MODERATE DEVIATIONS PRINCIPLE FOR SYSTEMS OF SLOW-FAST DIFFUSIONS

MATTHEW R. MORSE AND KONSTANTINOS SPILIOPoulos

Abstract. In this paper, we prove the moderate deviations principle (MDP) for a general system of slow-fast dynamics. We provide a unified approach, based on weak convergence ideas and stochastic control arguments, that cover both the averaging and the homogenization regimes. We allow the coefficients to be in the whole space and not just the torus and allow the noises driving the slow and fast processes to be correlated arbitrarily. Similar to the large deviation case, the methodology that we follow allows construction of provably efficient Monte Carlo methods for rare events that fall into the moderate deviations regime.

1. Introduction

The goal of this paper is to study moderate deviations for a large class of multiscale diffusion processes with small noise. In particular, we consider the system of slow-fast dynamics

(1)
\[\begin{align*}
dX_t^\varepsilon &= \left[ \frac{\varepsilon}{\delta}b(X_t^\varepsilon, Y_t^\varepsilon) + c(X_t^\varepsilon, Y_t^\varepsilon) \right] dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon, Y_t^\varepsilon) dW_t \\
dY_t^\varepsilon &= \frac{1}{\delta} \left[ \frac{\varepsilon}{\delta}f(X_t^\varepsilon, Y_t^\varepsilon) + g(X_t^\varepsilon, Y_t^\varepsilon) \right] dt + \frac{\sqrt{\varepsilon}}{\delta} \left[ \tau_1(X_t^\varepsilon, Y_t^\varepsilon) dW_t + \tau_2(X_t^\varepsilon, Y_t^\varepsilon) dB_t \right] \\
X_0^\varepsilon &= x_0, \quad Y_0^\varepsilon = y_0
\end{align*}\]

for \( t \in [0, 1] \) such that \((X_t^\varepsilon, Y_t^\varepsilon) \in \mathbb{R}^n \times \mathbb{R}^d\). For convenience, we refer to the state space of \( Y^\varepsilon \) as \( \mathcal{Y} \). The parameter \( \varepsilon \ll 1 \) represents the strength of the noise while \( \delta \ll 1 \) is the time-scale separation parameter. \( W_t \) and \( B_t \) are independent \( m \)-dimensional Brownian motions.

In (1), \( X^\varepsilon \) is the slow motion and \( Y^\varepsilon \) is the fast motion. Depending on the order in which \( \varepsilon, \delta \) go to zero, we get different behavior, and in particular

---

Department of Mathematics and Statistics, Boston University, Boston, MA 02215

E-mail addresses: mrmorse@bu.edu, kspiliop@math.bu.edu.

The present research was partially supported by the National Science Foundation (DMS 1550918).
we are interested in the following regimes:

\[
\lim_{\varepsilon \to 0} \frac{\varepsilon}{\varepsilon} = \begin{cases} 
\infty, & \text{Regime 1,} \\
\gamma \in (0, \infty), & \text{Regime 2.}
\end{cases}
\]

The goal of moderate deviations is to study the behavior of the process \((X^\varepsilon_t)\) in the regime between the central limit theorem behavior and the large deviation behavior. To be more precise, let \(h(\varepsilon) \to +\infty\) such that \(\sqrt{\varepsilon} h(\varepsilon) \to 0\) as \(\varepsilon \to 0\), denote by \(\bar{X}_t = \lim_{\varepsilon \to 0} X^\varepsilon_t\) (in the appropriate sense) the law of large numbers, and define the moderate deviation process

\[
\eta^\varepsilon_t = \frac{X^\varepsilon_t - \bar{X}_t}{\sqrt{\varepsilon} h(\varepsilon)}.
\]

The goal is to derive the large deviation principle for \(\eta^\varepsilon_t\), which is the moderate deviations principle for \(X^\varepsilon_t\). Notice that if \(h(\varepsilon) = 1\) then the limiting behavior of \(\eta^\varepsilon_t\) is that of the central limit theorem (CLT) whereas if \(h(\varepsilon) = 1/\sqrt{\varepsilon}\) then we would get the large deviation result.

Both large and moderate deviations theory have a long history. For general results on large deviations, we refer the interested reader to classical manuscripts such as [5, 11]. In regards to moderate deviations for diffusion processes, one of the first results was derived in [10] even though the analysis there was restricted to the setup with \(b = \sigma = 0\) and under abstract conditions. In [12] the author studies the moderate deviations for \(X^\varepsilon_t\) in the case of \(\varepsilon = \delta\) with \(b = 0\) (averaging regime) and with the fast process \(Y^\varepsilon_t\) being independent of the driving noise of the slow process \(X^\varepsilon_t\), using different methods. In [13] the authors study the MDP for integrated functionals of systems like \(\{\varepsilon\}^\varepsilon\), in the averaging regime (i.e., when \(b = 0\)) and with the fast process being independent of the driving noise of the slow process. In addition we also mention here the recent work of [6] where the moderate deviations principle (MDP) is derived for recursive stochastic algorithms (without multiple scales) using the weak convergence approach of [5].

We conclude this literature review by mentioning that the CLT for \(X^\varepsilon_t\), i.e. when \(h(\varepsilon) = 1\), has been derived in [16]. The LDP for \(X^\varepsilon_t\) is studied in a series of papers [7, 17, 18] for the cases of fast motion in periodic or in random stationary environments. Large deviations results for averaging problems have also been obtained in [4, 10, 20].

The novelty of this paper is fourfold. First, we obtain an explicit form for the action functional of the MDP which is given in terms of solutions to auxiliary, but specific, Poisson equations that can be solved either analytically or numerically. This makes the computation of the action functional possible for a wide range of models, in contrast to existing literature where that was possible only for a more restrictive class of models (for instance we provide the MDP also in the homogenization regime, i.e., in Regime 1 with \(b \neq 0\)). We also illustrate this with a number of examples. Second, the
method of our proof relies on the weak convergence approach of [3] which allows us to connect the moderate deviations problem with a stochastic control problem. As in the case of large deviations (see [5, 17]), the solution to the stochastic control problem gives vital information for the design of efficient Monte Carlo methods for estimation of moderate deviations probabilities of interest. We plan to address the design of Monte Carlo methods based on the moderate deviations principle in a subsequent work. Third, we treat both the averaging regime, Regime 2 or Regime 1 with \( b = 0 \), and the homogenization regime, Regime 1 with \( b \neq 0 \), in a unified way. Fourth, the fast process \( Y^\varepsilon_t \) is allowed to be both fully correlated with the slow process \( X^\varepsilon_t \) and is also allowed to take values in the whole Euclidean space and not just on the torus. The latter fact complicates the mathematical analysis significantly and in particular, the proof of tightness. We gather all of the necessary technical results in Appendixes B and C.

The rest of the paper is organized as follows. In Section 2, we introduce notation and model conditions and state the main result as Theorem 2.1. In Section 3, we present examples of the MDP. In Sections 4, 5, and 6 we prove Theorem 2.1. In Section 4, we introduce the stochastic control representation, show the connection with the MDP, and define the concept of viable pairs, which is essential to the proof. In Section 5, we prove the MDP for Regime 1. In Section 6, we discuss the changes in the proof necessary for Regime 2. In the Appendix, we prove several auxiliary lemmas which are used in the main proof and constitute the main technical challenges.

2. Notation, Conditions, and Main Results

2.1. Notation, Conditions, and Preliminaries. We work with the canonical filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\) equipped with a filtration \(\mathcal{F}_t\) that is right continuous and \(\mathcal{F}_0\) contains all \(\mathbb{P}\)-negligible sets.

For given sets \(A, B\), for \(i, j \in \mathbb{N}\) and \(\alpha \in (0, 1)\) we denote by \(\mathcal{C}^{i,j+\alpha}_b(A \times B)\), the space of functions with \(i\) bounded derivatives in \(x\) and \(j\) derivatives in \(y\), with all partial derivatives being \(\alpha\)-Hölder continuous with respect to \(y\), uniformly in \(x\).

We impose the following conditions on the SDE (1).

Condition 2.1. (i) Let \(\tilde{h}\) be either of the functions \(b\) or \(c\). \(\tilde{h}(\cdot, y) \in \mathcal{C}^2(\mathbb{R}^n)\) for all \(y \in \mathcal{Y}\), \(\nabla_y \nabla_x \tilde{h} \in \mathcal{C}(\mathbb{R}^n \times \mathcal{Y})\), \(\tilde{h}(x, \cdot) \in \mathcal{C}^{\alpha}(\mathcal{Y})\) uniformly in \(x \in \mathbb{R}^n\) for some \(\alpha \in (0, 1)\), and there exist \(K\) and \(q_b, q_c \geq 0\) such that

\[
|b(x, y)| + \|\nabla_x b(x, y)\| + \|\nabla_x \nabla_x b(x, y)\| \leq K(1 + |y|^{q_b})
\]

\[
|c(x, y)| + \|\nabla_x c(x, y)\| + \|\nabla_x \nabla_x c(x, y)\| \leq K(1 + |y|^{q_c})
\]
(ii) For every $N > 0$ there exists a constant $C(N)$ such that for all $x_1, x_2 \in \mathbb{R}^n$ and $|y| \leq N$, the diffusion matrix $\sigma$ satisfies
\[
\|\sigma(x_1, y) - \sigma(x_2, y)\| \leq C(N)|x_1 - x_2|.
\]
Moreover, there exists $K > 0$ and $q_\sigma \geq 0$ such that
\[
\|\sigma(x, y)\| \leq K(1 + |y|^{q_\sigma}).
\]

(iii) The functions $f(x, y), g(x, y), \tau_1(x, y),$ and $\tau_2(x, y)$ are $C^{2,2+\alpha}_b(\mathbb{R}^n \times \mathcal{Y})$ with $\alpha \in (0, 1)$. In addition, $g$ is uniformly bounded.

**Condition 2.2.** (i) The diffusion matrix $\tau_1 \tau_1^T + \tau_2 \tau_2^T$ is uniformly continuous and bounded, nondegenerate and there exist constants $\beta_1, \beta_2 > 0$ such that
\[
0 < \beta_1 \leq \frac{\langle (\tau_1 \tau_1^T(x, y) + \tau_2 \tau_2^T(x, y))y, y \rangle}{|y|^2} \leq \beta_2.
\]

(ii) There exists $R, \Gamma > 0$ and $r \geq 0$ such that in Regime 1,
\[
\sup_{x \in \mathbb{R}^n} f(x, y) \cdot y \leq -\Gamma|y|^{r+1} \text{ for } |y| > R,
\]
and in Regime 2,
\[
\sup_{x \in \mathbb{R}^n} (\gamma f(x, y) + g(x, y)) \cdot y \leq -\Gamma|y|^{r+1} \text{ for } |y| > R.
\]

For each Regime $i = 1, 2$, define an operator $\mathcal{L}_{i,x}$ (treating $x$ as a parameter) by
\[
\mathcal{L}_{1,x} F(y) = (\nabla_y F(y)) f(x, y) + \frac{1}{2} (\tau_1 \tau_1^T + \tau_2 \tau_2^T)(x, y) : \nabla_y \nabla_y F(y),
\]
\[
\mathcal{L}_{2,x} F(y) = (\nabla_y F(y)) (\gamma f(x, y) + g(x, y)) + \gamma \frac{1}{2} (\tau_1 \tau_1^T + \tau_2 \tau_2^T)(x, y) : \nabla_y \nabla_y F(y)
\]
where the notation $A : B$ for two $n \times k$ matrices means the trace of their product,
\[
A : B = \sum_{i=1}^n \sum_{j=1}^k a_{ij} b_{ij}.
\]

For a $k \times k$ matrix $A$ and a $n$-dimensional vector–valued function of a $k$-dimensional vector $f(x)$ define $A : \nabla \nabla f$ as a $n$-dimensional vector where component $i$ is equal to $A : \nabla \nabla f_i$. Also, for notational convenience we sometimes collect the variables at the end of the expression and we write
\[
\tau \tau^T(x, y) = \tau(x, y) \tau(x, y)^T.
\]

Operators $\mathcal{L}_{1,x}$ and $\mathcal{L}_{2,x}$ are the infinitesimal generators for the processes that play the role of the fast motion (and with respect to which averaging is being performed) in Regimes 1 and 2 respectively. Condition 2.2 guarantees
that the fast process in each Regime \( i = 1, 2 \) has a unique invariant measure, denoted by \( \mu_{i,x}(dy) \), for each \( x \in \mathbb{R}^n \).

Because the fast motion takes values in an unbounded space, \( \mathbb{R}^d \), the constants \( q_b, q_c, q_\sigma \) that determine the growth of the coefficients from Condition 2.1 and the constant \( r \) from Condition 2.2 that determines the recurrent properties of the fast component, will need to be related in order for the subsequent tightness argument to go through. In particular, we have Condition 2.3.

**Condition 2.3.** Consider the constants \( q_b, q_c, q_\sigma \) from Condition 2.1 and the constant \( r \) from Condition 2.2. Define \( q_{b,c} = \max\{q_b, q_c\} \) and \( q_F = \max\{q_b, q_c, (q_b + 1 - r)^+\} \), where for any \( x \in \mathbb{R} \) we have set \( (x)^+ = x_{1_{x \geq 0}} \).

Then in Regime 1, we assume that

\[
\max \left\{ (q_F + 1 - r)^+ + q_{b,c}, (q_F + 2(1 - r))^+ + q_{b,c}, (q_F + 1 - r)^+_2 + 2q_\sigma, \right.
\]

\[
(q_F + 2(1 - r))^+ + 2q_\sigma, (q_F + 3(1 - r))^+_2 + 2q_\sigma \right\} \leq r,
\]

\[
\max \{q_F, q_\sigma, (q_F + 1 - r)^+ \} < r.
\]

In Regime 2, we assume

\[
\max \left\{ (q_{b,c} + 1 - r)^+ + q_{b,c}, (q_{b,c} + 2(1 - r))^+ + q_{b,c}, (q_{b,c} + 1 - r)^+ + 2q_\sigma, \right. \]

\[
(q_{b,c} + 2(1 - r))^+ + 2q_\sigma, (q_{b,c} + 3(1 - r))^+_2 + 2q_\sigma \right\} \leq r,
\]

\[
\max \{q_{b,c}, q_\sigma, (q_{b,c} + 1 - r)^+ \} < r.
\]

**Remark 2.1.** Of course, it is clear that if the fast process is the Ornstein-Uhlenbeck process for example, where \( r = 1 \), Condition 2.3 can be dramatically simplified, see Example 1 in Section 3. In addition, it is also clear that Condition 2.3 places some restrictions on \( r \) as well. For example, if all the coefficients are bounded, in which case \( q_b = q_c = q_\sigma = 0 \), then we need to have that \( r \geq 4/5 \) for Regime 1 and \( r \geq 3/4 \) for Regime 2.

In addition, in Regime 1, we impose the following centering condition.

**Condition 2.4.** The drift term \( b \) satisfies

\[
\int_{\mathcal{Y}} b(x,y) \mu_{1,x}(dy) = 0.
\]

Then by the results in [14][15], which we collected in Theorem A.1 in the Appendix, for each \( \ell \in 1, \ldots, n \), there is a unique, twice differentiable function \( \chi_\ell(x,y) \) in the class of functions that grows at most polynomially in \( |y| \) that satisfies the equation

\[
(3) \quad \mathcal{L}_{1,x} \chi_\ell(x,y) = -b_\ell(x,y), \quad \int_{\mathcal{Y}} \chi_\ell(x,y) \mu_{1,x}(dy) = 0, \text{ for } \ell = 1, \ldots, n,
\]

where \( b_\ell(x,y) \) is the \( \ell \)th component of the vector \( b(x,y) = (b_1(x,y), \ldots, b_n(x,y)) \).

