity of ℓ. □

We now show that the function I, which is same as the function $\hat{I}$ defined in (2.18), is a rate function on $C([0,T]:\mathbb{R}^d)$.

Proposition 2.14. *For every $M \in (0,\infty)$, the set $\{\xi \in C([0,T]:\mathbb{R}^d) | I(\xi) \le M\}$ is compact, and consequently I is a rate function on $C([0,T]:\mathbb{R}^d)$.*

Proof. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence in $\{\xi \in C([0,T]:\mathbb{R}^d) | I(\xi) \le M\}$. Since $I(\xi_n) \le M$, we have from Proposition 2.10 that for each $n \in \mathbb{N}$, there exists some $Q_n \in \mathcal{P}_s(\xi_n)$, such that

$$\int_{\mathbb{H}_T}\left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right]Q_n(d\mathbf{v}) \le M + \frac{1}{n}. \tag{2.24}$$

Recall that $\mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ is the space of finite measures on $\mathbb{H}_T \equiv \mathbb{H}_{[0,T]}$ defined in (2.14) whose first marginal is the Lebesgue measure. It suffices to show that $\{\xi_n\}$ is pre-compact, and every limit point belongs to $\{\xi \in C([0,T]:\mathbb{R}^d) | I(\xi) \le M\}$. For this, we prove that:

(i) $\{Q_n, \xi_n\}_{n\in\mathbb{N}}$ is pre-compact in $\mathcal{P}_{\text{leb}}(\mathbb{H}_T) \times C([0,T]:\mathbb{R}^d)$.

(ii) Any limit point $\{Q,\xi\}$ satisfies the properties

  (a) $\int_{\mathbb{H}_T}\left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right]Q(d\mathbf{v}) \le M$,

  (b) (2.16) holds,

  (c) (2.17) holds.

We now prove (i). Since $\mathbb{L}$ is a finite (and hence compact) set and $[Q_n]_1 = \lambda_T$ for all n, in order to prove the pre-compactness of $\{Q_n\}$, it suffices to show that for every $\delta > 0$, there exists a $C_1 \in (0,\infty)$ such that

$$\sup_{n\in\mathbb{N}} Q_n\left\{(s,y,\eta,z) \in \mathbb{H}_T : \sum_{j\in\mathbb{L}}\eta_j[0,\zeta] + \|z\| > C_1\right\} \le \delta. \tag{2.25}$$

From (2.24)

$$\int_{\mathbb{H}_T}\|z\|^2 Q_n(d\mathbf{v}) \le 2(M+1), \ \int_{\mathbb{H}_T}\hat{\ell}(\eta)Q_n(d\mathbf{v}) \le M+1 \tag{2.26}$$

and using (2.22) with $\sigma = 1, u = 1$ and $v = \eta_j[0,\zeta]$,

$$\sum_{j\in\mathbb{L}}\int_{\mathbb{H}_T}\eta_j[0,\zeta]Q_n(d\mathbf{v}) \le |\mathbb{L}|Te + \int_{\mathbb{H}_T}\hat{\ell}(\eta)Q_n(d\mathbf{v}) \le |\mathbb{L}|Te + M + 1.$$

The inequality in (2.25) is now immediate from the last two displays. Thus $\{Q_n\}$ is pre-compact in $\mathcal{P}_{\text{leb}}(\mathbb{H}_T)$. Next we argue the pre-compactness of $\{\xi_n\}$. We first show that

$$\sup_{n\in\mathbb{N}}\sup_{0\le t\le T}\|\xi_n(t)\|^2 \doteq C_2 < \infty. \tag{2.27}$$

Since $Q_n \in \mathcal{P}_s(\xi_n)$, we have

$$\xi_n(t) = x_0 + \int_{\mathbb{H}_t} b(\xi_n(s), y)Q_n(d\mathbf{v}) + \int_{\mathbb{H}_t} a(\xi_n(s), y)zQ_n(d\mathbf{v}).$$

Using the linear growth property of a, b (Remark 2.2), we have

$$\begin{aligned}\|\xi_n(t)\|^2 \le 3\|x_0\|^2 &+ 3T\int_{\mathbb{H}_t}\kappa_1^2(\|\xi_n(s)\|+1)^2Q_n(d\mathbf{v})\\ &+ 3\kappa_1^2\int_{\mathbb{H}_t}(\|\xi_n(s)\|+1)^2Q_n(d\mathbf{v})\int_{\mathbb{H}_t}\|z\|^2Q_n(d\mathbf{v}).\end{aligned}$$

Thus from (2.26)

$$\|\xi_n(t)\|^2 \le 3\|x_0\|^2 + 6\kappa_1^2(T+2(M+1))\int_{[0,t]}(\|\xi_n(s)\|^2+1)ds.$$

The inequality in (2.27) now follows by Gronwall's inequality. Next, consider fluctuations of ξ_n. For $0 \le t_0 \le t_1 \le T$,

$$\begin{aligned}\|\xi_n(t_1) - \xi_n(t_0)\| &\le \int_{\mathbb{H}_{[t_0,t_1]}}\|b(\xi_n(s),y)\|Q_n(d\mathbf{v}) + \int_{\mathbb{H}_{[t_0,t_1]}}\|a(\xi_n(s),y)\|\|z\|Q_n(d\mathbf{v})\\ &\le \int_{\mathbb{H}_{[t_0,t_1]}}\kappa_1(\|\xi_n(s)\|+1)Q_n(d\mathbf{v})\\ &\quad + \left(\int_{\mathbb{H}_{[t_0,t_1]}}(\kappa_1(\|\xi_n(s)\|+1))^2Q_n(d\mathbf{v})\int_{\mathbb{H}_T}\|z\|^2Q_n(d\mathbf{v})\right)^{1/2}\\ &\le C_3|t_1-t_0|^{1/2},\end{aligned}$$

where the last inequality uses (2.27) and (2.26) and C_3 depends only on C_2, M, κ_1 and T. This estimate together with (2.27) shows that $\{\xi_n\}$ is pre-compact in $C([0,T]:\mathbb{R}^d)$. We now prove (ii). Let (Q,ξ) be a limit point of the sequence $\{(Q_n,\xi_n)\}_{n\in\mathbb{N}}$. Part (a) is immediate from (2.24) using Fatou's lemma and the lower semicontinuity of $\hat{\ell}$. Consider now part (b). We assume without loss of generality that the full sequence converges to (Q,ξ). From the Lipschitz property of a (Assumption 2.1), we have

$$\begin{aligned}\int_{\mathbb{H}_T}\|a(\xi_n(s),y)-a(\xi(s),y)\|\|z\|Q_n(d\mathbf{v}) &\le d_{\text{lip}}\int_{\mathbb{H}_T}\|\xi_n(s)-\xi(s)\|\|z\|Q_n(d\mathbf{v})\\ &\le d_{\text{lip}}\sup_{s\le T}\|\xi_n(s)-\xi(s)\|\int_{\mathbb{H}_T}\|z\|Q_n(d\mathbf{v})\\ &\le d_{\text{lip}}\sup_{s\le T}\|\xi_n(s)-\xi(s)\|\sqrt{T}\sqrt{2(M+1)}\\ &\to 0\end{aligned}$$

as $n\to\infty$. A similar calculation shows that as $n\to\infty$

$$\int_{\mathbb{H}_T}\|b(\xi_n(s),y)-b(\xi(s),y)\|Q_n(d\mathbf{v}) \to 0.$$

Since $(s,y,\eta,z) \mapsto (b(\xi(s),y),a(\xi(s),y))$ is a continuous and bounded map, and $\int_{\mathbb{H}_T} \|z\|^2 Q_n(d\mathbf{v}) \le 2(M+1)$, it follows from convergence of Q_n to Q that

$$\int_{\mathbb{H}_T} [b(\xi(s),y)+a(\xi(s),y)z]Q_n(d\mathbf{v}) \to \int_{\mathbb{H}_T} [b(\xi(s),y)+a(\xi(s),y)z]Q(d\mathbf{v}). \tag{2.28}$$

Combining the last three convergence statements we have (b). Next we consider part (c). By Lemma 2.11 we have that for $M_0 \in [1,\infty)$, $j \in \mathbb{L}$,

$$\begin{aligned} &\int_{\mathbb{H}_T} \left|\Gamma^\eta_{yj}(\xi_n(s)) - \Gamma^\eta_{yj}(\xi(s))\right| Q_n(d\mathbf{v}) \\ &\quad \le c_1 \left(e^{M_0}\kappa_2 \sup_{s\le T} \|\xi_n(s)-\xi(s)\| + \frac{1}{M_0}\int_{\mathbb{H}_T} \hat{\ell}(\eta) Q_n(d\mathbf{v}) \right). \end{aligned} \tag{2.29}$$

Sending $n \to \infty$ and then $M_0 \to \infty$ we see from (2.26) that the left side of (2.29) converges to 0 as $n \to \infty$. Finally, by the Skorohod representation theorem, (2.26) and Lemma 2.13,

$$\int_{\mathbb{H}_T} \Gamma^\eta_{yj}(\xi(s)) Q_n(d\mathbf{v}) \to \int_{\mathbb{H}_T} \Gamma^\eta_{yj}(\xi(s)) Q(d\mathbf{v}). \tag{2.30}$$

Combining this with (2.29) and recalling that $Q_n \in \mathcal{P}_s(\xi_n)$, we have (c). This completes the proof. □

3 Large deviation upper bound

The main result of this section is Theorem 3.5, which shows that for all $F \in C_b(C([0,T]:\mathbb{R}^d))$

$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \le -\inf_{\xi\in C([0,T]:\mathbb{R}^d)} \{F(\xi)+I(\xi)\}. \tag{3.1}$$

To do this we show a lower bound on the corresponding variational representations.

Let $\mathcal{PF}$ denote the predictable σ-field on $[0,T]\times\Omega$ associated with the filtration $\{\mathcal{F}_t : 0 \le t \le T\}$. Let

$$\mathcal{PF}[\mathbb{T}] \doteq \{\varphi = (\varphi_{ij})_{(i,j)\in\mathbb{T}} : \varphi_{ij} \text{ is } (\mathcal{PF}\otimes\mathcal{B}[0,\zeta]\backslash\mathcal{B}[0,\infty)) \text{ measurable for all } (i,j)\in\mathbb{T}\}.$$

For $h : [0,T] \to \mathbb{R}^m$, let $\tilde{L}_T(h) \doteq \frac{1}{2}\int_0^T \|h(s)\|^2\, ds$. Also, for $g = (g_{ij})_{(i,j)\in\mathbb{T}}$ such that $g_{ij} : [0,\zeta]\times[0,T] \to [0,\infty)$, let

$$\bar{L}_T(g) \doteq \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(g_{ij}(s,z))\lambda_\zeta(dz)ds.$$

Consider the following spaces. Let

$$S_2^M \doteq \{h : [0,T]\to\mathbb{R}^m : \tilde{L}_T(h) \le M\},\ S_\ell^M \doteq \{g\in\mathcal{R} : \bar{L}_T(g)\le M\},$$
$$S_2 \doteq \cup_{M=1}^\infty S_2^M, \quad S_\ell \doteq \cup_{M=1}^\infty S_\ell^M,$$

and

$$\mathcal{PF}_2^M \doteq \left\{\psi : \psi \text{ is } \mathcal{PF}\backslash\mathcal{B}(\mathbb{R}^m) \text{ measurable and } \psi\in S_2^M \ \mathbb{P}\text{- a.s. }\right\}, \quad \mathcal{PF}_2 \doteq \cup_{M=1}^\infty \mathcal{PF}_2^M,$$
$$\mathcal{PF}_\ell^M \doteq \left\{\varphi\in\mathcal{PF}[\mathbb{T}] : \varphi(\cdot,\omega)\in S_\ell^M,\ \mathbb{P}\text{- a.s. }\right\}, \quad \mathcal{PF}_\ell \doteq \cup_{M=1}^\infty \mathcal{PF}_\ell^M.$$

Set $\mathcal{U} = \mathcal{PF}_2 \times \mathcal{PF}_\ell$. For $f = (h,g) \in S_2\times S_\ell$, let

$$L_T(f) = \tilde{L}_T(h) + \bar{L}_T(g). \tag{3.2}$$

With this notation the variational representation of [6] says that

$$\begin{aligned} &-\varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \\ &= \inf_{u=(\psi,\varphi)\in\mathcal{U}} \mathbb{E}\left[L_T(u) + F\circ\mathcal{G}^\varepsilon\left(\sqrt{\varepsilon}W + \int_0^\cdot \psi(s)ds,\, \varepsilon N^{\varepsilon^{-1}\varphi}\right)\right]. \end{aligned} \tag{3.3}$$

In fact, a closer inspection of the proof of Theorem 2.8 of [6] (see [7, Theorem 2.4]) shows that (3.3) can be strengthened as follows. For $n \in \mathbb{N}$, define

$$\begin{aligned} \mathcal{PF}_\ell^b = \{\varphi = (\varphi_{ij}) \in \mathcal{PF}_\ell :\ &\text{for some } n \in \mathbb{N}, \varphi_{ij}(r,s,\omega) \in [n^{-1}, n], \\ &\text{for all } (r,s,\omega) \in [0,\zeta]\times[0,T]\times\Omega, (i,j)\in\mathbb{T}\}. \end{aligned}$$

Also let $\mathcal{U}_b = \mathcal{PF}_2 \times \mathcal{PF}_\ell^b$. Then in the equality in (3.3), $\mathcal{U}$ on the right side can be replaced by $\mathcal{U}_b$.

Since S_2^M is a closed ball in $L^2([0,T])$, it is compact under the weak topology. A $g \in S_\ell^M$ can be identified with $\theta_T^g = (\theta_T^{g_{ij}})_{(i,j)\in\mathbb{T}}$ such that each $\theta_T^{g_{ij}}$ is a measure on $[0,\zeta]\times[0,T]$, defined by

$$\theta_T^{g_{ij}}(C) = \int_C g_{ij}(r,s)\lambda_\zeta(dr)ds, \quad C \in \mathcal{B}([0,\zeta]\times[0,T]).$$

With the usual weak convergence topology on the space of finite measures on $[0,\zeta]\times[0,T]$, this identification induces a topology on S_ℓ^M under which it is a compact space. Throughout, we use these topologies on S_2^M and S_ℓ^M.

Controlled versions of processes will be denoted by an overbar, with the particular controls used clear from context. Thus for $(\psi,\varphi) \in \mathcal{U}_b$ we consider the coupled equations

$$\left.\begin{aligned} d\bar{X}^\varepsilon(t) &= b(\bar{X}^\varepsilon(t),\bar{Y}^\varepsilon(t))dt + \sqrt{\varepsilon}a(\bar{X}^\varepsilon(t),\bar{Y}^\varepsilon(t))dW(t) \\ &\quad + a(\bar{X}^\varepsilon(t),\bar{Y}^\varepsilon(t))\psi(t)dt \\ d\bar{Y}^\varepsilon(t) &= \sum_{(i,j)\in\mathbb{T}} \int_{r\in[0,\zeta]} (j-i)1_{\{\bar{Y}^\varepsilon(t-)=i\}} 1_{E_{ij}(\bar{X}^\varepsilon(t))}(r) N_{ij}^{\varepsilon^{-1}\varphi_{ij}}(dr\times dt) \end{aligned}\right\} \tag{3.4}$$

with $(\bar{X}^\varepsilon(0),\bar{Y}^\varepsilon(0)) = (x_0,y_0)$. Recall the map $\mathcal{G}^\varepsilon$ introduced below (2.3). From unique pathwise solvability of (3.4) and a standard argument based on Girsanov's theorem (see for example [7, Section 3.2]) it follows that $\mathcal{G}^\varepsilon\left(\sqrt{\varepsilon}W + \int_0^\cdot \psi^\varepsilon(s)ds,\, (\varepsilon N_{ij}^{\varphi_{ij}^\varepsilon/\varepsilon})\right)$ is the unique solution of (3.4) with ψ and φ replaced by ψ^ε and φ^ε. Thus the representation in (3.3) yields

$$-\varepsilon\log\mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] = \inf_{u=(\psi,\varphi)\in\mathcal{U}_b} \mathbb{E}\left[L_T(u) + F\left(\bar{X}^\varepsilon\right)\right]. \tag{3.5}$$

The following lemma will be used in proving a tightness property. In many places below we will consider controls u subject to an a.s. constraint of the form $L_T(u) \le M$. To simplify the notation, the almost sure qualification is omitted.

Lemma 3.1. *For every $M \in (0,\infty)$*

$$\sup_{\varepsilon\in(0,1)} \sup_{u=(\psi,\phi)\in\mathcal{U}_b : L_T(u)\le M} \mathbb{E}\left(\sup_{0\le t\le T} \|\bar{X}^\varepsilon(t)\|^2\right) < \infty.$$

Proof. Fix $M \in (0,\infty)$ and $u = (\psi,\phi) \in \mathcal{U}_b$ with $L_T(u) \le M$. Now write $\bar{X}^\varepsilon$ as

$$\bar{X}^\varepsilon(t) = x_0 + \bar{B}^\varepsilon(t) + \bar{A}^\varepsilon(t) + \bar{\mathcal{E}}^\varepsilon(t), \tag{3.6}$$

where

$$\bar{B}^\varepsilon(t) = \int_0^t b(\bar{X}^\varepsilon(s),\bar{Y}^\varepsilon(s))ds;$$

$$\bar{A}^\varepsilon(t) = \int_0^t a(\bar{X}^\varepsilon(s), \bar{Y}^\varepsilon(s))\psi(s)ds;$$
$$\bar{\mathcal{E}}^\varepsilon(t) = \sqrt{\varepsilon}\int_0^t a(\bar{X}^\varepsilon(s), \bar{Y}^\varepsilon(s))dW(s).$$

By Remark 2.2,

$$\begin{aligned}\mathbb{E}\left[\sup_{r\le t}\|\bar{A}^\varepsilon(r)\|^2\right] &\le \mathbb{E}\int_0^t \|a(\bar{X}^\varepsilon(s), \bar{Y}^\varepsilon(s))\|^2 ds \int_0^t \|\psi(s)\|^2 ds \\ &\le 4\kappa_1 M\int_0^t (1+\mathbb{E}\|\bar{X}^\varepsilon(s)\|^2)ds \\ &\le 4\kappa_1 MT + 4\kappa_1 M\int_0^t \mathbb{E}\left(\sup_{s\le v}\|\bar{X}^\varepsilon(s)\|^2\right)dv.\end{aligned}$$

Also, by Doob's maximal inequality we have with $C_1 = 8\kappa_1^2$,

$$\begin{aligned}\mathbb{E}\left(\sup_{r\le t}\|\bar{\mathcal{E}}^\varepsilon(r)\|^2\right) &\le 4\varepsilon\mathbb{E}\left[\int_0^t \|a(\bar{X}^\varepsilon(s), \bar{Y}^\varepsilon(s))\|^2 ds\right] \\ &\le 8\kappa_1^2\varepsilon\int_0^t \left(1+\mathbb{E}\left[\|\bar{X}^\varepsilon(s)\|^2\right]\right)ds \\ &\le C_1\varepsilon T + C_1\varepsilon\int_0^t \mathbb{E}\left(\sup_{s\le v}\|\bar{X}^\varepsilon(s)\|^2\right)dv.\end{aligned} \tag{3.7}$$

Similar calculation shows that

$$\mathbb{E}\left[\sup_{r\le t}\|\bar{B}^\varepsilon(r)\|^2\right] \le 2\kappa_1 T + 2\kappa_1 T\int_0^t \mathbb{E}\left(\sup_{s\le v}\|\bar{X}^\varepsilon(s)\|^2\right).$$

Hence from (3.6) it follows that with $C_2 = C_1(T+1) + 2\kappa_1(T+1)(1+2M)$

$$\mathbb{E}\left[\sup_{r\le t}\|\bar{X}^\varepsilon(r)\|^2\right] \le C_2 + C_2\int_0^t \mathbb{E}\left(\sup_{s\le v}\|\bar{X}^\varepsilon(s)\|^2\right)dv.$$

The lemma now follows from Gronwall's inequality. □

Proposition 3.2. *Fix $M \in (0,\infty)$. For $\varepsilon \in (0,1)$ let $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_b$ be such that $L_T(u^\varepsilon) \le M$. Let $(\bar{X}^\varepsilon, \bar{Y}^\varepsilon)$ solve* (3.4) *with ψ and φ replaced by ψ^ε and φ^ε. Then $\{\bar{X}^\varepsilon\}_{\varepsilon\in(0,1)}$ is tight in $C([0,T]:\mathbb{R}^d)$.*

Proof. Write $\bar{X}^\varepsilon$ as in (3.6). From (3.7) and Lemma 3.1

$$\mathbb{E}\sup_{t\in[0,T]}\|\bar{\mathcal{E}}^\varepsilon(t)\|^2 \le C_1\varepsilon T\left(1+\sup_{\varepsilon'\in(0,1)}\mathbb{E}\sup_{t\in[0,T]}\|\bar{X}^{\varepsilon'}(t)\|^2\right) \to 0, \tag{3.8}$$

as $\varepsilon\to 0$. It thus suffices to prove the tightness of $\{\bar{A}^\varepsilon\}$ and $\{\bar{B}^\varepsilon\}$. For this, note that for $0\le h\le T$ and $\varepsilon\in(0,1)$

$$\begin{aligned}&\mathbb{E}\sup_{0\le t_1\le t_2\le T, t_2\le t_1+h}\|\bar{A}^\varepsilon(t_2)-\bar{A}^\varepsilon(t_1)\|^2 \\ &= \mathbb{E}\sup_{0\le t_1\le t_2\le T, t_2\le t_1+h}\left\|\int_{t_1}^{t_2} a(\bar{X}^\varepsilon(s), \bar{Y}^\varepsilon(s))\psi^\varepsilon(s)ds\right\|^2 \\ &\le 2\kappa_1^2\mathbb{E}\sup_{0\le t_1\le T-h}\left(\int_{t_1}^{t_1+h}(1+\|\bar{X}^\varepsilon(s)\|^2)\,ds\int_0^T\|\psi^\varepsilon(s)\|^2 ds\right) \\ &\le 4\kappa_1^2 Mh\left(1+\sup_{\varepsilon\in(0,1)}\mathbb{E}\sup_{0\le t\le T}\|\bar{X}^\varepsilon(t)\|^2\right),\end{aligned}$$

where the last inequality used the fact that $L_T(u^\varepsilon) \le M$. Lemma 3.1 now gives that $\{\bar{A}^\varepsilon\}$ is tight in $C([0,T]:\mathbb{R}^d)$. Similar calculations show that $\{\bar{B}^\varepsilon\}$ is tight as well. The result follows. □

Remark 3.3. The estimate in Proposition 3.2 in particular shows that for every $M \in (0,\infty)$ there exists some $c_2(M) \in (0,\infty)$, such that

$$\sup_{\varepsilon\in(0,1)} \sup_{0\le s\le T-\delta} \sup_{u^\varepsilon\in\mathcal{U}_b: L_T(u^\varepsilon)\le M} \mathbb{E} \sup_{0\le t\le\delta} \|\bar{X}^\varepsilon(s+t) - \bar{X}^\varepsilon(s)\|^2 \le c_2(M)\delta$$

for any $\delta \in [0,T]$.

Given $\varepsilon > 0$, let $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_b$ and $(\bar{X}^\varepsilon, \bar{Y}^\varepsilon)$ be as in Proposition 3.2. Note that by (3.4) only the controlled rates $\varphi^\varepsilon_{\bar{Y}^\varepsilon(t-),\cdot}$ affect the evolution of $(\bar{X}^\varepsilon, \bar{Y}^\varepsilon)$. In proving the Laplace upper bound, we can (and will) assume without loss of generality that

$$\varphi^\varepsilon_{ij}(t,z) = 1 \text{ for all } (i,j) \in \mathbb{T} \text{ such that } i \in \mathbb{L}\setminus\{Y^\varepsilon(t-)\} \text{ and } (t,z)\in[0,T]\times[0,\zeta]. \quad (3.9)$$

Also, by convention we will take $\varphi^\varepsilon_{ij}(t,r) = 1$ if $(i,j) \notin \mathbb{T}$.