Let us set \( \chi(x,y) = (\chi_1(x,y), \ldots, \chi_n(x,y)) \). Define the function \( \lambda_i(x,y) : \mathbb{R}^n \times \)
\( \mathcal{Y} \rightarrow \mathbb{R}^n \) under Regime \( i \) by

\[
\lambda_1(x, y) = (\nabla_y \chi(x, y)) g(x, y) + c(x, y) \\
\lambda_2(x, y) = \gamma b(x, y) + c(x, y).
\]

Under Regime \( i \), for any function \( G(x, y) \), define the averaged function \( \bar{G} \) by

\[
(4) \quad \bar{G}(x) = \int_{\mathcal{Y}} G(x, y) \mu_{i,x}(dy).
\]

It follows that \( \bar{G} \) inherits the continuity and differentiability properties of \( G \). In particular, for each regime,

\[
\bar{\lambda}_i(x) = \int_{\mathcal{Y}} \lambda_i(x, y) \mu_{i,x}(dy).
\]

Then by an argument similar to that of Theorem 3.2 in [17], as \( \varepsilon \downarrow 0 \), in Regime \( i \) we have the averaging result \( X_{\varepsilon}^x \rightarrow \bar{X}_t \) in probability, where \( \bar{X}_t \) is defined by

\[
d\bar{X}_t = \bar{\lambda}_i(\bar{X}_t) dt, \quad \bar{X}_0 = x_0.
\]

Lastly, for Regime \( i = 1, 2 \), introduce the function \( \Phi_i(x, y) \), given by the PDE

\[
(5) \quad \mathcal{L}_{i,x} \Phi_i(x, y) = -(\lambda_i(x, y) - \bar{\lambda}_i(x)), \quad \int_{\mathcal{Y}} \Phi_i(x, y) \mu_{i,x}(dy) = 0.
\]

Under our assumptions, each one of \( \lambda_i - \bar{\lambda}_i \), for \( i = 1, 2 \), satisfy the assumptions of Theorem A.1 part (iii), and thus by Theorem A.1 (5) has a unique classical solution in the class of functions which grow at most polynomially in \( |y| \) for every \( x \).

Last but not least we assume uniqueness of a strong solution.

**Condition 2.5.** We assume that the SDE (1) has a unique strong solution.

**Remark 2.2.** Condition 2.5 holds for example if the coefficients are Lipschitz continuous with at most linear growth. However, these conditions can be significantly weakened, see for example [12]. Conditions 2.1, 2.3 and 2.5 should be considered together and it is clear that depending on the value of \( r \) in the recurrence Condition 2.2, Conditions 2.1, 2.3 will directly imply Condition 2.5. For example if \( r = 1 \) then the coefficients cannot grow faster than linearly in \( y \) and are always assumed to be bounded in \( x \), so in that case for instance Condition 2.5 instantly holds.
2.2. Main Results. By [3], the LDP for $\eta^\varepsilon_i$ is equivalent to the Laplace principle, which states that for any bounded continuous function $a: C([0,1];\mathbb{R}^n) \rightarrow \mathbb{R}$,

\[
\lim_{\varepsilon \downarrow 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right] = \inf_{\xi \in C([0,1];\mathbb{R}^n)} (S(\xi) + a(\xi))
\]

where $S(\xi)$ is called the action functional. In this paper we essentially prove [6] and Theorem 2.1 identifies the action functional $S(\xi)$. In order to state Theorem 2.1, we need to know the relative rates at which $\delta$, $\varepsilon$, and $1/h(\varepsilon)$ vanish. In particular, in Regime $i$, $i = 1, 2$, define $j_1$, $j_2$ by

\[
j_1 = \lim_{\varepsilon \downarrow 0} \frac{\delta/\varepsilon}{\sqrt{\varepsilon} h(\varepsilon)} < \infty, \quad j_2 = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon/\delta - \gamma}{\sqrt{\varepsilon} h(\varepsilon)} < \infty.
\]

$j_1$, $j_2$ specifies the relative rate at which $\varepsilon/\delta$ goes to its limit and $h(\varepsilon)$ goes to infinity. In order for a moderate deviations principle to hold, we require that $j_1$, $j_2$ be finite.

The main result of this paper is the following theorem.

**Theorem 2.1.** Let Conditions 2.1, 2.2, 2.3 and 2.5 be satisfied. Additionally, under Regime 1, let Condition 2.4 be satisfied. Then under Regime $i$, $i = 1, 2$, the process $\{X^\varepsilon, \varepsilon > 0\}$ from [4] satisfies the MDP, with the action functional $S(\xi)$ given by

\[
S(\xi) = \frac{1}{2} \int_0^1 \left( \xi - \kappa(\bar{X}_s, \xi_s) \right)^T q^{-1}(\bar{X}_s) \left( \xi - \kappa(\bar{X}_s, \xi_s) \right) ds
\]

if $\xi \in C([0,1];\mathbb{R}^n)$ is absolutely continuous, and $\infty$ otherwise. Under Regime 1, we have

\[
\kappa(x, \eta) = (\nabla_x \lambda_1(x))\eta + j_1 \int_\mathcal{Y} (\nabla_y \Phi_1(x, y)) g(x, y) \mu_{1,x}(dy)
\]

\[
q(x) = \int_\mathcal{Y} (\alpha_1 \alpha_1^T(x, y) + \alpha_2 \alpha_2^T(x, y)) \mu_{1,x}(dy)
\]

\[
\alpha_1(x, y) = \sigma(x, y) + (\nabla_y \chi(x, y)) \tau_1(x, y), \quad \alpha_2(x, y) = (\nabla_y \chi(x, y)) \tau_2(x, y).
\]

Under Regime 2, we have

\[
\kappa(x, \eta) = (\nabla_x \lambda_2(x))\eta + j_2 \int_\mathcal{Y} \left[ b(x, y) - \frac{1}{\gamma} (\nabla_y \Phi_2(x, y)) g(x, y) \right] \mu_{2,x}(dy)
\]

\[
q(x) = \int_\mathcal{Y} (\alpha_1 \alpha_1^T(x, y) + \alpha_2 \alpha_2^T(x, y)) \mu_{2,x}(dy)
\]

\[
\alpha_1(x, y) = \sigma(x, y) + (\nabla_y \Phi_2(x, y)) \tau_1(x, y), \quad \alpha_2(x, y) = (\nabla_y \Phi_2(x, y)) \tau_2(x, y),
\]

where the finite constants $j_1$, $j_2$ are defined in [4].
Remark 2.3. Note that in either Regime, the function $\kappa(x, \eta)$ is affine in $\eta$ and the function $q(x)$ is constant in $\eta$. This is expected by the nature of moderate deviations. In the large deviations case, see [7, 17], the corresponding $\kappa(x)$ and $q(x)$ are nonlinear functions of $x$. The affine structure of $\kappa(x, \eta)$ is what makes the moderate deviations very appealing for the design of Monte Carlo simulation methods, as it makes the solution to the associated Hamilton-Jacobi-Bellman equation much easier to obtain. We plan to explore this in detail in a follow up work.

3. Examples

In this section we present some concrete examples to illustrate Theorem 2.1.

3.1. Example 1. Consider the system of one-dimensional processes

$$dX_t^\varepsilon = b(X_t^\varepsilon, Y_t^\varepsilon) dt + \sqrt{\varepsilon} \sigma(X_t^\varepsilon, Y_t^\varepsilon) dW_t, \quad X_0^\varepsilon = x_0,$$

$$dY_t^\varepsilon = -\frac{1}{\varepsilon} Y_t^\varepsilon dt + \frac{1}{\sqrt{\varepsilon}} dB_t$$

where $B$ and $W$ are independent Brownian motions. The invariant measure of the fast process $Y$ is the Gaussian measure given by $\mu_{2,x}(dy) = \left(\frac{2}{\pi}\right)^{-1/2} \exp(-y^2/2) dy$. This system can be rewritten in terms of (1) with $\delta = \varepsilon$. In this case, the recurrence constant $r$ from Condition 2.2 is $r = 1$ and the restrictions on $q_b$, $q_\sigma$ from Condition 2.3 take the much simpler form $q_b \leq 1/2$ and $2q_\sigma + q_b \leq 1$. Notice that the limit $\bar{X}_t = \lim_{\varepsilon \downarrow 0} X_t^\varepsilon$ is given by

$$d\bar{X}_t = \bar{\lambda}_2(\bar{X}_t) dt, \quad \bar{X}_0 = x_0,$$

where $\bar{\lambda}_2(x) = \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} b(x, y) e^{-y^2/2} dy$.

In this case $\Phi_2(x, y)$, i.e. the solution to the PDE (5), takes the explicit form

$$\frac{\partial \Phi_2}{\partial y}(x, y) = -2e^{y^2/2} \int_{-\infty}^{y} b(x, z) e^{-z^2/2} dz$$

which then implies that the action functional $S(\xi)$ of Theorem 2.1 is defined with

$$\kappa(x, \eta) = \frac{1}{\sqrt{2\pi}} \eta \int_{\mathbb{R}} \frac{\partial b}{\partial x}(x, y) e^{-y^2/2} dy$$

$$q(x) = \int_{\mathbb{R}} \left[ \sigma(x, y)^2 + 4e^{y^2} \left( \int_{-\infty}^{y} b(x, z) e^{-z^2/2} dz \right)^2 \right] \mu_{2,x}(dy).$$

Remark 3.1. [12] presents a similar example under the additional assumption that $\int_{\mathbb{R}} b(x_0, y) \mu_{2,x}(dy) = 0$. By Theorem 1, this assumption is not necessary, and the results here extend the results of [12] to a much more general class of processes in a unified way.
3.2. Example 2. In the second example, we consider the first order Langevin equation under Regime 1,

\[ dX_\varepsilon^t = \left[ -\frac{\varepsilon}{\delta} \nabla Q \left( \frac{X_\varepsilon^t}{\delta} \right) - \nabla V(X_\varepsilon^t) \right] dt + \sqrt{\varepsilon \sqrt{2D}} dW_t, \quad X_\varepsilon^0 = x_0. \]

This equation has a number of applications and has been studied extensively, beginning with [21], see also [8]. In our notation, let \( Y_\varepsilon^t = X_\varepsilon^t / \delta \), \( b(x,y) = f(x,y) = -\nabla Q(y) \), and \( c(x,y) = g(x,y) = -\nabla V(x) \). The invariant density \( \mu(y) \) is the Gibbs measure

\[ \mu(y) = \frac{1}{Z} e^{-Q(y)/D}, \quad Z = \int e^{-Q(y)/D} dy. \]

In order to have closed form formulas, let us also assume that \( Q(y_1, y_2, \ldots, y_d) = Q_1(y_1) + Q_2(y_2) + \cdots + Q_d(y_d) \) and that \( \mathcal{Y} \) is the \( d \)-dimensional unit torus. Since the fast motion is restricted to be on a torus, the recurrence condition (part (ii)) of Condition 2.2 and Condition 2.3 are not needed.

Then \( \bar{X}_t = \lim_{\varepsilon \to 0} X_\varepsilon^t \) is given by

\[ \bar{X}_t = x_0 + \int_0^t \bar{\lambda}_1(\bar{X}_s) \, ds \]

where

\[ \bar{\lambda}_1(x) = -\Theta \nabla V(x), \quad \Theta = \text{diag} \left[ \frac{1}{Z_1 Z_1}, \ldots, \frac{1}{Z_d Z_d} \right] \]

and for \( i = 1, 2, \ldots, d \)

\[ Z_i = \int_T e^{-Q_i(y_i)/D} \, dy_i, \quad \hat{Z}_i = \int_T e^{Q_i(y_i)/D} \, dy_i. \]

\( \Phi_1(x, y) \) is given by

\[ \nabla_y \Phi_1(x, y) = \frac{1}{D} \Theta(y) \nabla V(x) \]

where

\[ \Theta(y) = \text{diag} \left[ \frac{e^{Q_1(y_1)/D}}{Z_i} \left( y_i - \frac{1}{Z_i} \int_0^{y_i} e^{-Q_i(\xi)/D} \, d\xi \right) \right. \\
\left. + \frac{1}{Z_i} \int_0^1 e^{Q_i(\rho)/D} \int_0^{\rho} \left( \frac{1}{Z_i} e^{-Q_i(\xi)/D} - 1 \right) \, d\xi \, d\rho \right]. \]

Then the action functional \( S(\xi) \) of Theorem 2.1 is defined with

\[ \kappa(x, \eta) = -\Theta \nabla \nabla V(x) \eta - j_1 \int_{\mathcal{Y}} \left( \frac{1}{D} \Theta(y) \nabla V(x) \right) \nabla V(x) \mu(dy) \]

\[ q(x) = 2D \Theta. \]
Remark 3.2. We remark that this example is not covered by previous results in the literature on moderate deviations. Here we are able to get a very explicit form for the action functional.

3.3. Example 3. To illustrate the case where \( Y_\varepsilon \) is a CIR (square-root) process, consider the following model:

\[
\begin{align*}
\frac{dX_\varepsilon}{d\varepsilon} &= c(X_\varepsilon, Y_\varepsilon) dt + \sqrt{\varepsilon} \sigma(X_\varepsilon, Y_\varepsilon) dW_t, \\
X_\varepsilon^0 &= x_0 \in \mathbb{R} \\
\frac{dY_\varepsilon}{d\varepsilon} &= \frac{\varepsilon}{\delta^2} a(b - Y_\varepsilon) dt + \tau Y_\varepsilon \sqrt{\varepsilon} dW_t, \\
Y_\varepsilon^0 &= y_0 \in \mathbb{R}
\end{align*}
\]

where \( a, b, \) and \( \tau \) are positive constants satisfying \( 2ab \geq \tau^2 \). Note that this model does not satisfy Condition 2.2 because the fast process noise is degenerate at \( y = 0 \). However, if \( y_0 > 0 \) and \( 2ab \geq \tau^2 \) then \( Y_\varepsilon > 0 \) for all \( t > 0 \) w.p.1., \( Y_\varepsilon \) has the gamma distribution as its unique invariant measure and so the results are expected to hold. For this model, the invariant measure and the limiting process \( \bar{X}_t \) do not depend on the regime. However, as we shall see the MDP for Regimes 1 and 2 do differ. The fast process has the gamma invariant density

\[
m(y) = \frac{(2a/\tau^2)^2ab/\tau^2}{\Gamma(2ab/\tau^2)} y^{2ab/\tau^2-1} e^{-2ay/\tau^2}.
\]

Then \( \bar{X}_t = \lim_{\varepsilon \downarrow 0} X_\varepsilon^t \) satisfies the ordinary differential equation

\[
\bar{X}_t = x_0 + \int_0^t \bar{\lambda}(\bar{X}_s) ds
\]

where

\[
\bar{\lambda}(x) = \int_0^\infty c(x, y) m(y) dy.
\]

Under Regime 1, the action functional \( S(\xi) \) of Theorem 2.1 is expected to be defined with

\[
\kappa(x, \eta) = \eta \frac{d}{dx} \bar{\lambda}(x) \\
q(x) = \int_0^\infty \sigma^2(x, y) m(y) dy.
\]

In contrast, under Regime 2, we let \( \Phi_2(x, y) \) be the unique solution to (5) with \( i = 2 \) and \( \lambda_2(x, y) = c(x, y) \). Then we have that

\[
\kappa(x, \eta) = \eta \frac{d}{dx} \bar{\lambda}(x) \\
q(x) = \int_0^\infty \left( \sigma(x, y) + \tau \sqrt{y} \frac{d\Phi_2(x, y)}{dy} \right)^2 m(y) dy.
\]
Hence, the two MDP’s differ on the formula for \( q(x) \). Again, we remark that this example can be covered with the results of this paper, but it is not clear whether existing previous results in the literature can address it or indicate how the action functional should look like.

4. The controlled processes

The proof of the Laplace principle [3] is based on a stochastic control representation given by Theorem 3.1 in [2]. This theorem is restated here for the convenience of the reader.