The proof of the Laplace upper bound relies on the asymptotic analysis of the following occupation measure. For $t \in [0,T]$, $\eta^\varepsilon(t) \in (\mathcal{M}_F[0,\zeta])^{|\mathbb{L}|}$ is defined to be $\eta^\varepsilon(t) = (\eta^\varepsilon_j(t))_{j\in\mathbb{L}}$, where

$$\eta^\varepsilon_j(t)(F) = \sum_{i\in\mathbb{L}} 1_{\{Y^\varepsilon(t-)=i\}} \int_F \varphi^\varepsilon_{ij}(t,r)\lambda_\zeta(dr), \quad F \in \mathcal{B}[0,\zeta]. \quad (3.10)$$

Then define $Q^\varepsilon \in \mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ by

$$Q^\varepsilon(A\times B\times C\times D) \doteq \int_{[0,T]} 1_A(s)1_B(\bar{Y}^\varepsilon(s))1_C(\eta^\varepsilon(s))1_D(\psi^\varepsilon(s))ds. \quad (3.11)$$

The main step in the proof will be to characterize the limit points of Q^ε. Note that, with $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_b$ and $\bar{X}^\varepsilon$ as in Proposition 3.2, it follows from (3.8) that

$$\lim_{\varepsilon\to0} \mathbb{E} \sup_{0\le t\le T} \left\| \bar{X}^\varepsilon(t) - x_0 - \int_{\mathbb{H}_t} b(\bar{X}^\varepsilon(s),y)Q^\varepsilon(d\mathbf{v}) - \int_{\mathbb{H}_t} a(\bar{X}^\varepsilon(s),y)zQ^\varepsilon(d\mathbf{v}) \right\| = 0. \quad (3.12)$$

For $\varepsilon \in (0,1)$, let $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_b$ and define Q^ε as in (3.11). Then note that

$$\int_0^T \|\psi^\varepsilon(s)\|^2 ds = \int_{\mathbb{H}_T} \|z\|^2 Q^\varepsilon(d\mathbf{v}) \quad (3.13)$$

and from (3.9) and (3.10)

$$\begin{aligned}\sum_{(i,j)\in\mathbb{T}} \int_{[0,T]\times[0,\zeta]} \ell(\varphi^\varepsilon_{ij}(r,s))\lambda_\zeta(dr)ds &= \sum_{j\in\mathbb{L}}\sum_{i\in\mathbb{L}} \int_{[0,T]\times[0,\zeta]} 1_{\{Y^\varepsilon(s)=i\}}\ell(\varphi^\varepsilon_{ij}(r,s))\lambda_\zeta(dr)ds \\ &= \int_{[0,T]} \hat{\ell}(\eta^\varepsilon(s))ds \\ &= \int_{\mathbb{H}_T} \hat{\ell}(\eta)Q^\varepsilon(d\mathbf{v}). \end{aligned} \quad (3.14)$$

Thus

$$L_T(u^\varepsilon) = \int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right] Q^\varepsilon(d\mathbf{v}). \quad (3.15)$$

We now prove the tightness of $\{(\bar{X}^\varepsilon, Q^\varepsilon)\}$ and characterize the limit points.

Proposition 3.4. *Fix $M \in (0,\infty)$. Define Q^ε as in* (3.11)*. Suppose for $\varepsilon \in (0,1)$, $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_b$ is such that $L_T(u^\varepsilon) \le M$. Then $\{(\bar{X}^\varepsilon, Q^\varepsilon)\}$ is a tight family of $C([0,T]:\mathbb{R}^d) \times \mathcal{M}_F(\mathbb{H}_T)$-valued random variables. Furthermore, if (ξ, Q) is a weak limit point of $\{(\bar{X}^\varepsilon, Q^\varepsilon)\}$, then*

1. $M \ge \int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right] Q(d\mathbf{v})$;
2. *Equations* (2.16) *and* (2.17) *hold a.s.*

Proof. Tightness of $\{\bar{X}^\varepsilon\}$ was shown in Proposition 3.2. Next we argue the tightness of $\{Q^\varepsilon\}$. From (3.15) we have

$$\int_{\mathbb{H}_T} \left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right] Q^\varepsilon(d\mathbf{v}) \le M. \tag{3.16}$$

To prove the tightness of $\{Q^\varepsilon\}$, it suffices to show that for any $\delta \in (0,\infty)$, there exists $C_1 \in (0,\infty)$ such that:

$$\sup_\varepsilon Q^\varepsilon \left\{ (s,y,\eta,z) \in \mathbb{H}_T : \sum_{j\in\mathbb{L}} \eta_j[0,\zeta] + \|z\| > C_1 \right\} \le \delta.$$

However, this is proved exactly as (2.25) using (3.16) instead of (2.24). The inequality in part 1 follows immediately from (3.16) using Fatou's Lemma and lower semicontinuity of $\eta \mapsto \hat{\ell}(\eta)$. We now prove 2. For this we assume without loss of generality (using the Skorokhod representation) that $(\bar{X}^\varepsilon, Q^\varepsilon)$ converges a.s. to (ξ, Q). Following similar steps as in the proof of Proposition 2.14 (see the proof of (2.28)), we conclude that

$$\int_{\mathbb{H}_T} [b(\bar{X}^\varepsilon(s), y) + a(\bar{X}^\varepsilon(s), y)z] Q^\varepsilon(d\mathbf{v}) \to \int_{\mathbb{H}_T} [b(\xi(s), y) + a(\xi(s), y)z] Q(d\mathbf{v}).$$

It now follows from (3.12) that (2.16) holds. We now prove that (2.17) holds. Note that

$$\int_{\mathbb{H}_t} \Gamma^\eta_{y,j}(\bar{X}^\varepsilon(s)) Q^\varepsilon(d\mathbf{v}) = \int_0^t \Gamma^{\eta^\varepsilon(u)}_{\bar{Y}^\varepsilon(u),j}(\bar{X}^\varepsilon(u))du, \ j \in \mathbb{L}, t \in [0,T]. \tag{3.17}$$

Recall the sets $E_{ij}(x)$ defined in (2.1). Then from (3.9) and (3.10), for any ϕ mapping $\mathbb{L}$ to $\mathbb{R}$

$$\begin{aligned}
&\phi(\bar{Y}^\varepsilon(t)) - \phi(y_0) \\
&\quad = \sum_{(i,j)\in\mathbb{T}} (\phi(j) - \phi(i)) \int_{[0,\zeta]\times[0,t]} 1_{\{\bar{Y}^\varepsilon(s-)=i\}} 1_{E_{ij}(\bar{X}^\varepsilon(s))}(r) N_{ij}^{\varepsilon^{-1}\varphi^\varepsilon_{ij}}(dr \times ds) \\
&\quad = \varepsilon^{-1} \sum_{(i,j)\in\mathbb{T}} (\phi(j) - \phi(i)) \int_{[0,\zeta]\times[0,t]} 1_{\{\bar{Y}^\varepsilon(s)=i\}} 1_{E_{ij}(\bar{X}^\varepsilon(s))}(r) \varphi^\varepsilon_{ij}(r,s) \lambda_\zeta(dr)ds + \bar{M}^\varepsilon_\phi(t) \\
&\quad = \varepsilon^{-1} \sum_{(i,j)\in\mathbb{T}} (\phi(j) - \phi(i)) \int_0^t 1_{\{\bar{Y}^\varepsilon(s)=i\}} \eta^\varepsilon_j(E_{ij}(\bar{X}^\varepsilon(s)))ds + \bar{M}^\varepsilon_\phi(t),
\end{aligned} \tag{3.18}$$

where $\bar{M}^\varepsilon_\phi$ is the martingale given by

$$\bar{M}^\varepsilon_\phi(t) \doteq \sum_{(i,j)\in\mathbb{T}} (\phi(j) - \phi(i)) \int_{[0,\zeta]\times[0,t]} 1_{\{\bar{Y}^\varepsilon(s-)=i\}} 1_{E_{ij}(\bar{X}^\varepsilon(s))}(r) \tilde{N}_{ij}^{\varepsilon^{-1}\varphi^\varepsilon_{ij}}(dr \times ds)$$

and $\tilde{N}_{ij}^{\varepsilon^{-1}\varphi_{ij}^{\varepsilon}}(dr\times ds) = N_{ij}^{\varepsilon^{-1}\varphi_{ij}^{\varepsilon}}(dr\times ds) - \varepsilon^{-1}\varphi_{ij}^{\varepsilon}(r,s)dr\,ds$. By Doob's inequality

$$\begin{aligned}\mathbb{E}\sup_{0\le s\le T}|\varepsilon\bar{M}_{\phi}^{\varepsilon}(s)|^2 &\le 4\varepsilon\sum_{(i,j)\in\mathbb{T}}(\phi(j)-\phi(i))^2\int_{[0,\zeta]\times[0,T]}\varphi_{ij}^{\varepsilon}(r,s)\lambda_{\zeta}(dr)ds\\ &\le 16\varepsilon\|\phi\|_{\infty}^2\sum_{(i,j)\in\mathbb{T}}\int_{[0,\zeta]\times[0,T]}\varphi_{ij}^{\varepsilon}(r,s)\lambda_{\zeta}(dr)ds\\ &\le 16\varepsilon\|\phi\|_{\infty}^2\left(\zeta Te+M\right),\end{aligned}$$

where the last inequality uses (2.22). It follows that $\sup_{0\le s\le T}|\varepsilon\bar{M}_{\phi}^{\varepsilon}(s)|$ converges to 0 in probability as $\varepsilon\to 0$. Next, from (3.18) we see that

$$\varepsilon(\phi(\bar{Y}^{\varepsilon}(t))-\phi(y_0)) = \sum_{(i,j)\in\mathbb{T}}(\phi(j)-\phi(i))\int_0^t 1_{\{\bar{Y}^{\varepsilon}(s)=i\}}\eta_j^{\varepsilon}(E_{ij}(\bar{X}^{\varepsilon}(s)))ds+\varepsilon\bar{M}_{\phi}^{\varepsilon}(t),\ t\in[0,T].$$

Since ϕ is bounded, we conclude that as $\varepsilon\to 0$,

$$\sup_{0\le t\le T}\left|\sum_{(i,j)\in\mathbb{T}}(\phi(j)-\phi(i))\int_0^t 1_{\{\bar{Y}^{\varepsilon}(s)=i\}}\eta_j^{\varepsilon}(E_{ij}(\bar{X}^{\varepsilon}(s)))ds\right|\to 0.$$

For fixed $j\in\mathbb{L}$, taking $\phi\doteq 1_{\{j\}}$, we now see from (2.7) that

$$\sup_{0\le t\le T}\left|\sum_{i\in\mathbb{L}}\int_0^t 1_{\{\bar{Y}^{\varepsilon}(s)=i\}}\eta_j^{\varepsilon}(E_{ij}(\bar{X}^{\varepsilon}(s)))ds\right| = \sup_{0\le t\le T}\left|\int_0^t\Gamma^{\eta^{\varepsilon}(s)}_{\bar{Y}^{\varepsilon}(s),j}(\bar{X}^{\varepsilon}(s))ds\right|$$

converges to 0 as $\varepsilon\to 0$. Hence from (3.17), $\int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(\bar{X}^{\varepsilon}(s))Q^{\varepsilon}(d\mathbf{v})\to 0$, uniformly in $t\in[0,T]$. Now as in the proof of Proposition 2.14 (see (2.29) and (2.30)), $\int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(\bar{X}^{\varepsilon}(s))Q^{\varepsilon}(d\mathbf{v})\to\int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(\xi(s))Q(d\mathbf{v})$. Thus (2.17) is satisfied and the result follows. □

We now prove the upper bound in (3.1). Recall that for $\varepsilon>0$, $(X^{\varepsilon},Y^{\varepsilon})$ is given as the unique pathwise solution of (2.3).

Theorem 3.5. *For any $F\in C_b(C([0,T]:\mathbb{R}^d))$, the inequality in* (3.1) *holds.*

Proof. Using (3.5) for every $\varepsilon>0$, we can find $(\psi^{\varepsilon},\varphi^{\varepsilon})\in\mathcal{U}_b$ such that

$$-\varepsilon\log\mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^{\varepsilon})\right)\right]\ge\mathbb{E}\left[L_T(\psi^{\varepsilon},\varphi^{\varepsilon})+F(\bar{X}^{\varepsilon})\right]-\varepsilon,$$

where $\bar{X}^{\varepsilon}$ solves (3.4) (with (ψ,φ) replaced with $(\psi^{\varepsilon},\varphi^{\varepsilon})$). Since F is bounded

$$\sup_{\varepsilon\in(0,1)}\mathbb{E}L_T(\psi^{\varepsilon},\varphi^{\varepsilon})\le 2\|F\|_{\infty}+1<\infty,$$

By a localization argument (see, e.g., [6, Section A.3]) we can assume without loss of generality that for some $M\in(0,\infty)$

$$\sup_{\varepsilon\in(0,1)}L_T(\psi^{\varepsilon},\varphi^{\varepsilon})\le M \text{ a.s.}.$$

Now Proposition 3.4 implies that $(\bar{X}^{\varepsilon},Q^{\varepsilon})$ is tight and any limit point (ξ,Q) satisfies (2.15)–(2.17) a.s. and consequently $Q\in\mathcal{P}_s(\xi)$ a.s., where $\mathcal{P}_s(\xi)$ was introduced above

(2.15). Assume without loss of generality that the convergence to (ξ, Q) holds along the full sequence. Then

$$\begin{aligned}\liminf_{\varepsilon\to 0} -\varepsilon \log \mathbb{E}_x\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] &\geq \mathbb{E}\left[\left(\int_{\mathbb{H}_T}\left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right]Q(d\mathbf{v})\right) + F(\xi)\right] \\ &\geq \mathbb{E}\left[\inf_{Q\in\mathcal{P}_s(\xi)}\left(\int_{\mathbb{H}_T}\left[\frac{1}{2}\|z\|^2 + \hat{\ell}(\eta)\right]Q(d\mathbf{v})\right) + F(\xi)\right] \\ &= \mathbb{E}\left[I(\xi) + F(\xi)\right] \\ &\geq \inf_{\xi\in C([0,T]:\mathbb{R}^d)}\left[I(\xi) + F(\xi)\right],\end{aligned}$$

where the first inequality uses Fatou's lemma and part (i) of Proposition 3.4, the second uses the property that $Q \in \mathcal{P}_s(\xi)$ a.s. and the third uses the definition of the rate function in (2.18) and Proposition 2.10. □

4 Near optimal paths with a unique characterization

In order to prove the large deviation lower bound (2.13), a natural approach is to consider a ξ that is a near infimum for the right side in (2.13) and construct a sequence of controls $(\psi^\varepsilon, \varphi^\varepsilon)$ such that $\bar{X}^\varepsilon \Rightarrow \xi$ where $\bar{X}^\varepsilon$ is as in (3.4) with (ψ, φ) replaced by $(\psi^\varepsilon, \varphi^\varepsilon)$ respectively. Along with an appropriate convergence of costs, the variational representation in (3.5) can then be used to argue that (2.13) holds. For a near optimal ξ, let $(u, \varphi, \pi) \in \mathcal{V}(\xi)$ be a near infimum for the expression on the right side of (2.8). The control pair (u, φ) suggests a natural sequence of controls $(\psi^\varepsilon, \varphi^\varepsilon)$ (see (5.4)) for the construction of controlled processes $\bar{X}^\varepsilon$ and an occupation measure Q^ε of the form in (3.11). Our strategy in the proof of the lower bound given in Section 5 will be to show that any limit point $\bar{\xi}$ of $\bar{X}^\varepsilon$ and a suitable marginal $\bar{\pi}$ of the limit point $\bar{Q}$ of Q^ε solves the system in (2.9)–(2.10) for the given (u, φ). The key result then needed in order to complete the proof is to argue that the system admits a unique solution for the given choice of (u, φ), thereby proving $(\bar{\xi}, \bar{\pi}) = (\xi, \pi)$ a.s. Although proving such a result for an arbitrary ξ and an arbitrary $(u, \varphi, \pi) \in \mathcal{V}(\xi)$ appears to be challenging, in this section we show that one can perturb ξ slightly to ξ^*, without affecting the cost too much, and find a near optimal $(u^*, \varphi^*, \pi^*) \in \mathcal{V}(\xi^*)$ such that the desired uniqueness property discussed above does in fact hold for (u^*, φ^*). See in particular parts 4 and 5 of the following proposition.

Proposition 4.1. *Let $\xi \in C([0,T] : \mathbb{R}^d)$ be such that $I(\xi) < \infty$. Fix $\gamma \in (0,1)$. Then there exists $\xi^* \in C([0,T] : \mathbb{R}^d)$ such that*

$$\|\xi - \xi^*\|_T \doteq \sup_{0\leq s\leq T}\|\xi(s) - \xi^*(s)\| < \gamma, \tag{4.1}$$

and there is $(u^*, \varphi^* = (\varphi^*_{ij}), \pi^* = (\pi^*_i)) \in \mathcal{V}(\xi^*)$ with the following properties.

1. *For some constants $m_2, m_3 \in (0,\infty)$ and all $(s,z) \in [0,T]\times[0,\zeta]$ and $(i,j) \in \mathbb{T}$,*
$$m_3 \geq \varphi^*_{ij}(s,z) \geq m_2.$$
2. *There is a measurable map $\varrho : [0,T]\times\mathbb{R}^d \to \mathcal{P}(\mathbb{L})$ such that for all $(s,x) \in [0,T]\times\mathbb{R}^d$*
$$\sum_{i\in\mathbb{L}} \varrho_i(s,x)\Gamma^{\varphi^*_i(s,\cdot)}_{ij}(x) = 0,$$
and for some $c_1 \in (0,\infty)$
$$\sup_{s\in[0,T]}\max_{i\in\mathbb{L}}|\varrho_i(s,x) - \varrho_i(s,\tilde{x})| \leq c_1\|x-\tilde{x}\|, \text{ for all } x,\tilde{x}\in\mathbb{R}^d,$$
$$\inf_{(s,x)\in[0,T]\times\mathbb{R}^d}\min_{i\in\mathbb{L}}\varrho_i(s,x) \geq c_1^{-1}. \tag{4.2}$$

3. *If for any* $(s,x) \in [0,T] \times \mathbb{R}^d$, $\bar{\pi} \in \mathcal{P}(\mathbb{L})$ *satisfies*

$$\sum_{i \in \mathbb{L}} \bar{\pi}_i \Gamma_{ij}^{\varphi_i^*(s,\cdot)}(x) = 0,$$

then $\bar{\pi} = \varrho(s,x)$. *In particular,* $\varrho(s,\xi^*(s)) = \pi^*(s)$.

4. *If for the given* u^* *and* φ^*, (2.9) *and* (2.10) *are satisfied for any other* $(\tilde{\xi},\tilde{\pi}) \in C([0,T]:\mathbb{R}^d) \times M([0,T]:\mathcal{P}(\mathbb{L}))$, *then* $(\tilde{\xi},\tilde{\pi}) = (\xi^*,\pi^*)$.

5. *The cost associated with* (u^*,φ^*) *satisfies:*

$$\sum_i \frac{1}{2}\int_0^T \|u_i^*(s)\|^2 \pi_i^*(s)ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}^*(s,z))\pi_i^*(s)\lambda_\zeta(dz)ds \le I(\xi) + \gamma. \tag{4.3}$$

Remark 4.2. We now give an outline of the proof strategy. Let $\xi \in C([0,T]:\mathbb{R}^d)$ be such that $I(\xi) < \infty$ and fix $\gamma \in (0,1)$. Let $(u,\varphi,\pi) \in \mathcal{V}(\xi)$ be such that

$$\sum_i \frac{1}{2}\int_0^T \|u_i(s)\|^2 \pi_i(s)ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds \le I(\xi) + \frac{\gamma}{2}. \tag{4.4}$$

Note that there are four time dependent objects appearing in the limit deterministic controlled dynamics: the trajectory ξ, the empirical measure on the fast variables π, the controls u that correspond to shifting the mean of the Brownian noises, and the thinning function φ that controls the rates for the fast variables. In addition, there is complete coupling of the fast and slow variables, and in particular the dynamics of the fast variables at time s will depend on both the controlled rates and ξ at s. The key issue regarding uniqueness is to make sure that these thinning functions φ can be bounded away from zero, which will imply ergodicity of the associated Markov processes giving the uniqueness of the corresponding π. If for a given collection (u,φ,π,ξ) the rates are not bounded away from zero, then we must show they can be perturbed so this is true, while at the same time making only a small change in ξ and the cost.

The steps are as follows. (a) We first perturb π to π^δ (see (4.6)), so that every state has strictly positive mass under π^δ. This positivity is used crucially in the remaining steps. (b) Replacing π with π^δ in (2.9) leads to a perturbation of the target trajectory ξ. To ensure that the trajectory perturbation is not too large, we modify the control u to u^δ in a way that compensates for the change in π (see (4.7)). (c) The perturbed measure π^δ need not be stationary (i.e., satisfy (2.10)) for the original thinning control φ and the new trajectory ξ^δ. In order to remedy this we next perturb φ to φ^δ (see (4.9)). (d) With the perturbed φ^δ (2.10) is satisfied so the π^δ would be stationary, but with ξ rather than ξ^δ. In particular the constructed $(u^\delta,\varphi^\delta,\pi^\delta)$ is not in general in $\mathcal{V}(\xi^\delta)$. This leads to our last modification where we change φ^δ to $\tilde{\varphi}^\delta$. It is at this point that the formulation of the original dynamics as the solution to an SDE driven by a collection of PRMs, and corresponding formulation of the control problem in terms of thinning functions, is very convenient (see (4.13)). With this change we now have a $(u^\delta,\tilde{\varphi}^\delta,\pi^\delta) \in \mathcal{V}(\xi^\delta)$. Furthermore with $\delta = \delta^*$ sufficiently small these perturbed quantities $(u^{\delta^*},\tilde{\varphi}^{\delta^*},\pi^{\delta^*},\xi^{\delta^*}) = (u^*,\varphi^*,\pi^*,\xi^*)$ will satisfy all the desired properties.

Proof of Proposition 4.1. Let ξ and $(u,\varphi,\pi) \in \mathcal{V}(\xi)$ be as in Remark 4.2. In particular,

$$\sum_i \frac{1}{2}\int_0^T \|u_i(s)\|^2 \pi_i(s)ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds \le I(\xi) + \frac{\gamma}{2}.$$

We claim that without loss of generality it can be assumed that for some $m_0 \in (0,\infty)$,

$$\sup_{(s,z)\in[0,T]\times[0,\zeta]} \max_{(i,j)\in\mathbb{T}} [\varphi_{ij}(s,z)\pi_i(s)] \le m_0. \tag{4.5}$$

We first prove the proposition assuming the claim. The proof of the claim is given at the end.