**Theorem 4.1.** Let \( Z \) be a \( 2m \)-dimensional Brownian motion with respect to the filtration \( \{ \mathcal{F}_t \} \) for \( 0 \leq t \leq 1 \). Let \( \mathcal{A} \) be the space of \( \mathcal{F}_t \)-progressively measurable \( 2m \)-dimensional processes \( v = (v_1, v_2) \) for \( 0 \leq t \leq 1 \) satisfying

\[
\mathbb{E} \int_0^1 |v(s)|^2 \, ds < \infty.
\]

Let \( F \) be a bounded, measurable, real–valued function defined on the space of \( \mathbb{R}^{2m} \)-valued continuous functions on \([0, 1]\). Then

\[
- \log \mathbb{E} \left[ \exp \{-F(Z(\cdot))\} \right] = \inf_{v \in \mathcal{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^1 |v(s)|^2 \, ds + F \left( Z(\cdot) + \int_0^1 v(s) \, ds \right) \right].
\]

In our case we set \( Z(\cdot) = (W(\cdot), B(\cdot)) \) and each one of \( v_1, v_2 \) are \( m \)-dimensional vectors. Under Condition 2.3 for each \( \varepsilon > 0 \), (11) has a unique strong solution. Therefore \( \eta^\varepsilon \) is a measurable function of \( Z \). Set \( F(Z(\cdot)) = h^2(\varepsilon)a(\eta^\varepsilon(\cdot)) \). Set \( u^\varepsilon_i = v_i/h(\varepsilon), u^\varepsilon = (u^\varepsilon_1, u^\varepsilon_2) \), and then divide by \( h^2(\varepsilon) \) to obtain

\[
- \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \{-h^2(\varepsilon)a(\eta^\varepsilon)\} \right] = \inf_{u^\varepsilon \in \mathcal{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^1 |u^\varepsilon(s)|^2 \, ds + a(\eta^\varepsilon, u^\varepsilon) \right]
\]

where the controlled deviations process \( \eta^{\varepsilon, u^\varepsilon} \) is defined by

\[
\eta^{\varepsilon, u^\varepsilon}_t = \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \left( X^{\varepsilon, u^\varepsilon}_t - \bar{X}_t \right)
\]
and the controlled processes \(X_t^{\varepsilon, u^\varepsilon}\) and \(Y_t^{\varepsilon, u^\varepsilon}\) are defined by

\[
dX_t^{\varepsilon, u^\varepsilon} = \left[ \frac{\varepsilon}{\delta} b(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) + c(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) + \sqrt{\varepsilon} h(\varepsilon) \sigma(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) u_1^\varepsilon(t) \right] dt + \sqrt{\varepsilon} \sigma(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) dW_t
\]

\[
dY_t^{\varepsilon, u^\varepsilon} = \frac{1}{\delta} \left[ f(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) + g(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) + \sqrt{\varepsilon} h(\varepsilon) \tau_1(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) u_1^\varepsilon(t) \right] dt + \sqrt{\varepsilon} \tau_2(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) u_2^\varepsilon(t) dt + \frac{\sqrt{\varepsilon}}{\delta} \left[ \tau_1(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) dW_t + \tau_2(X_t^{\varepsilon, u^\varepsilon}, Y_t^{\varepsilon, u^\varepsilon}) dB_t \right]
\]

\[
X_0^{\varepsilon, u^\varepsilon} = x_0, \quad Y_0^{\varepsilon, u^\varepsilon} = y_0.
\]

Note that we can rewrite \(\eta^{\varepsilon, u^\varepsilon}\) in the form

\[
\eta_t^{\varepsilon, u^\varepsilon} = \int_0^t \frac{1}{\sqrt{\varepsilon} h(\varepsilon)} \left[ \frac{\varepsilon}{\delta} b(X_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) + c(X_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) - \tilde{\lambda}_1(X_s) \right] ds + \int_0^t \sigma(X_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) u_1^\varepsilon(s) ds + \int_0^t \frac{1}{h(\varepsilon)} \sigma(X_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) dW_s.
\]

Define \(Z = \mathbb{R}^m\). This is the space in which the control processes \(u_1^\varepsilon\) and \(u_2^\varepsilon\) take values. Define \(\theta_i(x, \eta, y, z_1, z_2) : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{Y} \times Z \times Z \to \mathbb{R}^n\) by

\[
\theta_1(x, \eta, y, z_1, z_2) = (\nabla_y \kappa(x, y))(\tau_1(x, y) z_1 + \tau_2(x, y) z_2) + j_1(\nabla_y \Phi_1(x, y)) g(x, y)
\]

\[
+ (\nabla_x \tilde{\lambda}_1(x)) \eta + \sigma(x, y) z_1
\]

\[
\theta_2(x, \eta, y, z_1, z_2) = j_2 b(x, y) + (\nabla_y \Phi_2(x, y)) [\tau_1(x, y) z_1 + \tau_2(x, y) z_2] + (\nabla_x \tilde{\lambda}_2(x)) \eta
\]

\[
+ \sigma(x, y) z_1 + j_2 (\nabla_y \Phi_2(x, y)) f(x, y) + \frac{j_2}{2} \left( (\tau_1 \tau_1^T + \tau_2 \tau_2^T)(x, y) : \nabla_y \nabla_y \Phi_2(x, y) \right)
\]

Conditions \textit{2.1}, \textit{2.3} and Theorem \textit{A.1} guarantee that the functions \(\theta_1\) and \(\theta_2\) are bounded in \(x\), affine in \(\eta, z_1\) and \(z_2\) and bounded polynomially in \(|y|\) with order \(r \geq 0\) (\(r\) comes from Condition \textit{2.2}).

Next we introduce the occupation measure \(P^{\varepsilon, \Delta}\). Let \(\Delta = \Delta(\varepsilon) \downarrow 0\) as \(\varepsilon \downarrow 0\), whose role is to exploit a time-scale separation. Let \(A_1, A_2, B,\) and \(\Gamma\) be Borel sets of \(Z = \mathbb{R}^m, \mathcal{Y} = \mathbb{R}^d,\) and \([0,1]\) respectively. Let \((X^{\varepsilon, u^\varepsilon}, Y^{\varepsilon, u^\varepsilon})\) solve (13). Associate with \((X^{\varepsilon, u^\varepsilon}, Y^{\varepsilon, u^\varepsilon})\) and \(u^\varepsilon\) a family of occupation measures \(P^{\varepsilon, \Delta}\) defined by

\[
P^{\varepsilon, \Delta}(A_1 \times A_2 \times B \times \Gamma) = \int_\Gamma \left[ \frac{1}{\Delta} \int_0^{t+\Delta} 1_{A_1}(u_1^\varepsilon(s)) 1_{A_2}(u_2^\varepsilon(s)) 1_B(Y_s^{\varepsilon, u^\varepsilon}) ds \right] dt
\]

and assume \(u_1^\varepsilon(s) = 0\) if \(s > 1\).
Definition 4.1. Let \( \theta(x, \eta, y, z_1, z_2) : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{Y} \times \mathcal{Z} \times \mathcal{Z} \to \mathbb{R}^n \) be a function that has at most polynomial growth in \( |y| \) with order \( r \geq 0 \). For each \( x \in \mathbb{R}^n \), let \( \mathcal{L}_x \) be a second order elliptic partial differential operator and denote by \( D(\mathcal{L}_x) \) its domain of definition. A pair \( (\psi, P) \in \mathcal{C}([0,1]; \mathbb{R}^n) \times \mathcal{P}(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]) \) is called a viable pair with respect to \( (\theta, \mathcal{L}_x) \) if

- The function \( \psi \) is absolutely continuous.
- The measure \( P \) is integrable in the sense that
  \[
  \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} \left[ |z_1|^2 + |z_2|^2 + |y|^{2r} \right] P(dz_1 dz_2 dy ds) < \infty.
  \]
- For all \( t \in [0,1] \),
  \[
  \psi_t = \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,t]} \theta(\bar{X}_s, \psi_s, y, z_1, z_2) P(dz_1 dz_2 dy ds).
  \]
- For all \( t \in [0,1] \) and for every \( F \in D(\mathcal{L}_x) \),
  \[
  \int_0^t \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y}} \mathcal{L}_x F(y) P(dz_1 dz_2 dy ds) = 0.
  \]
- For all \( t \in [0,1] \),
  \[
  P(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,t]) = t.
  \]

We write \( (\psi, P) \in \mathcal{V}(\theta, \mathcal{L}_x) \).

Note that the last item is equivalent to stating that the last marginal of \( P \) is Lebesgue measure, or that \( P \) can be decomposed as \( P(dz_1 dz_2 dy dt) = P_t(dz_1 dz_2 dy) dt \). In comparison to the definition of viable pairs in the large deviations case (for example, \[7\]), \( \psi \) does not appear in (17), and so \( \psi \) and \( P \) are decoupled. Another difference with the definition of viable pair in \[7\] is that here we need to impose the condition \( \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} |y|^{2r} P(dz_1 dz_2 dy ds) < \infty \) which is due to the polynomial growth in \( |y| \) of the involved functions. As we will see in the convergence proof, due to the a priori bound of Lemma \[B.2\] this is a restriction that is satisfied.

The controlled process (12) and definition of viable pairs will be used to prove the following theorem:

**Theorem 4.2.** Let Conditions \[2.1, 2.2, 2.3 \text{ and } 2.5\] be satisfied. Additionally, under Regime 1, let Condition \[2.4\] be satisfied. Then under Regime \( i \), \( i = 1, 2 \), the family of processes \( \{X^\varepsilon, \varepsilon > 0\} \) from (11) satisfies the MDP, with the action functional \( S(\xi) = S_i(\xi) \) given by

\[
S_i(\xi) = \inf_{(\xi, P) \in \mathcal{V}(\theta_i, \mathcal{L}_i, i)} \left[ \frac{1}{2} \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} \left[ |z_1|^2 + |z_2|^2 \right] P(dz_1 dz_2 dy ds) \right]
\]

with the convention that the infimum over the empty set is \( \infty \).
Notice that Theorem 4.2 offers a compact way to write the MDP for both regimes in terms of the appropriate viable pairs each time. As will be shown during the proof, Theorem 2.1 follows directly from Theorem 4.2.

5. Proof in Regime 1

The proof is nearly identical for Regime 1 and for Regime 2, aside from some technical differences. In this section, we present the proof for Regime 1. In Section 6, we discuss the changes necessary for Regime 2. In Subsections 5.1 and 5.2, we prove tightness and convergence of the pair \((\eta^\varepsilon, u^\varepsilon, P^\varepsilon, \Delta)\) respectively. In Subsection 5.3, we prove the Laplace principle lower bound. In Subsection 5.4, we prove compactness of level sets of \(S(.\)\). Finally, in Subsection 5.5, we prove the Laplace principle upper bound and the representation formula of Theorem 2.1.

5.1. Proof of tightness. The main result of this section is the following proposition on tightness.

**Proposition 5.1.** Let Conditions 2.1, 2.2, 2.3, 2.4 and 2.5 be satisfied. Consider any family \(\{u^\varepsilon, \varepsilon > 0\}\) of controls in \(A\) satisfying for some \(N < \infty\)

\[
\sup_{\varepsilon > 0} \int_0^1 |u^\varepsilon(t)|^2 dt < N, \text{ almost surely}
\]

Then the following hold.

1. The family \(\{(X^\varepsilon, u^\varepsilon, P^\varepsilon, \Delta), \varepsilon > 0\}\) is tight.

2. Define the set

\[
B_{r,M} = \{ (z_1, z_2, y) \in Z \times Z \times Y : (|z_1| > M, |z_2| > M, |y|^r > M) \}.
\]

The family \(\{P^\varepsilon, \Delta, \varepsilon > 0\}\) is uniformly integrable in the sense that

\[
\lim_{M \to \infty} \sup_{\varepsilon > 0} \mathbb{E}_{x_0,y_0} \left[ \int_{\{(z_1, z_2, y) \in B_{r,M} \times [0,1]\}} (|z_1| + |z_2| + |y|^r) P^\varepsilon, \Delta(dz_1 dz_2 dy dt) \right] = 0.
\]

The proof of Proposition 5.1 is the subject of Sections 5.1.1 and 5.1.2.

5.1.1. Tightness of \(\{P^\varepsilon, \Delta, \varepsilon, \Delta > 0\}\) on \(\mathcal{P}(Z \times Z \times Y \times [0,1])\). The argument for tightness is similar to the argument for tightness in the proof of Theorem 3.2 in [17] (see also [7]), but with some differences due to the unboundedness of the space on which the fast motion takes values. We repeat here for completeness the argument emphasizing the differences.
By Lemmas B.1 and B.2 in the Appendix, we can restrict to a family \( \{ u^\varepsilon = (u^\varepsilon_1, u^\varepsilon_2), \varepsilon > 0 \} \) of controls in \( A \) satisfying

\[
\sup_{\varepsilon > 0} \mathbb{E} \int_0^1 \left[ |u^\varepsilon_1(s)|^2 + |u^\varepsilon_2(s)|^2 + |Y^\varepsilon,u^\varepsilon|^{2r} \right] ds < \infty.
\]

Recall that a tightness function \( \hat{g}(x) \) is a function mapping a space \( X \) to \( \mathbb{R} \cup \{ \infty \} \) which has a lower bound and for which for each \( M < \infty \), the level set \( Z_{\hat{g}}(M) = \{ x \in X : \hat{g}(x) \leq M \} \) is relatively compact in \( X \).

Consider \( q \in \mathcal{P}(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]) \) (not to be confused with the growth parameters of Condition 2.1). The function

\[
\hat{g}(q) = \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} \left[ |z_1|^2 + |z_2|^2 + |y|^{2r} \right] q(dz_1 dz_2 dy dt)
\]

is a tightness function on \( \mathcal{P}(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]) \) by the facts that it is nonnegative and that the level sets of \( \hat{g} \) are relatively compact. Then by Theorem A.3.17 in [5], for each \( M < \infty \), the set

\[
Z_{\hat{g}}(M) = \left\{ \theta \in \mathcal{P}(\mathcal{P}(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1])) : \int_{\mathcal{P}(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1])} \hat{g}(q) \theta(dq) \leq M \right\}
\]

is tight. Tightness of \( \{ \mathcal{P}^\varepsilon, \varepsilon, \Delta > 0 \} \) follows from the bound

\[
\sup_{\varepsilon \in (0,1]} \mathbb{E}[\hat{g}(\mathcal{P}^\varepsilon,\Delta)] = \sup_{\varepsilon \in (0,1]} \mathbb{E} \left[ \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} \left( |z_1|^2 + |z_2|^2 + |y|^{2r} \right) \mathcal{P}^\varepsilon,\Delta(dz_1 dz_2 dy dt) \right]
\]

\[
= \sup_{\varepsilon \in (0,1]} \mathbb{E} \int_0^1 \frac{1}{\Delta} \int_t^{t+\Delta} \left[ |u^\varepsilon_1(s)|^2 + |u^\varepsilon_2(s)|^2 + |Y^\varepsilon,u^\varepsilon|^{2r} \right] ds dt < \infty.
\]

Lastly the uniform integrability statement of Proposition 5.1 follows from the last display and the following observation

\[
\mathbb{E} \left[ \int_{(z_1,z_2,y) \in \mathcal{B}_r,M \times [0,1]} \left( |z_1| + |z_2| + |y|^{r} \right) \mathcal{P}^\varepsilon,\Delta(dz_1 dz_2 dy dt) \right]
\]

\[
\leq \frac{C}{M} \mathbb{E} \left[ \int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,1]} \left( |z_1|^2 + |z_2|^2 + |y|^{2r} \right) \mathcal{P}^\varepsilon,\Delta(dz_1 dz_2 dy dt) \right],
\]

for some unimportant constant \( C < \infty \).
5.1.2. **Tightness of** \( \{ \eta_{\varepsilon,u}, \varepsilon > 0 \} \) **on** \( C([0,1]; \mathbb{R}^n) \). Next, we prove tightness of the family \( \{ \eta_{\varepsilon,u} \} \). It is sufficient to prove that for every \( \zeta > 0 \)

\[
\lim_{\rho \downarrow 0} \limsup_{\varepsilon \downarrow 0} \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1 | t_2 - t_1 < \rho} | \eta_{t_2}^{\varepsilon,u} - \eta_{t_1}^{\varepsilon,u} | > \zeta \right] = 0.
\]

This proof is the main source of additional complexity as compared to the large deviations case. The proof depends on several technical lemmas which are stated and proved in the Appendix.