For $x \in \mathbb{R}^d$ let $\nu(x)$ as in Theorem 2.4 be the stationary distribution for the fast system when the slow variable equals x. Fix $\delta > 0$ and define

$$\pi_j^\delta(s) \doteq (1-\delta)\pi_j(s) + \delta\nu_j(\xi(s)). \tag{4.6}$$

Note that

$$\sup_{s\in[0,T]} \sum_j |\pi_j^\delta(s) - \pi_j(s)| \le 2\delta.$$

Define

$$u_j^\delta(s) \doteq u_j(s)\frac{\pi_j(s)}{\pi_j^\delta(s)},\ s\in[0,T], j\in\mathbb{L}. \tag{4.7}$$

Then $u_j^\delta(s)\pi_j^\delta(s) = u_j(s)\pi_j(s)$ for all s and j. Define

$$\xi^\delta(t) = x_0 + \sum_j \int_0^t b_j(\xi^\delta(s))\pi_j^\delta(s)ds + \sum_j \int_0^t a_j(\xi^\delta(s))u_j^\delta(s)\pi_j^\delta(s)ds. \tag{4.8}$$

From the Lipschitz properties of b_j, a_j, (4.8) has a unique solution for the given u^δ and π^δ. Note that with $M = I(\xi)+1$,

$$\int_0^T \left(\sum_j \pi_j^\delta(s)(|u_j^\delta(s)|+1)\right) ds \le (T + (2TM)^{1/2}) \doteq \Theta(M).$$

Then by Gronwall's lemma

$$\|\xi - \xi^\delta\|_T \le K\delta,$$

where $K = 2M_0\Theta(M)e^{d_{\text{lip}}\Theta(M)}$ and $M_0 = \sup_{0\le s\le T, j\in\mathbb{L}}(\|b_j(\xi(s))\| + \|a_j(\xi(s))\|)$. Now define, for $(i,j)\in\mathbb{T}$ and $(s,z)\in[0,T]\times[0,\zeta]$,

$$\varphi_{ij}^\delta(s,z) \doteq (1-\delta)\frac{\pi_i(s)}{\pi_i^\delta(s)}\varphi_{ij}(s,z) + \delta\frac{\nu_i(\xi(s))}{\pi_i^\delta(s)} \tag{4.9}$$

and

$$\Gamma_{ij}^\delta(s) \doteq \int_{E_{ij}(\xi(s))} \varphi_{ij}^\delta(s,z)\lambda_\zeta(dz),\ s\in[0,T], i\ne j, \tag{4.10}$$

and $\Gamma_{ii}^\delta(s) \doteq -\sum_{j:j\ne i}\Gamma_{ij}^\delta(s)$. Then since the π_i^δ will cancel and $(u,\varphi,\pi)\in\mathcal{V}(\xi)$,

$$\pi^\delta(s)\Gamma^\delta(s) = 0. \tag{4.11}$$

Thus $\pi^\delta(s)$ is stationary for $\Gamma^\delta(s)$. However, from (2.6) $\Gamma_{ij}^\delta(s) = \Gamma_{ij}^{\varphi_i^\delta(s,\cdot)}(\xi(s))$, and hence $(u^\delta,\varphi^\delta,\pi^\delta)$ is not in general in $\mathcal{V}(\xi^\delta)$. We now construct a further modification, $\tilde\varphi^\delta$, of φ^δ such that $(u^\delta,\tilde\varphi^\delta,\pi^\delta)\in\mathcal{V}(\xi^\delta)$, and such that the uniqueness of (2.9)–(2.10) (with (u,φ) replaced by $(u^\delta,\tilde\varphi^\delta)$) holds.

Let

$$\Gamma_{ij}(x) \doteq c_i(x)r_{ij}(x) = \lambda_\zeta(E_{ij}(x)), \tag{4.12}$$

where we recall that $c_i(x)$ is the overall rate of transitions out of i for the fast process when the slow process is in state x, and $r_{ij}(x)$ gives the probability of transition to state j. For $(i,j) \in \mathbb{T}$ define

$$\tilde{\varphi}_{ij}^{\delta}(s,z) = \begin{cases} \frac{\Gamma_{ij}^{\delta}(s)}{\Gamma_{ij}(\xi^{\delta}(s))}, & z \in E_{ij}(\xi^{\delta}(s)), \\ 1 & \text{otherwise .} \end{cases} \tag{4.13}$$

Then for such (i,j)

$$\Gamma_{ij}^{\delta}(s) = \int_{E_{ij}(\xi(s))} \varphi_{ij}^{\delta}(s,z)\lambda_{\zeta}(dz) = \int_{E_{ij}(\xi^{\delta}(s))} \tilde{\varphi}_{ij}^{\delta}(s,z)\lambda_{\zeta}(dz).$$

Then, from (4.8) and (4.11), $(u^{\delta}, \tilde{\varphi}^{\delta}, \pi^{\delta}) \in \mathcal{V}(\xi^{\delta})$ and $\Gamma_{ij}^{\delta}(s) = \Gamma_{ij}^{\tilde{\varphi}_i^{\delta}(s,\cdot)}(\xi^{\delta}(s))$, where $\Gamma_{ij}^{\psi}(x)$ was defined in (2.6), and $\tilde{\varphi}_i^{\delta}$ denotes the collection of controls $(\tilde{\varphi}_{ij}^{\delta}, j \in \mathbb{L})$. Next note that by construction, for $(i,j) \in \mathbb{T}$,

$$\varphi_{ij}^{\delta}(s,z) \geq \delta \frac{\nu_i(\xi(s))}{\pi_i^{\delta}(s)} \geq \delta \underline{\nu}, \text{ for all } (s,z) \in [0,T] \times [0,\zeta]$$

and, from (4.5), for all $\delta > 0$

$$\varphi_{ij}^{\delta}(s,z)\pi_i^{\delta}(s) \leq m_0 + 1 \doteq m_1.$$

Let $\underline{r} \doteq \underline{\varsigma}\kappa_3$, where $\underline{\varsigma}$ and κ_3 are defined above Assumption 2.3. Then for $(i,j) \in \mathbb{T}$ the definition (4.10) implies

$$\frac{m_1\zeta}{\delta\underline{\nu}} \geq \Gamma_{ij}^{\delta}(s) \geq \delta\underline{\nu}\,\underline{r}.$$

Also, from (4.12), for each s

$$\int_{[0,\zeta]} \ell(\tilde{\varphi}_{ij}^{\delta}(s,z))\lambda_{\zeta}(dz) = \ell\left(\frac{\Gamma_{ij}^{\delta}(s)}{\Gamma_{ij}(\xi^{\delta}(s))}\right)\Gamma_{ij}(\xi^{\delta}(s))$$

and using convexity

$$\int_{[0,\zeta]} \ell(\varphi_{ij}^{\delta}(s,z))\lambda_{\zeta}(dz) \geq \ell\left(\frac{\Gamma_{ij}^{\delta}(s)}{\Gamma_{ij}(\xi(s))}\right)\Gamma_{ij}(\xi(s)).$$

It is easy to check that for $a \geq 0$ and $b, c > 0$,

$$\left|\ell\left(\frac{a}{b}\right)b - \ell\left(\frac{a}{c}\right)c\right| \leq \left(1 + \frac{a}{b \wedge c}\right)|b - c|.$$

The Lipschitz properties of the underlying transition rates $\Gamma_{ij}(\cdot)$ (see (2.2)) yield the following inequalities, each of which is explained after the display:

$$\begin{aligned}
&\sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \pi_i^{\delta}(s)\ell(\tilde{\varphi}_{ij}^{\delta}(s,z))\lambda_{\zeta}(dz)ds - \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \pi_i^{\delta}(s)\ell(\varphi_{ij}^{\delta}(s,z))\lambda_{\zeta}(dz)ds \\
&\leq \kappa_2\|\xi - \xi^{\delta}\|_T \int_{[0,T]} \sum_{(i,j)\in\mathbb{T}} \pi_i^{\delta}(s)\left(\frac{\Gamma_{ij}^{\delta}(s)}{\underline{r}} + 1\right)ds \\
&\leq \kappa_2 T|\mathbb{L}|\|\xi - \xi^{\delta}\|_T + \frac{\kappa_2}{\underline{r}}\|\xi - \xi^{\delta}\|_T \int_{[0,\zeta]\times[0,T]} \sum_{(i,j)\in\mathbb{T}} \pi_i^{\delta}(s)\varphi_{ij}^{\delta}(s,z)\lambda_{\zeta}(dz)ds \\
&\leq \kappa_2 T|\mathbb{L}|\|\xi - \xi^{\delta}\|_T + \frac{\kappa_2}{\underline{r}}\|\xi - \xi^{\delta}\|_T \int_{[0,\zeta]\times[0,T]} \left(|\mathbb{L}| + \sum_{(i,j)\in\mathbb{T}} \pi_i(s)\varphi_{ij}(s,z)\right)\lambda_{\zeta}(dz)ds \\
&\leq \delta K_1.
\end{aligned}$$

The first inequality uses the previous three displays, the second uses the definition of Γ^{δ}_{ij} in (4.10), the third uses (4.9), and the final one uses the definition

$$K_1 \doteq K\kappa_2 \left(T|\mathbb{L}| + \frac{1}{\underline{r}} \left(|\mathbb{L}|\zeta T(1+e) + eM \right) \right),$$

$\|\xi - \xi^{\delta}\|_T \leq \delta$, and the fact that $x \leq e(1+\ell(x))$ for all $x \geq 0$. Hence we obtain

$$\begin{aligned}
&\int_{[0,\zeta]\times[0,T]} \sum_{(i,j)\in\mathbb{T}} \pi^{\delta}_i(s)\ell(\tilde{\varphi}^{\delta}_{ij}(s,z))\lambda_{\zeta}(dz)ds \\
&\quad \leq \int_{[0,\zeta]\times[0,T]} \sum_{(i,j)\in\mathbb{T}} \pi^{\delta}_i(s)\ell(\varphi^{\delta}_{ij}(s,z))\lambda_{\zeta}(dz)ds + K_1\delta \\
&\quad \leq \int_{[0,\zeta]\times[0,T]} \sum_{(i,j)\in\mathbb{T}} \pi_i(s)\ell(\varphi_{ij}(s,z))\lambda_{\zeta}(dz)ds + K_1\delta,
\end{aligned}$$

where the last line is a consequence of the fact that

$$\begin{aligned}
\pi^{\delta}_i(s)\ell(\tilde{\varphi}^{\delta}_{ij}(s,z)) &= \pi^{\delta}_i(s)\ell\left((1-\delta)\frac{\pi_i(s)}{\pi^{\delta}_i(s)}\varphi_{ij}(s,z) + \delta\frac{\nu_i(\xi(s))}{\pi^{\delta}_i(s)} \right) \\
&\leq \pi^{\delta}_i(s)(1-\delta)\frac{\pi_i(s)}{\pi^{\delta}_i(s)}\ell(\varphi_{ij}(s,z)) \leq \pi_i(s)\ell(\varphi_{ij}(s,z)).
\end{aligned} \tag{4.14}$$

Next note that if $\pi_i(s) \leq \pi^{\delta}_i(s)$ then

$$\|u^{\delta}_i(s)\|^2\pi^{\delta}_i(s) = \|u_i(s)\|^2\frac{\pi_i(s)}{\pi^{\delta}_i(s)}\pi_i(s) \leq \|u_i(s)\|^2\pi_i(s).$$

However if $\pi_i(s) \geq \pi^{\delta}_i(s)$ then $\pi_i(s) \geq \nu_i(\xi(s))$ follows, and thus $\pi^{\delta}_i(s) \geq \underline{\nu}$. Therefore

$$\left| \frac{\pi_i(s)}{\pi^{\delta}_i(s)} - 1 \right| \leq \frac{2\delta}{\pi^{\delta}_i(s)} \leq \frac{2\delta}{\underline{\nu}}.$$

Thus in this case

$$\|u^{\delta}_i(s)\|^2\pi^{\delta}_i(s) \leq \|u_i(s)\|^2\pi_i(s) + \frac{2\delta}{\underline{\nu}}\|u_i(s)\|^2\pi_i(s).$$

Combining the two cases we have from (4.4)

$$\frac{1}{2}\int_0^T \sum_i \|u^{\delta}_i(s)\|^2\pi^{\delta}_i(s)ds \leq \frac{1}{2}\int_0^T \sum_i \|u_i(s)\|^2\pi_i(s)ds + \frac{2\delta}{\underline{\nu}}M. \tag{4.15}$$

Taking $\delta^* \doteq \min\{\gamma/K, \gamma/4K_1, \gamma\underline{\nu}/8M\}$ we now see that with

$$(\xi^*, u^*, \varphi^*, \pi^*) \doteq (\xi^{\delta^*}, u^{\delta^*}, \tilde{\varphi}^{\delta^*}, \pi^{\delta^*}),$$