From (14), we have

\[
\eta_{t_2}^{\varepsilon,u} - \eta_{t_1}^{\varepsilon,u} = \int_{t_1}^{t_2} \frac{\varepsilon b(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) + c(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) - \bar{\lambda}_1(\bar{X}_s)}{\sqrt{\varepsilon h(\varepsilon)}} ds + \int_{t_1}^{t_2} \sigma(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) \frac{u_1(s)}{h(\varepsilon)} ds + \int_{t_1}^{t_2} \frac{\lambda_1(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u})}{\sqrt{\varepsilon h(\varepsilon)}} ds.
\]

We can rewrite the first term in the form

\[
\int_{t_1}^{t_2} \frac{\varepsilon b(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) + c(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) - \bar{\lambda}_1(\bar{X}_s)}{\sqrt{\varepsilon h(\varepsilon)}} ds
\]

\[
= \int_{t_1}^{t_2} \frac{\varepsilon b(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) + c(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u}) - \lambda_1(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u})}{\sqrt{\varepsilon h(\varepsilon)}} ds
\]

\[
+ \int_{t_1}^{t_2} \frac{\lambda_1(X_{s}^{\varepsilon,u}, Y_{s}^{\varepsilon,u})}{\sqrt{\varepsilon h(\varepsilon)}} ds + \int_{t_1}^{t_2} \frac{\bar{\lambda}_1(\bar{X}_s)}{\sqrt{\varepsilon h(\varepsilon)}} ds.
\]
Then we have

\[
\begin{align*}
\mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \eta_{t_2}^{\varepsilon,u^\varepsilon} - \eta_{t_1}^{\varepsilon,u^\varepsilon} \right| > \zeta \right] & \leq \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \int_{t_1}^{t_2} \frac{\bar{b}(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} ds \right| > \frac{\zeta}{5} \right] \\
+ \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \int_{t_1}^{t_2} \frac{\lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \bar{\lambda}_1(X_s^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} ds \right| > \frac{\zeta}{5} \right] \\
+ \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \int_{t_1}^{t_2} \sigma(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_1(s) ds \right| > \frac{\zeta}{5} \right] \\
+ \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \frac{1}{h(\varepsilon)} \int_{t_1}^{t_2} \sigma(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) dW_s \right| > \frac{\zeta}{5} \right] \\
= \sum_{i=1}^{5} J_{i}^{\varepsilon,\rho}.
\end{align*}
\]

By Lemma B.4, B.5, B.6, B.3 we have for \( i = 1, 2, 3, 4 \) respectively that

\[
\lim_{\rho \downarrow 0} \limsup_{\varepsilon \downarrow 0} J_{i}^{\varepsilon,\rho} = 0.
\]

It remains to study the term \( J_{5}^{\varepsilon,\rho} \). By the conditions on \( \sigma \) and Lemma B.2

\[
M_t = \int_0^t \sigma(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) dW_s
\]

is a local square integrable martingale with continuous paths. Then, using again Lemma B.2 we have for a constant \( C < \infty \) that may change from line
to line and for $\nu > 0$ small enough such that $q_d(1 + \nu) < r$, we have

$$
P \left[ \sup_{0 \leq t_1 \leq t + \rho} \left| \int_{t_1}^{t} \sigma(X_s^{\epsilon,u^\epsilon}, Y_s^{\epsilon,u^\epsilon}) \, dW_s \right| > h(\epsilon)^{\frac{\zeta}{5}} \right]$$

$$\leq C (h(\epsilon)\zeta)^{-(1+\nu)} \mathbb{E} \left[ \sup_{0 \leq t_1 \leq t + \rho} \left| \int_{t_1}^{t} \sigma(X_s^{\epsilon,u^\epsilon}, Y_s^{\epsilon,u^\epsilon}) \, dW_s \right|^{2(1+\nu)} \right]$$

$$\leq C (h(\epsilon)\zeta)^{-(1+\nu)} \mathbb{E} \left[ \sup_{0 \leq t_1 \leq t + \rho} \left| \int_{t_1}^{t} |\sigma(X_s^{\epsilon,u^\epsilon}, Y_s^{\epsilon,u^\epsilon})| \, ds \right|^{(1+\nu)} \right]$$

from which the result follows by Lemma 3.3. With this, the proof of (20) is completed.

5.2. Proof of existence of viable pair. In the previous section, we have shown that the family of processes $\{\eta^{\epsilon,u^\epsilon}, P^{\epsilon,\Delta}\}$ is tight. It follows that for any subsequence of $\epsilon$ converging to 0, there exists a subsequence of $(\eta^{\epsilon,u^\epsilon}, P^{\epsilon,\Delta})$ which is convergent in distribution to some limit $(\bar{\eta}, \bar{P})$. The goal of this section is to show that $(\bar{\eta}, \bar{P})$ is a viable pair with respect to $(\theta_1, \mathcal{L}_{1,x})$ according to Definition 4.4. For this purpose we use the martingale problem formulation.

By the Skorokhod Representation Theorem, we may assume that there exists a probability space in which the desired convergence occurs w.p.1. By the proof of tightness for $\{P^{\epsilon,\Delta}\}$ and Fatou’s lemma,

$$\mathbb{E} \int_{Z \times Z \times Y \times [0,1]} \left[ |z_1|^2 + |z_2|^2 + |y|^{2r} \right] \bar{P}(dz_1 dz_2 dy dt) < \infty$$

which then implies that $\int_{Z \times Z \times Y \times [0,1]} \left[ |z_1|^2 + |z_2|^2 + |y|^{2r} \right] \bar{P}(dz_1 dz_2 dy dt) < \infty$ w.p.1. Here $r$ is the order of polynomial bound in $|y|$ of the $\theta_1$ function.

Therefore, to show that the limit point $(\bar{\eta}, \bar{P})$ is a viable pair, we must show that it satisfies equations (16), (17), and (18).

We begin with (16). Let $p_1$ and $p_2$ be positive integers. Let $F$ be a real valued, smooth function with compact support on $\mathbb{R}^n$. Let $\phi_j$, $j = 1, \ldots, p_1$, be real valued, smooth functions with compact support on $Z \times Z \times Y \times [0,1]$. Let $S$, $T$, and $t_i$, $i = 1, \ldots, p_2$, be nonnegative real numbers such that $t_i \leq S < S + T \leq 1$. Let $\zeta$ be a real valued, bounded and continuous function with compact support on $(\mathbb{R}^n)^{p_2} \times \mathbb{R}^{p_1p_2}$. For a measure $\hat{r} \in \mathcal{P}(Z \times Z \times Y \times [0,1])$ and $t \in [0,1]$, define

$$(\hat{r}, \phi_j)_t = \int_{Z \times Z \times Y \times [0,t]} \phi_j(z_1, z_2, y, s) \hat{r}(dz_1 dz_2 dy ds).$$
Define the operator $\tilde{L}_t^\varepsilon,\Delta$ by
\[
\tilde{L}_t^\varepsilon,\Delta F(\eta) = \int_{Z \times Z \times Y} (\nabla F(\eta)) \theta_1(\bar{X}_t, \eta, y, z_1, z_2) P_t^\varepsilon,\Delta(dz_1 dz_2 dy)
\]
where
\[
P_t^\varepsilon,\Delta(dz_1 dz_2 dy) = \frac{1}{\Delta} \int_t^{t+\Delta} 1_{dz_1}(u_1^\varepsilon(s)) 1_{dz_2}(u_2^\varepsilon(s)) 1_{dy}(Y_s^\varepsilon, u^\varepsilon) ds.
\]
Then to prove (16), it is sufficient to prove that as $\varepsilon \downarrow 0$,
\[
(22)
\]
\[
\mathbb{E} \left[ \zeta(\bar{\eta}_t^\varepsilon, u^\varepsilon, (P_t^\varepsilon, \Delta, \phi_j)_{t, i}, i \leq p, j \leq p_1) \left[ F(\eta^\varepsilon, u^\varepsilon) - F(\eta^\varepsilon, u^\varepsilon) - \int_{S}^{S+T} \tilde{L}_t^\varepsilon,\Delta F(\eta_t^\varepsilon) dt \right] \right] \rightarrow 0
\]
and
\[
(23)
\int_{S}^{S+T} \tilde{L}_t^\varepsilon,\Delta F(\eta_t^\varepsilon) dt - \int_{Z \times Z \times Y \times [S, S+T]} (\nabla F(\eta_t^\varepsilon)) \theta_1(\bar{X}_t, \bar{\eta}_t, y, z_1, z_2) \tilde{P}(dz_1 dz_2 dy dt) \rightarrow 0.
\]
For every real valued, continuous function $\phi$ with compact support and $t \in [0, 1]$,
\[
(P_t^\varepsilon, \phi) \rightarrow (\tilde{P}, \phi) \quad \text{w.p.1}.
\]

**Lemma 5.1.** Let $S > 0$ and $T > 0$ be positive numbers such that $S + T \leq 1$. Consider a continuous function $\xi: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{Y} \times Z \times Z \to \mathbb{R}$ that is bounded in the first argument, affine in the second argument, not growing faster than $|y|^r$ in the third argument and affine in the last two arguments. Assume that $(\eta_t^\varepsilon, u^\varepsilon, P_t^\varepsilon, \Delta) \Rightarrow (\bar{\eta}, \bar{P})$ in distribution for some subsequence of $\varepsilon \downarrow 0$, and that Conditions (2.1) and (2.2) (and in Regime 1, Condition (2.4)) hold. Then the following limits are valid in distribution along this subsequence:
\[
\int_{Z \times Z \times Y \times [S, S+T]} \xi(\bar{X}_t, \eta_t^\varepsilon, u^\varepsilon, y, z_1, z_2) P_t^\varepsilon,\Delta(dz_1 dz_2 dy dt)
\rightarrow \int_{Z \times Z \times Y \times [S, S+T]} \xi(\bar{X}_t, \bar{\eta}_t, y, z_1, z_2) \tilde{P}(dz_1 dz_2 dy dt)
\]
and
\[
\int_{S}^{S+T} \xi(X_t^\varepsilon, \eta_t^\varepsilon, u_t^\varepsilon, Y_t^\varepsilon, u_t^\varepsilon(t), u_t^\varepsilon(t)) dt
- \int_{Z \times Z \times Y \times [S, S+T]} \xi(\bar{X}_t, \eta_t^\varepsilon, u^\varepsilon, y, z_1, z_2) P_t^\varepsilon,\Delta(dz_1 dz_2 dy dt) \rightarrow 0.
\]

Lemma 5.1 is similar to Lemma 3.2 from [7] with the difference however that the function $\xi$ is not bounded in $y$. The proof of Lemma 5.1 follows the same lines as that of Lemma 3.2 from [7], where here we need to make use of the uniform integrability of $P_t^\varepsilon,\Delta$ with respect to both $(z_1, z_2)$ and $y$ from the second part of Proposition 5.1 in the same way that the uniform
We apply this lemma with $\xi(x, \eta, y, z_1, z_2) = (\nabla F(\eta))\theta_1(x, \eta, y, z_1, z_2)$. The first statement of Lemma 5.1 is equivalent to (23), and the second is equivalent (after applying the Itô formula to $F(\eta)$) to (22), which proves (16).

To prove (17), introduce the operator $\tilde{L}_{z_1, z_2, x}^\varepsilon$ for functions $F \in C^2(\mathcal{Y})$ defined by

$$
\tilde{L}_{z_1, z_2, x}^\varepsilon F(y) = \frac{1}{\delta} (\nabla F(y)) \left[ \frac{\varepsilon}{\delta} f(x, y) + g(x, y) + \sqrt{\varepsilon} h(\tau_1(x, y)z_1 + \sqrt{\varepsilon} h(\tau_2(x, y)z_2)
+ \frac{\varepsilon}{\delta^2} (\tau_1 \tau_1^T + 2\tau_2 \tau_2^T))(x, y) : \nabla \nabla F(y).
\right]
$$

Consider $\{F_\ell : \mathcal{Y} \to \mathbb{R}, \ell \in \mathbb{N} \}$ to be a smooth and dense family of bounded functions with bounded derivatives in $C^2(\mathcal{Y})$. Then it is easy to see that

$$
M_\varepsilon^\ell = F_\ell (Y_{\varepsilon, u^\varepsilon} - F_\ell(y_0) - \int_0^t \tilde{L}_{u_1(s), u_2(s), X_s, u^\varepsilon} F_\ell(Y_s^{\varepsilon, u^\varepsilon}) \, ds
$$
is an $\mathcal{F}_t$ martingale. Let $G(\varepsilon) = \delta^2 / \varepsilon$ and notice that $G(\varepsilon) \tilde{L}_{z_1, z_2, x}^\varepsilon$ converges to $\mathcal{L}_{1, x, y}^{\varepsilon, u^\varepsilon}$ as $\varepsilon \downarrow 0$. Next, we define the operator

$$
G_{z_1, z_2, x}^\varepsilon F_\ell(y) = (\nabla F_\ell(y)) (\tau_1(x, y) z_1 + \tau_2(x, y) z_2)
$$
and write

$$
(24)
$$

$$
G(\varepsilon) M_\varepsilon^\ell - G(\varepsilon) (F_\ell(Y_{\varepsilon, u^\varepsilon} - F_\ell(y_0))
$$

$$
- G(\varepsilon) \left[ \int_0^t \frac{1}{\Delta} \left[ \int_s^{s+\Delta} \tilde{L}_{u_1(\rho), u_2(\rho), X_\rho, u^\varepsilon} F_\ell(Y_\rho^{\varepsilon, u^\varepsilon}) \, d\rho \right] \, ds
$$

$$
- \int_0^t \tilde{L}_{u_1(s), u_2(s), X_s, u^\varepsilon} F_\ell(Y_s^{\varepsilon, u^\varepsilon}) \, ds \right]
$$

$$
= - \frac{\delta}{\varepsilon} \sqrt{h(\varepsilon)} \int_0^t \frac{1}{\Delta} \left[ \int_s^{s+\Delta} \left[ G_{u_1(\rho), u_2(\rho), X_\rho, u^\varepsilon} F_\ell(Y_\rho^{\varepsilon, u^\varepsilon}) - G_{u_1(\rho), u_2(\rho), X_\rho, u^\varepsilon} F_\ell(Y_\rho^{\varepsilon, u^\varepsilon}) \right] \, d\rho \right] \, ds
$$

$$
- \frac{\delta}{\varepsilon} \sqrt{h(\varepsilon)} \int_{Z \times \mathcal{Z} \times \mathcal{Y} \times [0, t]} G_{z_1, z_2, x, u^\varepsilon} F_\ell(y) P^{\varepsilon, \Delta}(dz_1 dz_2 dy \, ds)
$$

$$
- \frac{\delta}{\varepsilon} \int_0^t \frac{1}{\Delta} \left[ \int_s^{s+\Delta} (\nabla F_\ell(Y_\rho^{\varepsilon, u^\varepsilon})) [g(X_\rho^{\varepsilon, u^\varepsilon}, Y_\rho^{\varepsilon, u^\varepsilon}) - g(X_\rho^{\varepsilon, u^\varepsilon}, Y_\rho^{\varepsilon, u^\varepsilon})] \, d\rho \right] \, ds
$$

$$
- \frac{\delta}{\varepsilon} \int_{Z \times \mathcal{Z} \times \mathcal{Y} \times [0, t]} (\nabla F_\ell(y)) g(X_s^{\varepsilon, u^\varepsilon}, y) P^{\varepsilon, \Delta}(dz_1 dz_2 dy \, ds)
$$

$$
- \int_0^t \frac{1}{\Delta} \left[ \int_s^{s+\Delta} \left[ \mathcal{L}_{1, x, u^\varepsilon} F_\ell(Y_\rho^{\varepsilon, u^\varepsilon}) - \mathcal{L}_{1, x, u^\varepsilon} F_\ell(Y_\rho^{\varepsilon, u^\varepsilon}) \right] \, d\rho \right] \, ds
$$

$$
- \int_{Z \times \mathcal{Z} \times \mathcal{Y} \times [0, t]} \mathcal{L}_{1, x, u^\varepsilon} F_\ell(y) P^{\varepsilon, \Delta}(dz_1 dz_2 dy \, ds).
Now consider each of these terms as $\varepsilon \downarrow 0$. The left hand side of (24) goes to zero since:

(a) $M^\varepsilon_t$ is square integrable, so $G(\varepsilon)M^\varepsilon_t \downarrow 0$ in probability as $\varepsilon \downarrow 0$,

(b) $F_\ell$ is bounded, $G(\varepsilon) \left[ F_\ell(Y^\varepsilon,u^\varepsilon) - F_\ell(y_0) \right]$ converges to zero uniformly as $\varepsilon \downarrow 0$, and

(c) $\Delta \downarrow 0$ as $\varepsilon \downarrow 0$,

$$G(\varepsilon) \left[ \int_0^t \frac{1}{\Delta} \left[ \int_s^{s+\Delta} \tilde{\mathcal{L}}_{u^\varepsilon_1(\rho),u^\varepsilon_2(\rho),X^\varepsilon_\rho,\mathfrak{U}^\varepsilon_\rho} F_\ell(Y^\varepsilon_\rho,u^\varepsilon_\rho) d\rho \right] ds - \int_0^t \tilde{\mathcal{L}}_{u^\varepsilon_1(s),u^\varepsilon_2(s),X^\varepsilon_\rho,u^\varepsilon_\rho} F_\ell(Y^\varepsilon_\rho,u^\varepsilon_\rho) ds \right]$$

converges to zero in probability.