$(u^*, \varphi^*, \pi^*) \in \mathcal{V}(\xi^*)$. Also, for all $(s,z) \in [0,T]\times[0,\zeta]$ and $i \in \mathbb{L}$,

$$m_3 \doteq \frac{m_1\zeta}{\delta^*\underline{r}\,\underline{\nu}} \vee 1 \geq \varphi^*_{ij}(s,z) \geq \frac{\delta^*\underline{\nu}\,\underline{r}}{\zeta} \wedge 1 \doteq m_2, \tag{4.16}$$

namely item 1 in the proposition holds. From (4.16) and the definition (2.6), for $(i,j) \in \mathbb{T}$, $s \in [0,T]$ and $x \in \mathbb{R}^d$

$$\Gamma^{\varphi^*_i(s,\cdot)}_{ij}(x) \in [m_2\underline{r}, m_3\zeta]. \tag{4.17}$$

Also, from (2.2), for all $x, x' \in \mathbb{R}^d$,

$$|\Gamma_{ij}^{\varphi_i^*(s,\cdot)}(x) - \Gamma_{ij}^{\varphi_i^*(s,\cdot)}(x')| \le m_3\kappa_2\|x - x'\|. \tag{4.18}$$

Using (4.17) and (4.18) it follows as in the proof of Lemma 2.5 that items 2 and 3 in the proposition hold. Next if (2.9) (with u replaced by u^*) and (2.10) (with φ replaced by φ^*) are satisfied for any other $(\tilde{\xi}, \tilde{\pi}) \in C([0,T] : \mathbb{R}^d) \times \mathbb{M}([0,T] : \mathcal{P}(\mathbb{L}))$, then we must have from item 3 in the proposition that $\tilde{\pi}(s) = \varrho(s, \tilde{\xi}(s))$ for all $s \in [0,T]$. The Lipschitz property of b, a and ϱ then gives that $\xi^*(t) = \tilde{\xi}(t)$, and thus also $\pi^*(t) = \tilde{\pi}(t)$, for all $t \in [0,T]$. This completes the proof of item 4. Finally consider item 5. Note that from (4.14), (4.15) and the choice of δ^*,

$$\begin{aligned}
&\frac{1}{2}\sum_i \int_0^T \|u_i^*(s)\|^2 \pi_i^*(s) ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}^*(s,z))\pi_i^*(s)\lambda_\zeta(dz)ds \\
&\le \sum_i \frac{1}{2}\int_0^T \|u_i(s)\|^2 \pi_i(s) ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}(s,z))\pi_i(s)\lambda_\zeta(dz)ds + \gamma/2 \\
&\le I(\xi) + \gamma,
\end{aligned}$$

where the last line is from (4.4). This proves 5. We now prove the claim made in (4.5). Note that we do not change the dynamics at all if we redefine φ_{ij} in the following way. With φ on the right equal to the old version and φ on the left the new, for $(i,j) \in \mathbb{T}$ set

$$\varphi_{ij}(s,z) = \begin{cases} \frac{\Gamma_{ij}^{\varphi_i(s,\cdot)}(\xi(s))}{\Gamma_{ij}(\xi(s))}, & z \in E_{ij}(\xi(s)), \\ \\ 1, & z \in [0,\zeta] \setminus E_{ij}(\xi(s)). \end{cases}$$

This amounts to assuming that outside $E_{ij}(\xi(s))$ the controlled jump rates are the same as for the original system, and that within $E_{ij}(\xi(s))$ they are constant in z, in such a way the overall jump rates do not change. Owing to convexity of ℓ and $\ell(1) = 0$, this can only lower the cost while preserving the dynamics, and could have been assumed for any candidate control for the jumps from the outset. Let

$$v(s) \doteq 1 + \sum_{(i,j)\in\mathbb{T}} \pi_i(s)\Gamma_{ij}^{\varphi_i}(\xi(s)),\ \bar{\Gamma}_{ij}^{\varphi_i}(\xi(s)) \doteq \Gamma_{ij}^{\varphi_i}(\xi(s))/v(s) \tag{4.19}$$

and for $\alpha > 0$ and $(i,j) \in \mathbb{T}$

$$\varphi_{ij}^\alpha(s,z) = \begin{cases} \alpha\frac{\bar{\Gamma}_{ij}^{\varphi_i}(\xi(s))}{\Gamma_{ij}(\xi(s))}, & z \in E_{ij}(\xi(s)), \\ \\ \frac{\alpha}{v(s)}, & z \in [0,\zeta] \setminus E_{ij}(\xi(s)). \end{cases}$$

Thus for $\alpha > 0$ the controlled jump rates are uniformly scaled by $\alpha/v(s)$, and therefore $\sum_{i\in\mathbb{L}} \pi_i(s)\Gamma^{\varphi_i^\alpha(s,\cdot)}(\xi(s)) = 0$. Since with $\alpha = v(s)$, $\varphi^\alpha = \varphi$,

$$\inf_{\alpha>0} \sum_{(i,j)\in\mathbb{T}} \pi_i(s) \int_{[0,\zeta]} \ell(\varphi_{ij}^\alpha(s,z))\lambda_\zeta(dz) \le \sum_{(i,j)\in\mathbb{T}} \pi_i(s) \int_{[0,\zeta]} \ell(\varphi_{ij}(s,z))\lambda_\zeta(dz).$$

We now compute the infimum on the left side. For notational simplicity, write $\bar{\Gamma}_{ij}^{\varphi_i(s,\cdot)}(\xi(s))$ as $\bar{\Gamma}_{ij}$, and $\Gamma_{ij}(\xi(s))$ as Γ_{ij}. Also let $\theta_{ij} \doteq \zeta - \lambda_\zeta(E_{ij}(\xi(s)))$. Note that $1 \le \theta_{ij} \le \zeta$. Differentiating with respect to α and setting the derivative to 0, we get

$$\begin{aligned}
&\log(\alpha) \sum_{(i,j)\in\mathbb{T}} \pi_i(s)\left(\bar{\Gamma}_{ij} + \frac{\theta_{ij}}{v(s)}\right) \\
&\quad = -\sum_{(i,j)\in\mathbb{T}} \pi_i(s)\left(\bar{\Gamma}_{ij}\log\left(\frac{\bar{\Gamma}_{ij}}{\Gamma_{ij}}\right) + \frac{\theta_{ij}}{v(s)}\log\left(\frac{1}{v(s)}\right)\right).
\end{aligned}$$

Thus

$$\log(\alpha) = -\frac{\sum_{(i,j)\in\mathbb{T}} \pi_i(s) \left(\bar{\Gamma}_{ij} \log\left(\frac{\bar{\Gamma}_{ij}}{\Gamma_{ij}}\right) + \frac{\theta_{ij}}{v(s)} \log\left(\frac{1}{v(s)}\right)\right)}{\sum_{(i,j)\in\mathbb{T}} \pi_i(s) \left(\bar{\Gamma}_{ij} + \frac{\theta_{ij}}{v(s)}\right)}.$$

It is easily checked that there is $m_6 \in (0,\infty)$ such that for all $s \in [0,T]$ the minimizing α satisfies $\alpha \le m_6$. Thus from the definition of φ^α and since from (4.19) $\pi_i(s)\bar{\Gamma}_{ij}(\xi(s)) \le 1$, we see that with this α, for all $(i,j) \in \mathbb{T}$, $(s,z) \in [0,T]\times[0,\zeta]$

$$\pi_i(s)\varphi^\alpha_{ij}(s,z) \le \max\{m_6/\underline{r}, m_6\}.$$

This proves the claim in (4.5) and completes the proof of the proposition. □

5 Large deviation lower bound

The goal of this section is to prove the following theorem.

Theorem 5.1. *For any $F \in C_b(C([0,T]:\mathbb{R}^d))$, the inequality in* (2.13) *holds, namely*

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \ge -\inf_{\xi\in C([0,T]:\mathbb{R}^d)}\{F(\xi)+I(\xi)\}, \tag{5.1}$$

where X^ε solves (2.3)*, and I is as defined in* (2.18)*.*

To do this we will show an upper bound on the corresponding variational representations. The situation is in some sense simpler than the corresponding lower bound, since a fixed control is being used. Thus the analysis is essentially just the law of large numbers for a two time scale system, with the main effort being to justify the replacement of the empirical measure of the fast component by the corresponding stationary measure in the limit. We begin by stating an elementary but useful lemma, whose proof we defer to the Appendix.

Lemma 5.2. *Let m_n, m be finite measures on $[0,T]\times\mathbb{L}$ such that the first marginal of m_n and m is Lebesgue measure:*

$$m_n([a,b]\times\mathbb{L}) = m([a,b]\times\mathbb{L}) = b-a \text{ for all } 0 \le a \le b \le T.$$

Suppose that m_n converges weakly to m. Let $v : [0,T] \to \mathbb{R}$ be an integrable map, i.e., $\int_0^T |v(s)|ds < \infty$. Then

$$\int_{\mathbb{L}\times[0,T]} v(s)1_{\{j\}}(y)m_n(dy\times ds) \to \int_{\mathbb{L}\times[0,T]} v(s)1_{\{j\}}(y)m(dy\times ds). \tag{5.2}$$

Proof of Theorem 5.1. Without loss of generality we can assume that F is Lipschitz continuous [10, Corollary 1.2.5], i.e., for some $C_F \in (0,\infty)$,

$$|F(\xi) - F(\tilde{\xi})| \le C_F\|\xi-\tilde{\xi}\|_T \text{ for all } \xi,\tilde{\xi} \in C([0,T]:\mathbb{R}^d).$$

Fix $\gamma \in (0,1)$. Let $\xi \in C([0,T]:\mathbb{R}^d)$ be such that

$$I(\xi) + F(\xi) \le \inf_{\zeta\in C([0,T]:\mathbb{R}^d)}[I(\zeta)+F(\zeta)] + \gamma. \tag{5.3}$$

Since $I(\xi) < \infty$, we can find $\xi^* \in C([0,T]:\mathbb{R}^d)$, $(u^*,\varphi^*,\pi^*) \in \mathcal{V}(\xi^*)$ and a measurable $\varrho : [0,T]\times\mathbb{R}^d \to \mathcal{P}(\mathbb{L})$ with properties stated in Proposition 4.1. Let $(\bar{X}^\varepsilon, \bar{Y}^\varepsilon)$ be defined through (3.4), where the controls $(\psi^\varepsilon,\varphi^\varepsilon) \in \mathcal{U}$ are defined in feedback form as

$$\psi^\varepsilon(s) \doteq \sum_{j=1}^{|\mathbb{L}|} 1_{\{\bar{Y}^\varepsilon(s-)=j\}}u^*_j(s), \quad \varphi^\varepsilon_{ij}(s) \doteq \varphi^*_{ij}(s)1_{\{\bar{Y}^\varepsilon(s-)=i\}} + 1_{\{\bar{Y}^\varepsilon(s-)\neq i\}}, \ s\in[0,T], (i,j)\in\mathbb{T}. \tag{5.4}$$

From (3.5) we get that

$$-\varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] \le \mathbb{E}\left[\tilde{L}_T(\psi^\varepsilon) + \bar{L}_T(\varphi^\varepsilon) + F(\bar{X}^\varepsilon)\right],$$

where $\tilde{L}_T$ and $\bar{L}_T$ were defined at the beginning of Section 3. Define $Q^\varepsilon \in \mathcal{P}_{\text{leb}}(\mathbb{H}_T)$ by (3.11) where η^ε is as introduced in (3.10). From (4.3) and (4.2) (note that (4.2) gives a lower bound on $\pi_i^*(s)$) it follows that for all $\varepsilon > 0$

$$\tilde{L}_T(\psi^\varepsilon) + \bar{L}_T(\varphi^\varepsilon) \le c_1(I(\xi)+1) \doteq M < \infty. \tag{5.5}$$

It then follows from Proposition 3.4 that $(\bar{X}^\varepsilon, Q^\varepsilon)$ is a tight family of $C([0,T]:\mathbb{R}^d) \times \mathcal{M}_F(\mathbb{H}_T)$-valued random variables, and if $(\bar{\xi}, \bar{Q})$ is a weak limit point of $\{(\bar{X}^\varepsilon, Q^\varepsilon)\}$ then equations (2.16) and (2.17) hold with (ξ, Q) replaced with $(\bar{\xi}, \bar{Q})$. Disintegrate $\bar{Q}$ as

$$\bar{Q}(ds \times \{y\} \times d\eta \times dz) = ds\bar{\pi}_y(s)[\bar{Q}]_{34|12}(d\eta \times dz). \tag{5.6}$$