We next study the right hand side of (24). Tightness of $\{X^\varepsilon,u^\varepsilon, \varepsilon > 0\}$ implies that the first term, third term, and fifth term on the right side converge to zero in probability as $\varepsilon \downarrow 0$. (Tightness of $\{X^\varepsilon,u^\varepsilon, \varepsilon > 0\}$ follows immediately from tightness of $\{\eta^\varepsilon,u^\varepsilon, \varepsilon > 0\}$ by (12).)

Uniform integrability of $P^\varepsilon, \Delta$ and the fact that $\delta/\varepsilon \downarrow 0$ imply that the second and fourth terms on the right side converge to zero in probability as $\varepsilon \downarrow 0$.

Therefore,

$$\int_{\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,t]} \mathcal{L}_{1,X^\varepsilon,u^\varepsilon} F_\ell(y) P^\varepsilon,\Delta(dz_1 dz_2 dy ds) \to 0$$

in probability as $\varepsilon \downarrow 0$.

This implies (17) by continuity in $t$ and density of $\{F_\ell, \ell \in \mathbb{N}\}$.

Proof of (18) is identical to [7] or [17]. More explicitly, by the fact that $P^\varepsilon,\Delta(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,t]) = t$, along with $P(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times \{t\}) = 0$ and the continuity of the mapping $t \to P(\mathcal{Z} \times \mathcal{Z} \times \mathcal{Y} \times [0,t])$, the property holds.

5.3. **Proof of Laplace principle lower bound.** We now prove the Laplace principle lower bound. We want to show that for all bounded, continuous functions $a$ mapping $C([0,1]; \mathbb{R}^n)$ into $\mathbb{R}$,

$$\liminf_{\varepsilon \downarrow 0} - \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right]$$

$$\geq \inf_{(\xi,P) \in \mathcal{V}(\theta_1,\mathcal{L}_{1,\alpha})} \left[ \frac{1}{2} \int \left[ |z_1|^2 + |z_2|^2 \right] P(dz_1 dz_2 dy ds) + a(\xi) \right].$$

It is sufficient to prove the lower limit along any subsequence such that

$$- \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right]$$
converges. Such a subsequence exists because \(-1/h^2(\varepsilon) \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right] \leq \|a\|_\infty. By Lemma 3.1 we may assume that
\[
\sup_{\varepsilon > 0} \mathbb{E} \int_0^1 |u^\varepsilon(s)|^2 ds \leq N.
\]
for some constant \(N\).

We construct the family of occupation measures \(P^{\varepsilon,\Delta}\), and the family \(\{(\eta^\varepsilon, u^\varepsilon, P^{\varepsilon,\Delta}) \mid \varepsilon > 0\}\) is tight. Hence, for any subsequence of \(\varepsilon \downarrow 0\) there is a further subsequence for which
\[
\liminf_{\varepsilon \downarrow 0} \left( -\frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right] - \varepsilon \right)
\]
\[
= \liminf_{\varepsilon \downarrow 0} \left( \mathbb{E} \left[ \frac{1}{2} \int_0^1 |u^\varepsilon(s)|^2 ds + a(\eta^\varepsilon, u^\varepsilon) \right] \right)
\]
\[
= \liminf_{\varepsilon \downarrow 0} \left( \mathbb{E} \left[ \frac{1}{2} \int_{Z \times Z \times Y} |z_1|^2 + |z_2|^2 P^{\varepsilon,\Delta}(dz_1 dz_2 dy dt) + a(\eta^\varepsilon, u^\varepsilon) \right] \right)
\]
\[
\geq \mathbb{E} \left[ \frac{1}{2} \int_{Z \times Z \times Y} |z_1|^2 + |z_2|^2 \tilde{P}(dz_1 dz_2 dy dt) + a(\eta) \right]
\]
\[
\geq \inf_{(\xi, P) \in \mathcal{V}(\theta_1, L_{1,x})} \left\{ \frac{1}{2} \int_{Z \times Z \times Y} |z_1|^2 + |z_2|^2 P(dz_1 dz_2 dy dt) + a(\xi) \right\}.
\]

This concludes the proof of the Laplace principle lower bound.

5.4. Proof of compactness of level sets of \(S(\cdot)\). We want to prove that for each \(s < \infty\), the set
\[
\Xi_s = \{ \xi \in \mathcal{C}([0, 1]; \mathbb{R}^n) : S(\xi) \leq s \}
\]
is a compact subset of \(\mathcal{C}([0, 1]; \mathbb{R}^n)\). The proof is analogous to the proof of the lower bound. We need to show precompactness of \(\Xi_s\) and that it is a closed set.

Precompactness of the pair \(\{(\xi^n, P^n), n > 0\}\) follows by standard arguments, see for example [7]. Next we must show that the limit of a sequence of viable pairs is a viable pair. Fix \(K < \infty\) and consider any convergent sequence \(\{(\xi^n, P^n), n > 0\}\) such that for every \(n > 0\), \((\xi^n, P^n) \in \mathcal{V}(\theta_1, L_{1,x})\) and
\[
\int_{Z \times Z \times Y \times [0, 1]} |z_1|^2 + |z_2|^2 + |y|^{2r} P^n(dz_1 dz_2 dy dt) < K,
\]
where \(r\) is the order of the polynomial bound in \(|y|\) of \(\theta_1\).
Since \((\xi^n, P^n)\) is a viable pair, we get that
\[
\xi^n_t = \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,t]} \theta_1(\bar{X}_{s}, \xi^n_s, y, z_1, z_2) P^n(dz_1 \, dz_2 \, dy \, ds)
\]
and
\[
\int_0^t \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y}} \mathcal{L}_{1,\bar{X}_s} F(y) P^n(dz_1 \, dz_2 \, dy \, ds) = 0
\]
for every \(t \in [0,1]\) and every \(F \in C^2(\mathbb{Y})\). Then by the convergence of \((\xi^n, P^n)\) to \((\xi, P)\), we get that \((\xi, P) \in \mathcal{V}(\theta_1, \mathcal{L}_{1,x})\).

Finally, we must prove lower semicontinuity, that is
\[
\liminf_{n \to \infty} S(\xi^n) \geq S(\xi)
\]
Without loss of generality, we may assume that there is some \(M < \infty\) such that \(\liminf_{n \to \infty} S(\xi^n) \leq M\). Also, by the definition of \(S(\xi^n)\), we obtain that one can find measures \(\{P^n, n < \infty\}\) such that \((\xi^n, P^n) \in \mathcal{V}(\theta_1, \mathcal{L}_{1,x})\),
\[
\sup_{n < \infty} \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,1]} [|z_1|^2 + |z_2|^2] P^n(dz_1 \, dz_2 \, dy \, dt) < M + 1
\]
and
\[
S(\xi^n) \geq \frac{1}{2} \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,1]} [|z_1|^2 + |z_2|^2] P^n(dz_1 \, dz_2 \, dy \, dt) - \frac{1}{n}
\]
Then by Fatou’s lemma we have
\[
\liminf_{n \to \infty} S(\xi^n) \geq \liminf_{n \to \infty} \left[ \frac{1}{2} \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,1]} [|z_1|^2 + |z_2|^2] P^n(dz_1 \, dz_2 \, dy \, dt) - \frac{1}{n} \right]
\]
\[
\geq \frac{1}{2} \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,1]} [|z_1|^2 + |z_2|^2] P(dz_1 \, dz_2 \, dy \, dt)
\]
\[
\geq \inf_{(\xi, P) \in \mathcal{V}(\theta_1, \mathcal{L}_{1,x})} \left\{ \frac{1}{2} \int_{\mathbb{Z} \times \mathbb{Z} \times \mathbb{Y} \times [0,1]} [|z_1|^2 + |z_2|^2] P(dz_1 \, dz_2 \, dy \, dt) \right\} = S(\xi).
\]

5.5. Proof of Laplace principle upper bound and representation formula. The first step is to establish the equivalence of the control formulation to the relaxed control formulation, as in [7]. Let us briefly recall how this is done.

The action functional \(S(\xi)\) can be written in terms of a local action functional, i.e.,
\[
S(\xi) = \int_0^1 L^{\text{rel}}(\bar{X}_s, \xi_s, \dot{\xi}_s) \, ds.
\]
This follows from the definition of a viable pair by setting
\[
L_{\text{rel}}(x, \eta, \beta) = \inf_{P \in \mathcal{A}_{x, \eta, \beta}^{\text{rel}}} \int_{Z \times Z \times Y} \frac{1}{2} \left[ |z_1|^2 + |z_2|^2 \right] P(dz_1 dz_2 dy)
\]
where
\[
\mathcal{A}_{x, \eta, \beta}^{\text{rel}} = \left\{ P \in \mathcal{P}(Z \times Z \times Y) : \int_{Z \times Z \times Y} L_{1,x} F(y) P(dz_1 dz_2 dy) = 0 \right. \\
\left. \text{for all } F \in C^2(Y), \int_{Z \times Z \times Y} \left[ |z_1|^2 + |z_2|^2 + |y|^{2r} \right] P(dz_1 dz_2 dy) < \infty, \right.
\]
and \( \beta = \int_{Z \times Z \times Y} \theta_1(x, \eta, y, z_1, z_2) P(dz_1 dz_2 dy) \). The constant \( r \) in the above expression is the order of the polynomial bound of \( |y| \) in \( \theta_1 \).

Note that any measure \( P \in \mathcal{P}(Z \times Z \times Y) \) can be decomposed in the form
\[
(25) \quad P(dz_1 dz_2 dy) = \nu(dz_1 dz_2) \mu(dy)
\]
where \( \mu \) is a probability measure on \( Y \) and \( \nu \) is a stochastic kernel on \( Z \times Z \) given \( Y \). Following the terminology of [7], we refer to this as a “relaxed” formulation.

Inserting (25) into (17) and noticing that \( L_{1,x} \) does not depend on the control variables, we obtain that for every \( F \in C^2(Y) \),
\[
\int_Y L_{1,x} F(y) \mu(dy) = 0.
\]

The nondegeneracy of \( (\tau_1 \tau_1^T + \tau_2 \tau_2^T) \) and the previous equation show that \( \mu(dy) \) is the unique invariant measure corresponding to the operator \( L_{1,x} \) (i.e. \( \mu(dy) = \mu_{1,x}(dy) \)).

Because the cost is convex in \( z \) and \( \theta_1 \) is affine in \( z \), the relaxed control formulation is equivalent to the ordinary control formulation of the local rate function:
\[
(26) \quad L_0^o(x, \eta, \beta) = \inf_{(v, \mu) \in \mathcal{A}_{x, \eta, \beta}^o} \frac{1}{2} \int_Y |v(y)|^2 \mu_{1,x}(dy)
\]
where
\[
\mathcal{A}_{x, \eta, \beta}^o = \left\{ (v(\cdot) = (v_1(\cdot), v_2(\cdot)) : Y \to \mathbb{R}^{2m}, \mu \in \mathcal{P}(Y) : (v, \mu) \text{ satisfy} \right. \\
\left. \int_Y L_{1,x} F(y) \mu_{1,x}(dy) = 0 \text{ for all } F \in C^2(Y), \int_Y \left[ |v(y)|^2 + |y|^{2r} \right] \mu(dy) < \infty, \right.
\]
and \( \beta = \int_Y \theta_1(x, \eta, y, v_1(y), v_2(y)) \mu_{1,x}(dy) \).
The equivalence of $L^{rel}(x, \eta, \beta)$ and $L^o(x, \eta, \beta)$ follows from Jensen’s inequality and the fact that $\theta_1(x, \eta, y, z_1, z_2)$ and $L_{1,x}$ are affine in $z_1$ and $z_2$.

The following result is a key statement for the equivalence of Theorems 2.1 and 4.2.

**Theorem 5.1.** Under Conditions 2.1, 2.2, and Condition 2.4, the infimization problem (26) has the explicit solution

$$L^o(x, \eta, \beta) = \frac{1}{2} (\beta - \kappa(x, \eta))^T q^{-1}(x)(\beta - \kappa(x, \eta))$$

where $\kappa(x, \eta)$ and $q(x)$ are given by (2) and (1). Furthermore, with $\alpha_1(x, y)$, $\alpha_2(x, y)$ given by (11), the control $v(y) = (v_1(y), v_2(y))$ defined by

$$v_1(y) = \alpha_1(x, y)^T q^{-1}(x)(\beta - \kappa(x, \eta))$$

$$v_2(y) = \alpha_2(x, y)^T q^{-1}(x)(\beta - \kappa(x, \eta))$$

attains the infimum in the variational problem (26).

**Proof.** Observe that for any $v \in A^o_{x,\eta,\beta}$,

$$\int_\mathcal{Y} |v(y)|^2 \mu_{1,x}(dy) \geq (\beta - \kappa(x, \eta))^T q^{-1}(x)(\beta - \kappa(x, \eta)).$$

This follows because any $v \in A^o_{x,\eta,\beta}$ satisfies

$$\beta = \int_\mathcal{Y} \theta_1(x, \eta, y, v_1(y), v_2(y)) \mu_{1,x}(dy)$$

$$= \kappa(x, \eta) + \int_\mathcal{Y} \left[ \sigma(x, y)v_1(y) + (\nabla_y \chi(x, y))(\tau_1(x, y)v_1(y) + \tau_2(x, y)v_2(y)) \right] \mu_{1,x}(dy)$$

$$= \kappa(x, \eta) + \int_\mathcal{Y} \left[ \alpha_1(x, y)v_1(y) + \alpha_2(x, y)v_2(y) \right] \mu_{1,x}(dy).$$

Then treating $x$ and $\eta$ as parameters and applying Lemma 5.1 from [4] to the relation above, we get the claim. Next, observe that by choosing (with $x$ and $\eta$ treated as parameters)

$$v_1(y) = \alpha_1(x, y)^T q^{-1}(x)(\beta - \kappa(x, \eta))$$

$$v_2(y) = \alpha_2(x, y)^T q^{-1}(x)(\beta - \kappa(x, \eta))$$

we have

$$\int_\mathcal{Y} |v(y)|^2 \mu_{1,x}(dy) = (\beta - \kappa(x, \eta))^T q^{-1}(x)(\beta - \kappa(x, \eta)).$$

This completes the proof of the theorem. \qed

Now we can prove the Laplace principle upper bound. We must show that for all bounded, continuous functions $a$ mapping $\mathcal{C}([0, 1]; \mathbb{R}^n)$ into $\mathbb{R}$

$$\limsup_{\varepsilon \downarrow 0} - \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon)a(\eta^\varepsilon) \right\} \right] \leq \inf_{\xi \in \mathcal{C}([0, 1]; \mathbb{R}^n)} [S(\xi) + a(\xi)].$$
Let $\zeta > 0$ be given and consider $\psi \in C([0, 1]; \mathbb{R}^n)$ with $\psi_0 = 0$ such that

$$S(\psi) + a(\psi) \leq \inf_{\xi \in C([0, 1]; \mathbb{R}^n)} [S(\xi) + a(\xi)] + \zeta < \infty.$$ 

Since $a$ is bounded, this implies that $S(\psi) < \infty$, and thus $\psi$ is absolutely continuous. Theorem 5.1 shows that $L^\alpha(x, \eta, \beta)$ is continuous and finite at each $(x, \eta, \beta) \in \mathbb{R}^{3n}$. By a mollification argument we can assume that $\psi$ is piecewise continuous, see Section 6.5 in [5]. Given this $\psi$ define

$$\bar{u}_1(t, x, \eta, y) = \alpha_1(x, y)^T q^{-1}(x)(\dot{\psi}_t - \kappa(x, \eta))$$

$$\bar{u}_2(t, x, \eta, y) = \alpha_2(x, y)^T q^{-1}(x)(\dot{\psi}_t - \kappa(x, \eta))$$

with $\alpha_1$ and $\alpha_2$ defined as in Theorem 5.1. Define a control in feedback form by

$$\bar{u}^\varepsilon(t) = (\bar{u}_1(t), \bar{u}_2(t)) = (\bar{u}_1(t, \bar{X}_t, \eta_\varepsilon^\varepsilon, Y_\varepsilon^\varepsilon), \bar{u}_2(t, \bar{X}_t, \eta_\varepsilon^\varepsilon, Y_\varepsilon^\varepsilon)).$$

Then $\eta_\varepsilon^\varepsilon, \bar{u}^\varepsilon \to \bar{\eta}$ in distribution, where w.p.1

$$\bar{\eta}_t = \int_0^t \kappa(\bar{X}_s, \bar{\eta}_s) ds + \int_0^t \left[ \int \left[ (\sigma(\bar{X}_s, y) + (\nabla y \chi(X_s, y)) \tau_1(X_s, y) \right] \bar{u}_1(s) \\
+ (\nabla y \chi(X_s, y)) \tau_2(X_s, y) \bar{u}_2(s) \right] \mu_{1,\bar{X}_s}(dy) \right] ds$$

$$= \int_0^t \kappa(\bar{X}_s, \bar{\eta}_s) ds + \int_0^t \left[ \int \left[ \alpha_1 \alpha_1^T(\bar{X}_s, y) + \alpha_2 \alpha_2^T(\bar{X}_s, y) \right] \mu_{1,\bar{X}_s}(dy) \right] q^{-1}(\bar{X}_s)(\dot{\psi}_s - \kappa(\bar{X}_s, \bar{\eta}_s)) ds$$

$$= \int_0^t \kappa(\bar{X}_s, \bar{\eta}_s) ds + \int_0^t q(\bar{X}_s) q^{-1}(\bar{X}_s)(\dot{\psi}_s - \kappa(\bar{X}_s, \bar{\eta}_s)) ds$$

$$= \int_0^t \dot{\psi}_s ds = \psi_t.$$

The cost satisfies

$$\mathbb{E} \left( \frac{1}{2} \int_0^1 |\bar{u}_s^\varepsilon|^2 ds - \frac{1}{2} \int_0^1 \int \bar{u}(s, \bar{X}_s, \bar{\eta}_s, y)^2 \mu_{1,\bar{X}_s}(dy) ds \right)^2 \to 0 \text{ as } \varepsilon \downarrow 0.$$ 

Theorem 5.1 implies that

$$\mathbb{E} \frac{1}{2} \int_0^1 \int |\bar{u}(s, \bar{X}_s, \bar{\eta}_s, y)|^2 \mu_{1,\bar{X}_s}(dy) ds = \mathbb{E} \mathbb{S}(\bar{\eta}) = S(\psi).$$
Then we obtain
\[
\limsup_{\varepsilon \downarrow 0} \frac{1}{h^2(\varepsilon)} \log \mathbb{E} \left[ \exp \left\{ -h^2(\varepsilon) a(\eta^\varepsilon) \right\} \right] = \limsup_{\varepsilon \downarrow 0} \inf_{u \in A} \mathbb{E} \left[ \frac{1}{2} \int_0^1 |u(t)|^2 \, dt + a(\eta^\varepsilon, u) \right]
\]
\[
\leq \limsup_{\varepsilon \downarrow 0} \mathbb{E} \left[ \frac{1}{2} \int_0^1 |\tilde{u}(t)|^2 \, dt + a(\eta^\varepsilon, \tilde{u}^\varepsilon) \right]
\]
\[
= \mathbb{E} \left[ \frac{1}{2} \int_0^1 \int_y |\tilde{u}(s, \bar{X}_s, \bar{\eta}_s, y)|^2 \mu_{1, \bar{X}_s}(dy) \, ds + a(\bar{\eta}) \right]
\]
\[
= \left[ S(\psi) + a(\psi) \right] \inf_{\xi \in C([0,1]; \mathbb{R}^n)} [S(\xi) + a(\xi)] + \zeta.
\]

Since \( \zeta > 0 \) is arbitrary, the upper bound is proved. Furthermore, we have an explicit representation formula for the action functional, given by
\[
S(\xi) = \frac{1}{2} \int_0^1 \left( \dot{\xi}_s - \kappa (\bar{X}_s, \xi_s) \right)^T q^{-1}(\bar{X}_s) \left( \dot{\xi}_s - \kappa (\bar{X}_s, \xi_s) \right) \, ds
\]
if \( \xi \in C([0,1]; \mathbb{R}^n) \) is absolutely continuous, and \( \infty \) otherwise.

6. Comments on the Proofs for Regime 2

The structure of the proof for Regime 2 is identical to that of Regime 1, after replacing \( \lambda_1, \theta_1, L_1, \Phi_1 \), and \( \mu_1 \) by \( \lambda_2, \theta_2, L_2, \Phi_2 \), and \( \mu_2 \) respectively. Hence we do not repeat it here. For example in Regime 2, applying the Itô formula to \( \Phi_2(X_t^{\varepsilon, u^\varepsilon}, \xi_t^{\varepsilon, u^\varepsilon}) \) and some term rearranging shows that
\[
\eta_t^{\varepsilon, u^\varepsilon} = \int_0^t \left[ j_2 b(\bar{X}_s, Y_s^{\varepsilon, u^\varepsilon}) + (\nabla_y \Phi_2(\bar{X}_s, Y_s^{\varepsilon, u^\varepsilon})) \left[ \tau_1 (\bar{X}_s, Y_s^{\varepsilon, u^\varepsilon}) u_1(s) + \tau_2 (\bar{X}_s, Y_s^{\varepsilon, u^\varepsilon}) u_2(s) \right] \right. 
\]
\[
+ (\nabla_2 \lambda_2(\bar{X}_s)) u_s^{\varepsilon, u^\varepsilon} + \sigma(\bar{X}_s, Y_s^{\varepsilon, u^\varepsilon}) u_1(s) + j_2 (\nabla_y \Phi_2(\bar{X}_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon})) f(\bar{X}_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) 
\]
\[
\left. + \frac{j_2}{2} \left( \left( \tau_1 \tau_1^T + \tau_2 \tau_2^T \right)(\bar{X}_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) : \nabla_y \nabla_y \Phi_2(\bar{X}_s^{\varepsilon, u^\varepsilon}, Y_s^{\varepsilon, u^\varepsilon}) \right) \right] \, ds + \tilde{R}^\varepsilon,
\]
where \( \tilde{R}^\varepsilon \) contains the additional terms which go to zero as \( \varepsilon \downarrow 0 \). However, the necessary statements that were needed for Regime 1 and which are proved in Appendix B do need some special care. We address these in Appendix C.

APPENDIX A. Regularity results

The following theorem collects results from [14] and [15] that are used in this paper.

**Theorem A.1.** Let Conditions 2.1 and 2.2 be satisfied. In Regime \( i = 1, 2 \) we have that,
(i) There exists a unique invariant measure $\mu_{i,x}(dy)$ associated with the operator $L_{i,x}$. For all $x \in \mathbb{R}^n$ and $q \in \mathbb{N}$,
\[
\int |y|^q \mu_{i,x}(dy) < \infty.
\]
Moreover, $\mu_{i,x}$ has a density which is twice differentiable in $x$.

(ii) Assume that $G(x,y) \in C^{2,\alpha}(\mathbb{R}^n \times \mathcal{Y})$. Then
\[
\bar{G}(x) = \int_{\mathcal{Y}} G(x,y) \mu_{i,x}(dy)
\]
is twice differentiable in $x$.

(iii) Assume that $F(x,y) \in C^{2,\alpha}(\mathbb{R}^n \times \mathcal{Y})$, 
\[
\int_{\mathcal{Y}} F(x,y) \mu_{i,x}(dy) = 0,
\]
and that for some positive constants $K$ and $q_F$,
\[
|F(x,y)| + \|\nabla_x F(x,y)\| + \|\nabla_x \nabla_x F(x,y)\| \leq K(1 + |y|^{q_F}).
\]
Then there is a unique solution from the class of functions which grow at most polynomially in $|y|$ to
\[
L_{i,x}u(x,y) = -F(x,y), \quad \int_{\mathcal{Y}} u(x,y) \mu_{i,x}(dy) = 0.
\]
Moreover, the solution satisfies $u(\cdot,y) \in C^2$ for every $y \in \mathcal{Y}$, $\nabla_x \nabla_x u \in C(\mathbb{R}^n \times \mathcal{Y})$, and there exist positive constants $K'$ that change from line to line such that
\[
|u(x,y)| \leq K'(1 + |y|)^{(q_F+1-r)+},
\]
\[
\|\nabla_y u(x,y)\| \leq K'(1 + |y|^{(q_F+1-r)+} + |y|^{q_F})
\]
\[
\|\nabla_x u(x,y)\| \leq K'(1 + |y|^{(q_F+1-r)+} + |y|^{(q_F+2(1-r)))+}),
\]
\[
\|\nabla_x \nabla_x u(x,y)\| \leq K'(1 + |y|^{(q_F+1-r)+} + |y|^{(q_F+2(1-r)))+} + |y|^{(q_F+3(1-r)))+})
\]
where $r$ is as defined in Condition 2.2.

Proof. (i) Conditions 2.1 and 2.2 imply that $L_{i,x}$ satisfies the conditions for Proposition 1 in [14] and Theorem 1 in [15]. The first statement is due to Proposition 1 in [14] and the second statement is due to Theorem 1 in [15].

(ii) Conditions 2.1 and 2.2 and the condition on $G$ imply that Theorem 2 in [15] holds, so (ii) holds.

(iii) Conditions 2.1 and 2.2 and the conditions on $F$ imply that Theorem 3 in [15] holds, which implies the existence and smoothness of $u$. The corresponding growth conditions follow from Theorem 2 of [14], appropriately translated to our case. Notice that Theorem 2 of [14] has a statement for the growth only for the solution and its $y$–derivative. The statements for
the \(x\)-derivatives follow, for example, by differentiating the equation and re-applying Theorem 2 of [14] to the new equation.  

**Appendix B. Lemmas for Regime 1**

**Lemma B.1.** Assume Conditions 2.1, 2.2 and 2.5. Let \((X_t^{\varepsilon,u^{\varepsilon}}, Y_t^{\varepsilon,u^{\varepsilon}})\) be the strong solution to (13). Then the infimum of the representation in (11) can be taken over all controls such that
\[
\int_0^1 |u^{\varepsilon}(s)|^2 \, ds < N, \text{ almost surely,}
\]
where the constant \(N\) does not depend on \(\varepsilon\) or \(\delta\).

**Proof.** The proof is standard, but we recall it here for the readers convenience. Without loss of generality, we can consider a function \(a(x)\) that is bounded and uniformly Lipschitz continuous in \(C([0,1];\mathbb{R}^n)\). Namely, there exists a constant \(L_a\) such that
\[
|a(x) - a(y)| \leq L_a \|x - y\|
\]
(where \(\|\cdot\|\) is the supremum norm) and \(\|a\|_{\infty} = \sup_{x\in C([0,1];\mathbb{R}^n)} |a(x)| < \infty\).

Fix \(\zeta > 0\). There exists a family of controls \(\{u^{\varepsilon}, \varepsilon > 0\}\) in \(A\) such that for every \(\varepsilon > 0\),
\[
-\frac{1}{h^2(\varepsilon)} \log \mathbb{E} [\exp\{-h^2(\varepsilon)a(\eta^{\varepsilon})\}] \geq \mathbb{E} \left[\frac{1}{2} \int_0^1 |u^{\varepsilon}(s)|^2 \, ds + a(\eta^{\varepsilon,u^{\varepsilon}})\right] - \zeta.
\]
Then each control \(u^{\varepsilon}\) satisfies
\[
\sup_{\varepsilon > 0} \mathbb{E} \left[\frac{1}{2} \int_0^1 |u^{\varepsilon}(s)|^2 \, ds\right] \leq 2\|a\|_{\infty} + \zeta.
\]
By the proof of Theorem 4.4 in [3], it is enough to assume that for given \(\zeta > 0\) the controls satisfy the bound
\[
\int_0^1 |u^{\varepsilon}(s)|^2 \, ds < N \text{ where } N \geq \frac{4\|a\|_{\infty}(4\|a\|_{\infty} + \zeta)}{\zeta}.
\]

**Lemma B.2.** Let Conditions 2.1, 2.2, 2.3, and 2.5 be satisfied. For \(N \in \mathbb{N}\), let \(u^{\varepsilon} \in A\) such that almost surely
\[
\sup_{\varepsilon > 0} \int_0^1 |u^{\varepsilon}(s)|^2 \, ds < N.
\]

Then, with \(r > 0\) from Condition 2.2 we have for any \(T \leq 1\) there exist \(\varepsilon_0 > 0\) small enough such that
\[
\sup_{\varepsilon \in (0,\varepsilon_0)} \mathbb{E} \int_0^T |Y_{s}^{\varepsilon,u^{\varepsilon}}|^{2r} \, ds \leq K(N,T,r)
\]
for some finite constant $K(N,T,r)$ that may depend on $(N,T,r)$, but not on $\varepsilon, \delta(\varepsilon)$.

**Proof.** By the Markov property it is enough to check what happens when the process $Y^{\varepsilon,u^\varepsilon}$ is outside a compact subset of $Y$. For this purpose, with $R < \infty$ to be chosen, let us define

$$\tau_R = \inf \left\{ t > 0 : |Y_t^{\varepsilon,u^\varepsilon}| < R \right\}$$

and assume that the initial condition is such that $|y_0| > R$. For notational convenience, we will write $(X,Y)$ instead of $(X^{\varepsilon,u^\varepsilon}, Y^{\varepsilon,u^\varepsilon})$. Without loss of generality and for exposition purposes we shall set $g = 0$ (since it is assumed to be bounded) and $\tau_1 = 0$ (the argument is exactly the same if both $\tau_1$ and $\tau_2$ are non-zero). By Condition 2.2 we have that uniformly in both $x$ and $y$ and for any constant $\beta$

$$\beta - 2 \frac{\langle \tau_2 \tau_2^T(x,y) y, y \rangle}{|y|^2} + \frac{1}{2} \text{Tr}(\tau_2 \tau_2^T)(x,y) \leq \rho,$$

for some fixed constant $\rho > 0$. Hence, considering $t \leq T$, the Itô formula gives for $\beta > 0$ (to be chosen)

$$E|Y_{t\wedge \tau_R}|^\beta = |y_0|^\beta + \beta \frac{\varepsilon}{\delta^2} E \int_0^{t\wedge \tau_R} |Y_s|^{\beta - 2} \left( \langle Y_s, f(X_s, Y_s) \rangle + \frac{\beta - 2}{2} \frac{\langle \tau_2 \tau_2^T(X_s,Y_s) Y_s, Y_s \rangle}{|Y_s|^2} + \frac{1}{2} \text{Tr}(\tau_2 \tau_2^T)(X_s,Y_s) \right) ds +$$

$$+ \frac{\sqrt{\varepsilon h(\varepsilon)}}{\delta} \beta E \int_0^{t\wedge \tau_R} |Y_s|^{\beta - 2} \langle Y_s, \tau_2(X_s,Y_s) u^\varepsilon_2(s) \rangle ds +$$

$$\leq |y_0|^\beta + \beta \frac{\varepsilon}{\delta^2} E \int_0^{t\wedge \tau_R} |Y_s|^{\beta - 2} (-\Gamma |Y_s|^r + 1 + \rho) ds +$$

$$+ \frac{\sqrt{\varepsilon h(\varepsilon)}}{\delta} \beta E \int_0^{t\wedge \tau_R} \left( |Y_s|^{2\beta - 2} + \|\tau_2\|^2 |u^\varepsilon(s)|^2 \right) ds$$

where Condition 2.3 was used.