We will now show that, a.s.,

$$(\bar{\xi}(s), \bar{\pi}(s)) = (\xi^*(s), \pi^*(s)) \text{ for a.e. } s \in [0,T]. \tag{5.7}$$

We can assume without loss of generality that convergence of $(\bar{X}^\varepsilon, Q^\varepsilon)$ to $(\bar{\xi}, \bar{Q})$ holds a.s. along the full sequence. We begin by showing that for every $y \in \mathbb{L}$, $t \in [0,T]$, and any continuous map $h : [0,T] \to \mathbb{R}^m$, with $h(s)'$ denoting the transpose,

$$\int_{\mathbb{H}_t} h(s)' 1_{\{j\}}(y) z \bar{Q}(d\mathbf{v}) = \int_0^t h(s)' u_j^*(s) \bar{\pi}_j(s) ds. \tag{5.8}$$

As in the proof of Proposition 3.4, using (5.5) we have that for all $t \in [0,T]$ and $j \in \mathbb{L}$

$$\int_{\mathbb{H}_t} h(s)' 1_{\{j\}}(y) z Q^\varepsilon(d\mathbf{v}) \to \int_{\mathbb{H}_t} h(s)' 1_{\{j\}}(y) z \bar{Q}(d\mathbf{v}). \tag{5.9}$$

The left side of (5.9) equals

$$\int_{[0,t]} h(s)' 1_{\{\bar{Y}^\varepsilon(s)=j\}} u_j^*(s) ds = \int_{[0,t]\times\mathbb{L}} h(s)' u_j^*(s) 1_{\{j\}}(y) [Q^\varepsilon]_{12}(dy \times ds).$$

From Lemma 5.2 and the fact that u_j^* is square integrable, this converges a.s. to

$$\int_{[0,t]\times\mathbb{L}} h(s)' u_j^*(s) 1_{\{j\}}(y) [\bar{Q}]_{12}(dy \times ds) = \int_{[0,t]} h(s)' u_j^*(s) \bar{\pi}_j(s) ds,$$

where the equality follows from (5.6). This proves (5.8).

Applying (5.8) to the rows of $a_j(\bar{\xi}(s))$ and summing over j, we have that

$$\int_{\mathbb{H}_t} a(\bar{\xi}(s), y) z \bar{Q}(d\mathbf{v}) = \sum_{j=1}^{|\mathbb{L}|} \int_{[0,t]} a_j(\bar{\xi}(s)) u_j^*(s) \bar{\pi}_j(s) ds.$$

Also, recalling the representation of $\bar{Q}$ in (5.6)

$$\int_{\mathbb{H}_t} b(\bar{\xi}(s), y) \bar{Q}(d\mathbf{v}) = \sum_{j=1}^{|\mathbb{L}|} \int_{[0,t]} b_j(\bar{\xi}(s)) \bar{\pi}_j(s) ds.$$

Thus we have shown that (2.9) is satisfied with (ξ, π, u) replaced with $(\bar{\xi}, \bar{\pi}, u^*)$.

We next show that (2.10) is satisfied as well, i.e.,

$$\sum_{i\in\mathbb{L}} \bar{\pi}_i(s)\Gamma_{ij}^{\varphi^*_{i,\cdot}(s,\cdot)}(\bar{\xi}(s)) = 0 \text{ for a.e. } s\in[0,T] \text{ and all } j\in\mathbb{L}. \tag{5.10}$$

For this recall that (2.17) is satisfied with (ξ, Q) replaced with $(\bar{\xi}, \bar{Q})$, i.e.,

$$\int_{\mathbb{H}_t} \Gamma^{\eta}_{y,j}(\bar{\xi}(s))\bar{Q}(d\mathbf{v}) = 0. \tag{5.11}$$

Also, from Lemma 2.13 (see, e.g., the proof of (2.30)), for any continuous function $g:[0,T]\to\mathbb{R}^d$,

$$\lim_{\varepsilon\to 0}\int_{\mathbb{H}_t}\eta_j(E_{yj}(g(s)))Q^{\varepsilon}(d\mathbf{v}) = \int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(g(s))\bar{Q}(d\mathbf{v}), \tag{5.12}$$

Furthermore, for every $t\in[0,T]$,

$$\int_{\mathbb{H}_t}\eta_j(E_{yj}(g(s)))Q^{\varepsilon}(d\mathbf{v}) = \sum_{i\in\mathbb{L}}\int_0^t 1_{\{\bar{Y}^{\varepsilon}(s)=i\}}\int_{E_{ij}(g(s))}\varphi^*_{ij}(s,z)\lambda_{\zeta}(dz)ds. \tag{5.13}$$

From the equalities in (5.12) and (5.13) it now follows that

$$\sum_{i\in\mathbb{L}}\int_{[0,t]} 1_{\{\bar{Y}^{\varepsilon}(s)=i\}}\int_{E_{ij}(g(s))}\varphi^*_{ij}(s,z)\lambda_{\zeta}(dz)ds \to \int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(g(s))\bar{Q}(d\mathbf{v})$$

as $\varepsilon\to 0$. However, the expression on the left side in the previous display is the same as

$$\int_{[0,t]\times\mathbb{L}}\left(\int_{E_{yj}(g(s))}\varphi^*_{yj}(s,z)\lambda_{\zeta}(dz)\right)[Q^{\varepsilon}]_{12}(dy\times ds).$$

From (5.5)

$$\int_{[0,T]}\left(\int_{[0,\zeta]}\varphi^*_{yj}(s,z)\lambda_{\zeta}(dz)\right)ds < \infty.$$

Thus from Lemma 5.2

$$\begin{aligned}
&\int_{[0,t]\times\mathbb{L}}\left(\int_{E_{yj}(g(s))}\varphi^*_{yj}(s,z)\lambda_{\zeta}(dz)\right)[Q^{\varepsilon}]_{12}(dy\times ds)\\
&\quad\to\int_{[0,t]\times\mathbb{L}}\left(\int_{E_{yj}(g(s))}\varphi^*_{yj}(s,z)\lambda_{\zeta}(dz)\right)[\bar{Q}]_{12}(dy\times ds)\\
&\quad=\int_{[0,t]}\sum_{i\in\mathbb{L}}\bar{\pi}_i(s)\Gamma_{ij}^{\varphi^*_{i,\cdot}(s,\cdot)}(g(s))ds
\end{aligned}$$

for all $t\in[0,T]$. Thus we have shown that for every continuous $g:[0,T]\to\mathbb{R}^d$ and every $t\in[0,T]$,

$$\int_{\mathbb{H}_t}\Gamma^{\eta}_{y,j}(g(s))\bar{Q}(d\mathbf{v}) = \int_{[0,t]}\sum_{i\in\mathbb{L}}\bar{\pi}_i(s)\Gamma_{ij}^{\varphi^*_{i,\cdot}(s,\cdot)}(g(s))ds.$$

Taking $g=\bar{\xi}$ and using (5.11) we have (5.10).

We have therefore shown that both (2.9) and (2.10) hold with (ξ,π,u,φ) replaced by $(\bar{\xi},\bar{\pi},u^*,\varphi^*)$. Thus from part 4 of Proposition 4.1, we have (5.7).

Finally, we consider the convergence of costs. Note that

$$\begin{aligned}
\tilde{L}_T(\psi^{\varepsilon}) &= \frac{1}{2}\int_0^T\|\psi^{\varepsilon}(s)\|^2ds = \frac{1}{2}\sum_{j\in\mathbb{L}}\int_0^T 1_{\{\bar{Y}^{\varepsilon}(s)=j\}}\|u^*_j(s)\|^2ds\\
&= \frac{1}{2}\sum_{j\in\mathbb{L}}\int_{[0,T]\times\mathbb{L}}\|u^*_j(s)\|^2 1_{\{j\}}(y)[Q^{\varepsilon}]_{12}(dy\times ds)
\end{aligned}$$

which from Lemma 5.2 and the finiteness of $\int_{[0,T]} \|u_j^*(s)\|^2 ds$ converges to

$$\int_{[0,T]} \|u_j^*(s)\|^2 \pi_j^*(s) ds.$$

The second term in the cost can be handled in a similar manner. Note that

$$\begin{aligned}\bar{L}_T(\varphi^\varepsilon) &= \sum_{(i,j)\in\mathbb{T}} \int_{[0,T]\times[0,\zeta]} \ell(\varphi_{ij}^\varepsilon(s,z))\lambda_\zeta(dz)ds \\ &= \sum_{(i,j)\in\mathbb{T}} \int_{[0,T]\times[0,\zeta]} 1_{\{\bar{Y}^\varepsilon(s)=i\}} \ell(\varphi_{ij}^*(s,z))\lambda_\zeta(dz)ds.\end{aligned}$$

For $j \in \mathbb{L}$ write

$$\begin{aligned}&\sum_{i\in\mathbb{L}} \int_{[0,T]\times[0,\zeta]} 1_{\{\bar{Y}^\varepsilon(s)=i\}} \ell(\varphi_{ij}^*(s,z))\lambda_\zeta(dz)ds \\ &\quad= \int_{[0,T]\times\mathbb{L}} \left(\int_{[0,\zeta]} \ell(\varphi_{yj}^*(s,z))\lambda_\zeta(dz) \right) [Q^\varepsilon]_{12}(dy\times ds).\end{aligned}$$

Then using the fact that $\int_{[0,T]\times[0,\zeta]} \ell(\varphi_{ij}^*(s,z))\lambda_\zeta(dz)ds < \infty$ we see as before using Lemma 5.2 that the first term in the last display converges to

$$\int_{[0,T]\times[0,\zeta]} \ell(\varphi_{ij}^*(s,z))\pi_i^*(s)\lambda_\zeta(dz)ds.$$

Combining the above observations we have, as $\varepsilon \to 0$,

$$\bar{L}_T(\varphi^\varepsilon) \to \sum_{(i,j)\in\mathbb{T}} \int_{[0,T]\times[0,\zeta]} \ell(\varphi_{ij}^*(s,z))\pi_i^*(s)\lambda_\zeta(dz)ds$$

in probability. Using (5.5) the dominated convergence theorem gives

$$\begin{aligned}&\mathbb{E}\left[\tilde{L}_T(\psi^\varepsilon) + \bar{L}_T(\varphi^\varepsilon)\right] \\ &\to \sum_i \frac{1}{2}\int_0^T \|u_i^*(s)\|^2 \pi_i^*(s)ds + \sum_{(i,j)\in\mathbb{T}} \int_{[0,\zeta]\times[0,T]} \ell(\varphi_{ij}^*(s,z))\pi_i^*(s)\lambda_\zeta(dz)ds \\ &\le I(\xi) + \gamma.\end{aligned}$$

Using this and the convergence of $\bar{X}^\varepsilon$ to ξ^* (see (5.7)), we have from (5.3) that

$$\begin{aligned}\limsup_{\varepsilon\to 0} -\varepsilon \log \mathbb{E}\left[\exp\left(-\varepsilon^{-1}F(X^\varepsilon)\right)\right] &\le F(\xi^*) + I(\xi) + \gamma \\ &\le F(\xi) + I(\xi) + \gamma + C_F\gamma \\ &\le \inf_{\zeta\in C([0,T]:\mathbb{R}^d)} [I(\zeta) + F(\zeta)] + 2\gamma + C_F\gamma,\end{aligned}$$

where the second inequality uses (4.1) in Proposition 4.1. Sending $\gamma \to 0$ gives the desired lower bound in (5.1). □