Choosing now $R$ large enough such that $R^{1+r} > \frac{2\rho}{\varepsilon}$ and recalling that $\sup_{\varepsilon \in (0,1)} E \int_0^T |u^\varepsilon(s)|^2 ds \leq N$, we can continue the last inequality as follows

$$E|Y_{t\wedge \tau_R}|^\beta \leq |y_0|^\beta - \beta \frac{\varepsilon}{2} E \int_0^{t\wedge \tau_R} |Y_s|^{\beta - 1} ds + \frac{\sqrt{\varepsilon h(\varepsilon)}}{2\delta} \beta \|\tau_2\|^2 N$$

$$+ \frac{\sqrt{\varepsilon h(\varepsilon)}}{2\delta} \beta E \int_0^{t\wedge \tau_R} |Y_s|^{2\beta - 2} ds$$
Choosing now $\beta \leq r + 1$, we obtain
\[
\mathbb{E}|Y_{1\land \tau_R}|^\beta \leq |y_0|^\beta - \frac{\beta \Gamma}{2} \mathbb{E} \int_0^{t^\land \tau_R} |Y_s|^{|\beta - r - 1}| \, ds + \frac{\sqrt{\varepsilon} h(\varepsilon)}{2\delta} \beta \|\tau_2\|^2 N
\]
\[
+ \frac{\sqrt{\varepsilon} h(\varepsilon)}{2\delta} \mathbb{E} \int_0^{t^\land \tau_R} |Y_s|^{|\beta - r - 1}| \, ds
\]
Choosing next $\varepsilon, \delta$ sufficiently small such that $\delta \sqrt{\varepsilon} h(\varepsilon) < \Gamma/4$ we obtain
\[
\mathbb{E}|Y_{1\land \tau_R}|^\beta \leq |y_0|^\beta - \frac{\beta \Gamma}{4} \mathbb{E} \int_0^{t^\land \tau_R} |Y_s|^{|\beta - r - 1}| \, ds + \frac{\sqrt{\varepsilon} h(\varepsilon)}{\sqrt{\varepsilon}} \beta \|\tau_2\|^2 N
\]
which then gives by comparison
\[
\mathbb{E} \int_0^{t^\land \tau_R} |Y_s|^{|\beta - r - 1}| \, ds \leq \frac{4}{\beta \Gamma} \frac{\delta^2}{\varepsilon} |y_0|^\beta + \frac{2}{\Gamma} \frac{\delta h(\varepsilon)}{\sqrt{\varepsilon}} \|\tau_2\|^2 N.
\]
Since $\beta \leq r + 1$, we choose $\beta = r + 1$, which concludes the proof of the lemma. \(\square\)

**Lemma B.3.** Let Conditions 2.1 and 2.2 be satisfied. Let $N \in \mathbb{N}$ be finite and $u^\varepsilon = (u_1^\varepsilon, u_2^\varepsilon) \in \mathcal{A}$ such that almost surely
\[
\sup_{\varepsilon > 0} \mathbb{E} \int_0^1 |u^\varepsilon(s)|^2 \, ds < N.
\]
Let $A(x, y)$ and $B(x, y)$ be matrix-valued functions and $K, \theta \in (0, r)$, where $r$ is as defined in Condition 2.2, be constants such that each components $A_{ij}$ and $B_{ij}$ satisfy
\[
|A_{ij}(x, y)| \leq K(1 + |y|^\theta), \text{ and } |B_{ij}(x, y)| \leq K(1 + |y|^{2\theta}).
\]
Then for $\alpha \in \{1, 2\}$:

(i) For any $p \in (1, r/\theta]$, there exists a $C < \infty$ such that for fixed $\rho > 0$ and all $0 \leq t_1 < t_1 + \rho \leq 1$,
\[
\mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1 \atop |t_2 - t_1| < \rho} \left| \int_{t_1}^{t_2} A(X_s^\varepsilon, u_s^\varepsilon, Y_s^\varepsilon, u_s) u^\varepsilon(s) \, ds \right|^p \leq C|\rho|^{r/\theta - 1}.
\]
and
\[
\mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1 \atop |t_2 - t_1| < \rho} \left| \int_{t_1}^{t_2} B(X_s^\varepsilon, u_s^\varepsilon, Y_s^\varepsilon) \, ds \right|^p \leq C|\rho|^{r/\theta - 1}.
\]

(ii) For all $\zeta > 0$
\[
\lim_{\rho \downarrow 0} \lim_{\varepsilon \downarrow 0} \mathbb{E} \left[ \sup_{0 \leq t_1 < t_2 \leq 1 \atop |t_2 - t_1| < \rho} \left| \int_{t_1}^{t_2} A(X_s^\varepsilon, u_s^\varepsilon, Y_s^\varepsilon, u_s) u^\varepsilon(s) \, ds \right| > \zeta \right] = 0.
\]
and

\[
\lim_{\rho \downarrow 0} \lim_{\varepsilon \downarrow 0} \sup_{0 \leq t_1 < t_2 \leq 1} \mathbb{P} \left[ \sup_{\lvert t_2 - t_1 \rvert < \rho} \left| \int_{t_1}^{t_2} B(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \, ds \right| > \zeta \right] = 0.
\]

**Proof.** We shall only prove the statement for \( A(x, y) \), as the proof for \( B(x, y) \) is the same but simpler. Applying Hölder inequality with \( 1/m + 1/q = 1 \) and \( q = p > 1 \) gives

\[
\mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \left| \int_{t_1}^{t_2} A(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_\alpha(s) \, ds \right|^{2p} \leq \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \left( \int_{t_1}^{t_2} \left| A(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right|^2 \, ds \right)^p \left( \int_{t_1}^{t_2} \left| u_\alpha(s) \right|^2 \, ds \right)^p \leq N^p \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \left( \int_{t_1}^{t_2} \left| A(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right|^2 \, ds \right)^p \leq N^p \rho^{p/m} \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \left( \int_{t_1}^{t_2} \left| A(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right|^{2q} \, ds \right)^{p/q} \leq N^p \rho^{p/m} \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \left( \int_{t_1}^{t_2} \left| A(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right|^{2p} \, ds \right) \leq N^p \rho^{p/m} \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \int_{t_1}^{t_2} \left( 1 + \left| Y_s^{\varepsilon,u^\varepsilon} \right|^{2q} \right) \, ds \leq N^p \rho^{p-1} \mathbb{E} \int_0^1 \left( 1 + \left| Y_s^{\varepsilon,u^\varepsilon} \right|^{2r} \right) \, ds
\]

and the result follows by Lemma B.2 and by the choice of \( p \).

The second claim follows from the first statement and Markov’s inequality. \( \square \)

**Lemma B.4.** Assume Conditions \( 2.1 \) \( 2.2 \) \( 2.3 \) and \( 2.4 \) and define the function \( \chi(x, y) \) by \( B \) and the processes \( X^{\varepsilon,u^\varepsilon} \) and \( Y^{\varepsilon,u^\varepsilon} \) by \( A \). Let \( N < \infty \) such that almost surely

\[
\sup_{\varepsilon > 0} \int_0^1 \left| u_\varepsilon(s) \right|^2 \, ds < N.
\]

We have
(i)

\[
\mathbb{E} \sup_{t \in [0, 1]} \left| \int_0^t \frac{\varepsilon b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} \, ds 
- \int_0^t \left( \nabla y \chi(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \tau_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) + \tau_2(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_2^\varepsilon(s) \right] \, ds \right|^2
\]

\[< (\delta C_1)^2 + \left( \frac{1}{h(\varepsilon)} C_2 \right)^2 + o \left( \delta^2 + \frac{1}{h^2(\varepsilon)} \right)\]

(ii) For every \( \zeta > 0 \),

\[
\lim_{\rho \downarrow 0} \lim_{\varepsilon \downarrow 0} \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \sup_{|t_2 - t_1| < \rho} \left| \int_{t_1}^{t_2} \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \left( \frac{\varepsilon b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} \right) \, ds \right| > \zeta \right] = 0.
\]

(iii) There exists an \( M > 0 \) such that for all sufficiently small \( \varepsilon \),

\[
\mathbb{E} \sup_{t \in [0, 1]} \left| \int_0^t \frac{\varepsilon b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} \, ds \right|^2 < M.
\]
Proof. (i). Applying the Itô formula to $\chi(X_t^{\varepsilon,u^\varepsilon}, Y_t^{\varepsilon,u^\varepsilon})$ and rearranging gives

\[
\frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left[ \frac{\varepsilon}{\delta} b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right] ds
\]

\[
= \frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} \left( \chi(X_t^{\varepsilon,u^\varepsilon}, Y_t^{\varepsilon,u^\varepsilon}) - \chi(x_0, y_0) \right)
\]

\[
+ \frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_x \chi(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \frac{\varepsilon}{\delta} b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right] ds
\]

\[
+ \int_0^t \left( \nabla_y \chi(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \tau_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) + \tau_2(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_2^\varepsilon(s) \right] ds
\]

Using Lemmas 14.2 and 14.3 and Doob’s martingale inequality, along with the facts that the integrands that appear in the previous display grow no more than polynomially in $|y|$ (Condition 2.3 is being used here), we have

\[
\mathbb{E} \sup_{t \in [0,1]} \left| \int_0^t \frac{\varepsilon}{\sqrt{\varepsilon h(\varepsilon)}} \left[ \frac{\varepsilon}{\delta} b(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) - \lambda_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right] ds \right|^2
\]

\[
< (\delta C_1)^2 + \left( \frac{1}{h(\varepsilon)} C_2 \right)^2 + o \left( \delta^2 + \frac{1}{h^2(\varepsilon)} \right)
\]

where the constants $C_1, C_2 < \infty$ do not depend on $\varepsilon$.

(ii). Separate the integral in (27) as in (26), and then most terms go to zero in probability as $\varepsilon$ goes to zero. The only exception is the term

\[
\int_{t_1}^{t_2} \left( \nabla_y \chi(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \tau_1(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) + \tau_2(X_s^{\varepsilon,u^\varepsilon}, Y_s^{\varepsilon,u^\varepsilon}) u_2^\varepsilon(s) \right] ds.
\]

Condition 2.3 and Theorem A.1 imply that Lemma 14.3 can be applied, concluding the proof of the statement.
(iii). Rewrite (28) as in (29) and use the triangle inequality. Take expectations, and for fixed $\varepsilon$, all terms are bounded by Lemmas $B.1$, $B.2$, and $B.3$ and Doob’s martingale inequality. 

Lemma B.5. Assume Conditions $2.1$, $2.2$, $2.3$ and $2.4$ and define the processes $X_{\varepsilon,u}$ and $Y_{\varepsilon,u}$ by (13). Let $N < \infty$ such that almost surely

$$\sup_{\varepsilon > 0} \int_0^1 |u^\varepsilon(s)|^2 \, ds < N.$$ 

Also, define the function $\Phi_1(x,y)$ by (5). Then

(i) 

$$\mathbb{E} \sup_{t \in [0,1]} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_1(X^\varepsilon,u^\varepsilon, Y^\varepsilon,u^\varepsilon) - \bar{\lambda}_1(X^\varepsilon,u^\varepsilon) \right) \, ds \right|^2 \leq \left( \frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} C \right)^2 + o \left( \frac{\delta^2}{\varepsilon h^2(\varepsilon)} \right)$$

where the constant $C < \infty$ does not depend on the choice of $\varepsilon$.

(ii) For every $\zeta > 0$,

$$\lim_{\rho \downarrow 0} \limsup_{\varepsilon \downarrow 0} \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_{t_1}^{t_2} \left( \lambda_1(X^\varepsilon,u^\varepsilon, Y^\varepsilon,u^\varepsilon) - \bar{\lambda}_1(X^\varepsilon,u^\varepsilon) \right) \, ds \right| > \zeta \right] = 0.$$

(iii) There exists an $M > 0$ such that for all sufficiently small $\varepsilon$,

$$\mathbb{E} \sup_{t \in [0,1]} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_1(X^\varepsilon,u^\varepsilon, Y^\varepsilon,u^\varepsilon) - \bar{\lambda}_1(X^\varepsilon,u^\varepsilon) \right) \, ds \right|^2 < M.$$ 

Proof. (i). Note that

$$\int_Y \left( \lambda_1(x,y) - \bar{\lambda}_1(x) \right) \mu_\varepsilon(dy) = 0$$

for fixed $x$ by the definition of $\bar{\lambda}_1$ in (4). Then by Theorem $A.1$ (5) has a unique smooth solution for every $x$ in the space of functions with at most polynomial growth in $y$. 


Apply the Itô formula to $\Phi_1(X_t^{\varepsilon,u^\varepsilon},Y_t^{\varepsilon,u^\varepsilon})$ and rearrange to obtain

$$
\frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) - \tilde{\lambda}_1(X_s^{\varepsilon,u^\varepsilon}) \right) ds = - \frac{\delta^2/\varepsilon}{\sqrt{\varepsilon h(\varepsilon)}} \left( \Phi_1(X_t^{\varepsilon,u^\varepsilon},Y_t^{\varepsilon,u^\varepsilon}) - \Phi_1(x_0,y_0) \right)
$$

$$
+ \frac{\delta^2/\varepsilon}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_x \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \left( \frac{\varepsilon}{6} b(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) + c(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) ds
$$

$$
+ \frac{\delta^2/\varepsilon}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \varepsilon \sigma \sigma^T (X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) : \nabla_x \nabla_x \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) ds
$$

$$
+ \frac{\delta}{\varepsilon} \int_0^t \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \sigma(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) ds
$$

$$
+ \frac{\delta}{\varepsilon} \int_0^t \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) g(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) ds
$$

$$
+ \frac{\delta}{\varepsilon} \int_0^t \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \tau_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) + \tau_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) u_2^\varepsilon(s) \right] ds
$$

$$
+ \frac{\delta^2}{\varepsilon h(\varepsilon)} \int_0^t \left( \nabla_x \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \sigma(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) dW_s
$$

$$
+ \frac{\delta}{\varepsilon h(\varepsilon)} \int_0^t \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \tau_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) dW_s
$$

$$
+ \frac{\delta}{\varepsilon h(\varepsilon)} \int_0^t \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \tau_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) dB_s.
$$

Due to Conditions 2.1 and 2.2, $\lambda_1(x,y) - \tilde{\lambda}_1(x)$ satisfies the condition of Theorem A.1. Notice that a polynomial bound on $|y|$ of $\nabla_x \nabla_x \nabla_y \chi$ is needed. However, due to Condition 2.1 this follows by Theorems 1 and 2 in [14].

Using Theorem A.1, Lemmas B.2 and B.3 and Doob’s martingale inequality, along with the facts that the integrands that appear in the previous display grow no more than polynomially in $|y|^r$ (Condition 2.3 is being used here), we have

$$
\mathbb{E} \sup_{t \in [0,1]} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) - \tilde{\lambda}_1(X_s^{\varepsilon,u^\varepsilon}) \right) ds \right|^2 \leq \left( \frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} C \right)^2 + o \left( \frac{\delta^2}{\varepsilon h^2(\varepsilon)} \right)
$$

where $C$ does not depend on $\varepsilon$.