Appendix

A Proofs of some auxiliary results

Proof of Lemma 2.5. Since $\alpha > 0$, the $\mathbb{L}$-valued Markov chain with transition probabilities

$$p_{yy'}^x = \frac{1}{|\mathbb{L}|}\sum_{n=1}^{|\mathbb{L}|} \hat{r}_{yy'}^n(x),\ y, y' \in \mathbb{L},$$

is ergodic for each $x \in \mathbb{R}^d$. Since from Assumption 2.1 $x \mapsto r_{yy'}(x)$ is Lipschitz continuous, it follows that $p^x_{yy'}$ is Lipschitz continuous in x and $\inf_{y,y' \in \mathbb{L}} \inf_{x \in \mathbb{R}^d} p^x_{yy'} > 0$. Denote the unique invariant measure of this chain by $\pi(x)$. From Lemma 3.1 in [13], $\pi(x)$ is given as a ratio of polynomials in $\{p^x_{yz}\}_{y,z \in \mathbb{L}}$. Thus $x \mapsto \pi_y(x)$ is Lipschitz continuous for every $y \in \mathbb{L}$ (with Lipschitz constant depending on κ_2, κ_3). The lemma now follows on observing that $\nu_y(x) \propto \frac{\pi_y(x)}{c(x,y)}$, and hence the assertion follows from the Lipschitz property of $x \mapsto c(x,y)$ and the properties $\underline{\varsigma} > 0$ and $\bar{\varsigma} < \infty$. □

Proof of Lemma 2.6. The linear growth property is clear from the Lipschitz property of f. The local Lipschitz property follows by noting that for any compact $K \subset \mathbb{R}^d$ and $x, x' \in K$

$$
\begin{aligned}
|\hat{f}(x) - \hat{f}(x')| &\leq \sum_{y \in \mathbb{L}} \left[\nu_y(x) |f(x,y) - f(x',y)| + \sup_{x \in K} |f(x,y)| |\nu_y(x) - \nu_y(x')| \right] \\
&\leq \left(L^f_{\text{lip}} + \max_{y \in \mathbb{L}} \sup_{x \in K} |f(x,y)| L^\nu_{\text{lip}} \right) |\mathbb{L}| \, \|x - x'\|. \qquad \square
\end{aligned}
$$

Proof of Lemma 5.2. Clearly (5.2) holds if v is continuous. Now suppose that v is bounded and let $M_0 = \sup_{s \in [0,T]} |v(s)|$. Fix $\gamma > 0$. Then by Lusin's theorem, there is a continuous function $\bar{v} : [0,T] \to \mathbb{R}$ such that $\sup_{s \in [0,T]} |\bar{v}(s)| \leq M_0$ and $\lambda_T(s \in [0,T] : v(s) \neq \bar{v}(s)) \leq \gamma$. Since (5.2) holds with v replaced with $\bar{v}$, we have

$$
\begin{aligned}
&\limsup_{n \to \infty} \left| \int_{\mathbb{L} \times [0,T]} v(s) 1_{\{j\}}(y) m_n(dy \times ds) - \int_{\mathbb{L} \times [0,T]} v(s) 1_{\{j\}}(y) m(dy \times ds) \right| \\
&\leq \limsup_{n \to \infty} \left| \int_{\mathbb{L} \times [0,T]} \bar{v}(s) 1_{\{j\}}(y) m_n(dy \times ds) - \int_{\mathbb{L} \times [0,T]} \bar{v}(s) 1_{\{j\}}(y) m(dy \times ds) \right| + 2M_0 \gamma \\
&= 2M_0 \gamma.
\end{aligned}
$$

Sending $\gamma \to 0$ we see that (5.2) holds for all bounded v. Finally, consider a general integrable v. Then (5.2) holds with v replaced with $v_M \doteq v 1_{\{|v| \leq M\}}$. Also, as $M \to \infty$

$$
\sup_{n \in \mathbb{N}} \int_{\mathbb{L} \times [0,T]} |v(s)| 1_{\{|v(s)| \geq M\}} 1_{\{j\}}(y) m_n(dy \times ds) \leq \int_{\{|v| \geq M\}} |v(s)| ds \to 0.
$$

This shows that (5.2) holds for a general integrable v and completes the proof of the lemma. □

B Frequently used notation

$\mathbb{L}$	state space of the fast process (Section 2)
$a(x,y) = a_y(x), b(x,y) = b_y(x)$ $c(x,y) = c_y(x), r(x,y,y') = r_{yy'}(x)$	model coefficients (Section 2, Assumption 2.1)
$E_{yy'}(x) = [0, c_y(x) r_{yy'}(x)]$	(Section 2, Equation (2.1))
$\kappa_1, \bar{\varsigma}, \underline{\varsigma}, \zeta, \alpha, \kappa_3$	constants associated with the model (Section 2)
Π_x	generator for the discrete process Y frozen at x (Section 2)
$\mathbb{T}$	collection of (i,j) with $r_{ij}(x) > 0$ (Section 2)
$N^{\varepsilon^{-1}}_{ij}$	Poisson random measures (Section 2)
$X^\varepsilon, Y^\varepsilon$	State Processes (Section 2, Equation (2.3))
$\nu(x)$	invariant measure associated Π_x (Section 2, Theorem 2.4)

$\hat{b}(x) = \sum_y b(x,y)\nu_y(x)$	averaged drift (Section 2)
$(\Gamma^{\psi}_{ij}(x)), (\Gamma^{\eta}_{ij}(x))$	controlled generators (Section 2.1, Equations (2.6) and (2.7))
$\mathcal{R}$	controlled rates (Section 2.1, Equation (2.7))
$I(\xi), \hat{I}(\xi)$	rate function and its alternate form (Section 2.1, Equations (2.8) and (2.18))
$\mathcal{V}(\xi)$	set of controls (u, φ, π) in definition of $I(\xi)$ (Section 2.1)
$\mathcal{P}_s(\xi)$	set of probability measures Q in definition of $\hat{I}(\xi)$ (Section 2.2)
$\hat{\ell}(\eta)$	cost for rate control (Section 2.2)
$\mathbb{H}_{[a,b]}, \mathbb{H}_T$	space for occupation measures (Section 2.2, Equation (2.14))
$\mathcal{P}_{\text{leb}}(\mathbb{H}_T)$	space of probability measures on $\mathbb{H}_T$ with Lebesgue marginal (Section 2.2)
$\mathcal{U}$,$\mathcal{U}_b$	classes of controls (Section 3)
$L_T, \tilde{L}_T, \bar{L}_T$	integrated cost functions (Section 3, Equation (3.2))
$\eta^{\varepsilon}(t)$	rate control measure (Section 3, Equation (3.10))
Q^{ε}	controlled occupation measure (Section 3, Equation (3.11))
$\varrho_i(s,x)$	invariant measure for fixed (s,x) (Section 4, Proposition 4.1)

References

[1] Vladimir I. Arnold, Valery V. Kozlov, and Anatoly I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*, volume 3 of *Encyclopedia of Mathematical Sciences*. Springer-Verlag, Berlin, third edition, 2006. ISBN 978-3-540-28246-4; 3-540-28246-7. [Dynamical systems. III], Translated from the Russian original by E. Khukhro. MR-2269239

[2] Karen Ball, Thomas G. Kurtz, Lea Popovic, and Greg Rempala. Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, 16(4):1925–1961, 2006. ISSN 1050-5164. doi: 10.1214/105051606000000420. URL http://dx.doi.org/10.1214/105051606000000420.

[3] N. N. Bogoliubov and Y. A. Mitropolsky. *Asymptotic Methods in the Theory of Nonlinear Oscillations*. Translated from the second revised Russian edition. International Monographs on Advanced Mathematics and Physics. Hindustan Publishing Corp., Delhi, Gordon and Breach Science Publishers, New York, 1961. MR-0141845

[4] Michelle Boué and Paul Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998. ISSN 0091-1798. MR-1675051

[5] Amarjit Budhiraja, Paul Dupuis, and Vasileios Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *Ann. Probab.*, 36(4):1390–1420, 2008. MR-2435853

[6] Amarjit Budhiraja, Paul Dupuis, and Vasileios Maroulas. Variational representations for continuous time processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(3):725–747, 2011. ISSN 0246-0203. MR-2841073

[7] Amarjit Budhiraja, Jiang Chen, and Paul Dupuis. Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stochastic Process. Appl.*, 123 (2):523–560, 2013. ISSN 0304-4149. MR-3003362

[8] Amarjit Budhiraja, Paul Dupuis, and Arnab Ganguly. Moderate deviation principles for stochastic differential equations with jumps. *Ann. Probab.*, 44(3):1723–1775, 2016. ISSN 0091-1798. MR-3502593

[9] A. Crudu, A. Debussche, and O. Radulescu. Hybrid stochastic simplifications for multiscale gene networks. *BMC Systems Biology*, 3(89), 2009.

[10] Paul Dupuis and Richard S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1997. A Wiley-Interscience Publication. MR-1431744

[11] Paul Dupuis, Konstantinos Spiliopoulos, and Hui Wang. Importance sampling for multiscale diffusions. *Multiscale Modeling & Simulation*, 10(1):1–27, 2012. MR-2902596

[12] Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006. ISBN 978-0-8218-4145-7; 0-8218-4145-9. MR-2260560

[13] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. ISBN 0-387-98362-7. Translated from the 1979 Russian original by Joseph Szücs. MR-1652127

[14] R. Z. Hasminskii. On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968. ISSN 0023-5954. MR-0260052

[15] Qi He and G. Yin. Large deviations for multi-scale Markovian switching systems with a small diffusion. *Asymptot. Anal.*, 87(3-4):123–145, 2014. ISSN 0921-7134. MR-3195726

[16] Qi He, George Yin, and Qing Zhang. Large deviations for two-time-scale systems driven by nonhomogeneous Markov chains and associated optimal control problems. *SIAM J. Control Optim.*, 49(4):1737–1765, 2011. ISSN 0363-0129. MR-2837573

[17] Gang Huang, Michel Mandjes, and Peter Spreij. Large deviations for Markov-modulated diffusion processes with rapid switching. *Stochastic Processes and their Applications*, 126(6): 1785–1818, 2016. ISSN 0304-4149. doi: http://dx.doi.org/10.1016/j.spa.2015.12.005. URL http://www.sciencedirect.com/science/article/pii/S0304414915003221. MR-3483737

[18] H. W. Kang and T. G. Kurtz. Separation of time-scales and model reduction for stochastic reaction networks. *Annals of Applied Probability*, 23(2):529–583, 2013. ISSN 1050-5164. MR-3059268

[19] Hye-Won Kang, Thomas G. Kurtz, and Lea Popovic. Central limit theorems and diffusion approximations for multiscale Markov chain models. *Ann. Appl. Probab.*, 24(2):721–759, 2014. ISSN 1050-5164. MR-3178496

[20] Kifer, Yuri. Averaging in dynamical systems and large deviations. *Inventiones Mathematicae*, 110, no. 1: 337–370, 1992. MR-1185587

[21] Kifer, Yuri. Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging. *American Mathematical Soc.*, 2009. MR-2547839

[22] Rohini Kumar and Lea Popovic. Large deviations for multi-scale jump-diffusion processes. *Stochastic Processes and their Applications*, pages –, 2016. ISSN 0304-4149. doi: http://dx.doi.org/10.1016/j.spa.2016.07.016. URL http://www.sciencedirect.com/science/article/pii/S0304414916301284. To appear. MR-3619272

[23] Robert Liptser. Large deviations for two scaled diffusions. *Probab. Theory Related Fields*, 106(1):71–104, 1996. ISSN 0178-8051. MR-1408417

[24] A. Neĭshtadt. Averaging, capture into resonances, and chaos in nonlinear systems. pages 261–273. Amer. Inst. Phys., New York, 1990.

[25] Anatolii A. Puhalskii. On large deviations of coupled diffusions with time scale separation. *Ann. Probab.*, 44(4):3111–3186, 2016. ISSN 0091-1798. MR-3531687

[26] A. V. Skorokhod. *Asymptotic Methods in the Theory of Stochastic Differential Equations*, volume 78 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1989. ISBN 0-8218-4531-4. Translated from the Russian by H. H. McFaden. MR-1020057

[27] A. Yu. Veretennikov. On an averaging principle for systems of stochastic differential equations. *Mat. Sb.*, 181(2):256–268, 1990. ISSN 0368-8666. MR-1046602

[28] A. Yu. Veretennikov. On large deviations in the averaging principle for SDEs with a "full dependence". *Ann. Probab.*, 27(1):284–296, 1999. ISSN 0091-1798. MR-1681106

[29] A. Yu. Veretennikov. On large deviations for SDEs with small diffusion and averaging. *Stochastic Process. Appl.*, 89(1):69–79, 2000. ISSN 0304-4149. MR-1775227

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System `http://www.vtex.lt/en/ejms.html`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`