(ii). Again using (30), most terms go to zero in probability as $\varepsilon$ goes to zero. The possible exception is the term

$$
\frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} \int_{t_1}^{t_2} \left( \nabla_y \Phi_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) g(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) ds.
$$
Condition 2.3 and Theorem A.1 imply that Lemma B.3 can be applied, which together with the constraint \( \frac{h / \varepsilon}{\sqrt{\varepsilon h (\varepsilon)}} \to j_1 < \infty \) conclude the proof of the statement.

(iii). Again, use (30) and use the triangle inequality. For fixed \( \varepsilon \), by Lemmas B.1, B.2, and B.3 all terms are bounded in \( L^2([0,1] \times \mathbb{P}) \). \( \square \)

Lemma B.6. Assume Conditions 2.1, 2.2, 2.3 and 2.4 and define the processes \( X^{\varepsilon,u \varepsilon} \) and \( Y^{\varepsilon,u \varepsilon} \) by (13). Let \( N < \infty \) such that almost surely

\[
\sup_{\varepsilon > 0} \int_0^1 |u^{\varepsilon}(s)|^2 \, ds < N.
\]

Also, define \( \eta^{\varepsilon,u \varepsilon} \) by (12). Then

(i) \( \mathbb{E} \sup_{t \in [0,1]} |\eta^{\varepsilon,u \varepsilon}_t|^2 \leq K \exp(L_\lambda^2) \)

where \( L_\lambda \) is the Lipschitz constant for \( \bar{\lambda}_1 \) and the constant \( K \) does not depend on \( \varepsilon \).

(ii) For every \( \zeta > 0 \),

\[
\lim_{\rho \downarrow 0} \lim_{\varepsilon \downarrow 0} \sup_{\varepsilon > 0} \mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \left| \frac{1}{\sqrt{\varepsilon h (\varepsilon)}} \int_{t_1}^{t_2} (\bar{\lambda}_1(X^{\varepsilon,u \varepsilon}_s) - \bar{\lambda}_1(X_s)) \, ds \right| > \zeta \right] = 0.
\]

Proof. (i). \( \bar{\lambda}_1(x) \) is Lipschitz with Lipschitz constant \( L_\lambda \) by the fact that its first derivative is bounded. Write (12) as

\[
X^{\varepsilon,u \varepsilon}_t = \bar{X}_t + \sqrt{\varepsilon h (\varepsilon)} \eta^{\varepsilon,u \varepsilon}_t
\]

and then

\[
|\bar{\lambda}_1(X^{\varepsilon,u \varepsilon}_t) - \bar{\lambda}_1(\bar{X}_t)| \leq L_\lambda |\sqrt{\varepsilon h (\varepsilon)} \eta^{\varepsilon,u \varepsilon}_t|.
\]

Therefore we have

\[
(31) \sup_{0 \leq t_1 \leq 1} \left| \frac{1}{\sqrt{\varepsilon h (\varepsilon)}} \int_{0}^{t_1} (\bar{\lambda}_1(X^{\varepsilon,u \varepsilon}_s) - \bar{\lambda}_1(X_s)) \, ds \right|^2 \leq L_\lambda^2 \int_0^1 \left| \eta^{\varepsilon,u \varepsilon}_s \right|^2 \, ds.
\]
Using the decomposition in (14) and (21), we have up to some multiplicative constant $C < \infty$
\[
|\eta_t^{\varepsilon,u^\varepsilon}|^2 \leq \left| \int_0^t \frac{2b(X_{s}^{\varepsilon,u^\varepsilon},Y_{s}^{\varepsilon,u^\varepsilon}) + c(X_{s}^{\varepsilon,u^\varepsilon},Y_{s}^{\varepsilon,u^\varepsilon}) - \lambda_1(X_{s}^{\varepsilon,u^\varepsilon},Y_{s}^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} \, ds \right|^2 + \left| \int_0^t \frac{\lambda_1(X_{s}^{\varepsilon,u^\varepsilon},Y_{s}^{\varepsilon,u^\varepsilon}) - \tilde{\lambda}_1(X_{s}^{\varepsilon,u^\varepsilon})}{\sqrt{\varepsilon h(\varepsilon)}} \, ds \right|^2 + \left| \int_0^t \frac{\tilde{\lambda}_1(X_{s}^{\varepsilon,u^\varepsilon}) - \tilde{\lambda}_1(X_{s})}{\sqrt{\varepsilon h(\varepsilon)}} \, ds \right|^2 + \left| \int_0^t \frac{1}{h(\varepsilon)}\sigma(X_{s}^{\varepsilon,u^\varepsilon},Y_{s}^{\varepsilon,u^\varepsilon}) \, dW_s \right|^2.
\]

Take supremum in $t \in [0,1]$ and then expectations of both sides. For sufficiently small $\varepsilon$, the first term is bounded by Lemma B.4. The second term is bounded by Lemma B.5. The third term is bounded by Lemma B.3. Finally, the expectation of the fifth term is bounded due to Doob’s martingale inequality and the bound on $\sigma$ together with Lemma B.2. Combining these, we get
\[
\mathbb{E} \sup_{t \in [0,1]} |\eta_t^{\varepsilon,u^\varepsilon}|^2 \leq K + L_\lambda^2 \int_0^1 \mathbb{E} \sup_{s \in [0,t]} |\eta_s^{\varepsilon,u^\varepsilon}|^2 \, dt
\]
where $K$ is the sum of the bounds on the expectations of terms one, two, four, and five. Then by Gronwall’s lemma, we have the required statement,
\[
(32) \quad \mathbb{E} \sup_{t \in [0,1]} |\eta_t^{\varepsilon,u^\varepsilon}|^2 \leq K \exp(L_\lambda^2)
\]
for all sufficiently small $\varepsilon > 0$.

(ii) From (31) and the Markov inequality,
\[
\mathbb{P} \left[ \sup_{0 \leq t_1 < t_2 \leq 1} \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_{t_1}^{t_2} (\tilde{\lambda}_1(X_{s}^{\varepsilon,u^\varepsilon}) - \tilde{\lambda}_1(X_{s})) \, ds > \zeta^2 \right] \leq \frac{L_\lambda^2}{\zeta^2} \mathbb{E} \sup_{0 \leq t_1 < t_2 \leq 1} \int_{t_1}^{t_2} |\eta_{s}^{\varepsilon,u^\varepsilon}|^2 \, ds.
\]
Since $\mathbb{E} \sup_{t \in [0,1]} |\eta_t^{\varepsilon,u^\varepsilon}|^2$ is uniformly bounded by (32) for $\varepsilon$ small enough, this probability goes to zero as $|t_2 - t_1|$ goes to zero, completing the proof.

\[\square\]

**Appendix C. Lemmas for Regime 2**

Notice that Lemmas B.4, B.5 and B.6 are also valid for Regime 2. Statements and proofs for the lemmas corresponding to Lemmas B.4, B.5 and B.6 are similar to those in Regime 1, by considering $\lambda_2$ in place of $\lambda_1$. The
only difference is in the proof of the statement that corresponds to Lemma B.5(i), which we now state and prove.

**Lemma C.1.** Assume Conditions 2.1, 2.2, and 2.3 and define the processes $X^{\varepsilon,u^\varepsilon}$ and $Y^{\varepsilon,u^\varepsilon}$ by (13). Let $N < \infty$ such that almost surely

$$\sup_{\varepsilon > 0} \int_0^1 |u^\varepsilon(s)|^2 ds < N.$$ 

Also, define the function $\Phi_2(x,y)$ by (5) with $j_2 < \infty$. Then

$$\mathbb{E} \sup_{t \in [0,1]} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) - \bar{\lambda}_2(X_s^{\varepsilon,u^\varepsilon}) \right) ds 
- \int_0^t \left( \nabla_y \Phi_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) \left[ \tau_1(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) u_1^\varepsilon(s) + \tau_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) u_2^\varepsilon(s) \right] ds 
- \frac{\varepsilon/\delta - \gamma}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_y \Phi_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) \right) f(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) ds 
- \frac{1}{2} \frac{\varepsilon/\delta - \gamma}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \tau_1 \tau_1^T + \tau_2 \tau_2^T \right) (X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) : \nabla_y \nabla_y \Phi_2(X_s^{\varepsilon,u^\varepsilon},Y_s^{\varepsilon,u^\varepsilon}) ds \right|^2 
\leq (\delta C_1)^2 + \left( \frac{1}{h(\varepsilon)} C_2 \right)^2 + o \left( \delta^2 + \frac{1}{h^2(\varepsilon)} \right)$$

where the constants $C_1$ and $C_2$ do not depend on the choice of $\varepsilon$.

**Proof.** Note that in Regime 2

$$\int_y \left( \lambda_2(x,y) - \bar{\lambda}_2(x) \right) \mu_{2,x}(dy) = 0$$

for fixed $x$ by the definition of $\bar{\lambda}_2$ in (4). Then (3) has a unique, smooth solution for every $x$ that is bounded in $x$ and grows at most polynomially in $|y|$ as in Theorem A.1.
Apply the Itô formula to $\Phi_2(X_{t}^{\epsilon,u^\epsilon}, Y_{t}^{\epsilon,u^\epsilon})$ and rearrange to show

$$
\frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) - \tilde{\lambda}_2(X_{s}^{\epsilon,u^\epsilon}) \right) ds = -\frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} \left( \Phi_2(X_{t}^{\epsilon,u^\epsilon}, Y_{t}^{\epsilon,u^\epsilon}) - \Phi_2(x_0, y_0) \right)
$$

$$
+ \frac{\delta}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_x \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \left( \frac{\varepsilon}{\delta} b(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) + c(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) ds
$$

$$
+ \frac{\delta \varepsilon}{h(\varepsilon) \sqrt{\varepsilon h(\varepsilon)}} \int_0^t \sigma^T(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) : \nabla_x \nabla_x \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) ds
$$

$$
+ \delta \int_0^t \left( \nabla_x \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \sigma(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) u_1(s) ds
$$

$$
+ \int_0^t \left( \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \left[ \tau_1(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) u_1^\epsilon(s) + \tau_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) u_2^\epsilon(s) \right] ds
$$

$$
+ \frac{\delta}{h(\varepsilon)} \int_0^t \left( \nabla_x \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \tau_1(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) dW_s
$$

$$
+ \frac{1}{h(\varepsilon)} \int_0^t \left( \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \tau_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) dB_s
$$

$$
+ \frac{\varepsilon}{\delta} - \gamma \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) f(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) ds
$$

$$
+ \frac{1}{2} \frac{\varepsilon}{h(\varepsilon)} \int_0^t \left( \tau_1 T_1 + \tau_2 T_2 \right) (X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) : \nabla_y \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) ds.
$$

Using Theorem A.1, Lemmas B.2 and B.3 and Doob's martingale inequality, along with the facts that the integrands that appear in the previous display grow no more than polynomially in $|y|^2$ (Condition 2.3 is being used here), we have

$$
\mathbb{E} \sup_{t \in [0,1]} \left| \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \lambda_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) - \tilde{\lambda}_2(X_{s}^{\epsilon,u^\epsilon}) \right) ds \right|
$$

$$
- \int_0^t \left( \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) \left[ \tau_1(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) u_1^\epsilon(s) + \tau_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) u_2^\epsilon(s) \right] ds
$$

$$
- \frac{\varepsilon}{\delta} - \gamma \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) \right) f(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) ds
$$

$$
- \frac{1}{2} \frac{\varepsilon}{\delta} - \gamma \frac{1}{\sqrt{\varepsilon h(\varepsilon)}} \int_0^t \left( \tau_1 T_1 + \tau_2 T_2 \right) (X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) : \nabla_y \nabla_y \Phi_2(X_{s}^{\epsilon,u^\epsilon}, Y_{s}^{\epsilon,u^\epsilon}) ds
$$

$$
\leq (\delta C_1)^2 + \left( \frac{1}{h(\varepsilon)} C_2 \right)^2 + o \left( \delta^2 + \frac{1}{h^2(\varepsilon)} \right),
$$

where $C_1$ and $C_2$ do not depend on $\varepsilon$. \qed
REFERENCES

[1] D. Baier and M.I. Freidlin, Theorems on large deviations and stability under random perturbations, *Soviet Math. Dokl.*, Vol. 18, No. 4 (1977), pp. 905–909.

[2] M. Boué and P. Dupuis, A variational representation for certain functionals of Brownian motion, *The Annals of Probability*, Vol. 26, No. 4 (1998), pp. 1641–1659.

[3] A. Budhiraja and P. Dupuis, A variational representation for positive functionals of infinite dimensional Brownian motion, *Probability and Mathematical Statistics*, Vol. 20, Fasc. 1 (2000), pp. 39–61.

[4] P. Chigansky and R. Lipster, Moderate deviations for a diffusion-type process in a random environment, *Theory of Probability and Applications*, Vol. 54, No. 1 (2010), pp. 29–50.

[5] P. Dupuis and R.S. Ellis, *A weak convergence approach to the theory of large deviations*, John Wiley & Sons, New York, 1997.

[6] P. Dupuis and D. Johnson, Moderate deviations for recursive stochastic algorithms, *Stochastic Systems*, Vol. 5, No. 1 (2015), pp. 87–119.

[7] P. Dupuis and K. Spiliopoulos, Large deviations for multiscale problems via weak convergence methods, *Stochastic Processes and their Applications*, Vol. 122, (2012), pp. 1947–1987.

[8] P. Dupuis, K. Spiliopoulos, and H. Wang, Importance sampling for multiscale diffusions, *SIAM Journal on Multiscale Modeling and Simulation*, Vol. 12, No. 1 (2012), pp. 1–27.

[9] M.I. Freidlin, The averaging principle and theorems on large deviations, *Russian Mathematical Surveys* Vol. 33, No. 5 (1978), pp. 117–176.

[10] M.I. Freidlin and R. Sowers, A comparison of homogenization and large deviations, with applications to wavefront propagation, *Stochastic Process and Their Applications*, Vol. 82, Issue 1, (1999), pp. 23–52.

[11] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems*. Translated by J. Szücs. Springer-Verlag, New York, 1984.

[12] A. Guillin, Averaging principle of SDE with small diffusion: moderate deviations, *The Annals of Probability*, Vol. 31, No. 1 (2003), pp. 413–443.

[13] A. Guillin and R. Liptser, MDP for integral functionals of fast and slow processes with averaging, *Stochastic Processes and their Applications*, Vol. 115, No. 7 (2005), pp. 1187–1207.

[14] E. Pardoux and A.Yu. Veretennikov, On the Poisson equation and diffusion approximation I, *The Annals of Probability*, Vol. 29, No. 3 (2001), pp. 1061–1085.

[15] E. Pardoux and A.Yu. Veretennikov, On Poisson equation and diffusion approximation 2, *The Annals of Probability*, Vol. 31, No. 3 (2003), pp. 1166–1192.

[16] K. Spiliopoulos, Fluctuation analysis and short time asymptotics for multiple scales diffusion processes, *Stochastics and Dynamics*, Vol. 14, No. 3 (2014), pp. 1350026.

[17] K. Spiliopoulos, Large deviations and importance sampling for systems of slow-fast motion, *Applied Mathematics and Optimization*, Vol. 67 (2013), pp. 123–161.

[18] K. Spiliopoulos, Quenched large deviations for multiscale diffusion processes in random environments, *Electronic Journal of Probability*, Vol. 20 (2015), pp. 1–29.

[19] A.Yu. Veretennikov, On strong solutions and explicit formulas for solutions of stochastic integral equations, *Math. USSR-Sb* Vol. 39 (1981), pp. 387–403.

[20] A. Yu. Veretennikov, On large deviations in the averaging principle for SDEs with a “full dependence”, correction, [arXiv:math/0502098v1 [math.PR]] (2005). Initial article in *Annals of Probability*, Vol. 27, No. 1, (1999), pp. 284–296.

[21] R. Zwanzig, Diffusion in a rough potential, *Proc. Natl. Acad. Sci. USA*, 85 (1988), pp. 2029–2030.