LARGE DEVIATIONS FOR INTERACTING MULTISCALE PARTICLE SYSTEMS

Z.W. BEZEMEK, AND K. SPILIOPOULOS

ABSTRACT. We consider a collection of weakly interacting diffusion processes moving in a two-scale locally periodic environment. We study the large deviations principle of the empirical distribution of the particles’ positions in the combined limit as the number of particles grow to infinity and the time-scale separation parameter goes to zero simultaneously. We make use of weak convergence methods providing a convenient representation for the large deviations rate function, which allow us to characterize the effective controlled mean field dynamics. In addition, we obtain in certain special cases equivalent representations for the large deviations rate function.

1. INTRODUCTION

The goal of this article is to obtain the large deviations principle (LDP) for interacting particle systems of diffusion type in multiscale environments. We use methods from weak convergence and stochastic control, ultimately making connections with mean field stochastic control problems [12].

In particular, we consider the interacting particle system

\begin{equation}
\begin{aligned}
\frac{dX_{t}^{i,N}}{dt} &= \frac{1}{\delta} f(X_{t}^{i,N}, X_{t}^{i,N}/\delta, \mu_{t}^{N}) + b(X_{t}^{i,N}, X_{t}^{i,N}/\delta, \mu_{t}^{N}) \frac{dW_{t}^{i}}{d}, \\
X_{0}^{i,N} &= \omega^{i,N}
\end{aligned}
\end{equation}

where \( t \in [0, 1], W_{t}^{i}, i = 1, ..., N \) are \( m \)-dimensional independent \( \mathcal{F}_{t} \)-Brownian motions,

\[ \mu_{t}^{N}(\omega) := \frac{1}{N} \sum_{i=1}^{N} \delta X_{t}^{i,N}(\omega), \]

\( X_{t}^{i,N}(\omega), f(x, y, \mu), b(x, y, \mu) \in \mathbb{R}^{d}, \sigma(x, y, \mu) \in \mathbb{R}^{d \times m} \) and all coefficients are \( 1 \)-periodic in the second coordinate. Suppose also that \( \delta > 0, N \in \mathbb{N} \) and \( \delta(N) \to 0 \) as \( N \to \infty \).

Our goal is to obtain the large deviations principle for the measure-valued process \( \{\mu_{t}^{N}, t \in [0, 1]\} \) in the combined limit \( N \to \infty \) and \( \delta \downarrow 0 \). Here, \( \delta \) is the time scale separation parameter. One can regard \( X_{t}^{i,N} \) as the slow \( i^{th} \) component and \( Y_{t}^{i,N} = X_{t}^{i,N}/\delta \) as the fast \( i^{th} \) component.

Systems of interacting diffusions arise in many areas of science, finance and engineering, see for example [6, 30, 32, 33, 38, 44, 46] to name just a few. On the other hand, diffusions in multiscale environments are also common in many applications ranging from chemical physics to finance and climate modeling, see for example [1, 7, 23–25, 37, 45, 64] for a representative, but by no means complete, list. Our goal in this paper is to study the combined effect of weak mean field interactions in a fast oscillating multiscale environment from the point of view of large deviations for the empirical measure of particles.

In the case \( \delta = 1 \), i.e. in the absence of multiple scales, the limiting problem of \( N \to \infty \) has been very well studied in the literature. Typical behavior, fluctuations as well as large deviations have been obtained, see for example [1] for related classical works. Analogously, if \( N = 1 \), i.e., in the single particle case, the limiting behavior as \( \delta \downarrow 0 \) has also been extensively studied in the literature under various modeling assumptions, see for example [2, 9, 20, 26, 31, 45, 50, 51, 56, 60] and the references therein. In this paper, we study the combined limit as \( N \to \infty \) and \( \delta(N) \downarrow 0 \). The main result of the paper is Theorem 3.4 that gives the large deviations principle of the empirical distribution of the particles in the combined limit \( N \to \infty \) and \( \delta \downarrow 0 \). As a byproduct we also obtain in Theorem 3.7 the typical behavior, i.e. the law of large numbers.

We use weak convergence methods of [21] which leads to the study of related mean field stochastic control.
problems \cite{29,42}. In addition, in Section 5 we connect, in a simplified regime, the variational form of the action functional that we obtain in Theorem 3.4 with the expected “dual” form based on the classical work of Dawson and Gärtner \cite{16} in the \( \delta = 1 \) case and \cite{20} in the \( N = 1 \) case. As pointed out in \cite{29}, while the limit problems here are not mean-field games, but rather optimal control problems of McKean–Vlasov-type, the modern advances in control theory for mean-field models and McKean-Vlasov equations resulting from mean-field games’ current popularity may provide a rigorous means of proving the formal results provided in Section 4. This analysis will appear elsewhere.

As an example, we consider in Section 5 a system of interacting processes in dimension one in a multiscale confining potential, i.e. a Curie-Weiss type of interaction. This example in motivated by the work of Dawson and Gärtner \cite{16}, in the \( \delta = 1 \) case of \cite{16} and the LDP in the \( N = 1 \) case. As pointed out in \cite{29}, while the large deviations principle satisfies the required assumptions for an LDP to hold, derive the large deviations principles and discuss equivalent formulations. We provide an alternative variational form of the rate function for this particular system in Theorem 5.2.

To our knowledge, this is the first large deviations result for the combined \( \delta \downarrow 0 \) and \( N \to \infty \) limit. Some similar results include the proof of an averaging principal for slow-fast McKean-Vlasov SDEs found in \cite{30}, in which the authors establish an averaging principal for the empirical mean a system of mean field multiscale diffusions at the level of large deviations. What is meant by this is that they study the known large deviations rate functions for the \( N \to \infty \) limit of the empirical measures with fixed \( \delta, J^k \), and are able to prove \( \Gamma \)-convergence of the sequence \( \{J^k\}_{k\geq 0} \) to a functional \( J \) as \( \delta \downarrow 0 \). Lastly, in \cite{17}, a result similar to Theorem 3.7 appears (only typical behavior, not LDP). A key difference between the regime of \cite{17} and the regime of our paper is that rather than depending on the slow process \( X_{t}^{i,N} \), the fast process \( X_{t}^{i,N}/\delta \), and the empirical measure \( \mu^N_t \), their coefficients depend on the fast process \( X_{t}^{i,N}/\delta \) and the “fast empirical measure” \( \mu^N_t \). As a result, the invariant measure \( \pi \) (see Equation 3) depends not on the parameter \( \mu = \mathcal{L}(X) \) in the limit, as in our regime, but implicitly on itself as \( \mu = \pi \). Consequently, in \cite{17}, multiple steady states can exist, potentially affecting the way in which the limits \( \delta \to 0 \) and \( N \to \infty \) interact. We discuss this further in our conclusion section 11.

The rest of the paper is organized as follows. In Section 2 we lay out notation and main assumptions in regards to the model \cite{11}. In addition, we introduce the corresponding controlled particle system and controlled McKean-Vlasov process which will be crucial components of the large deviations analysis. In Section 3 we present our main result on large deviations for the measure-valued process \( \{\mu^N_t, t \in [0,1]\} \) in the combined limit \( N \to \infty \) and \( \delta \downarrow 0 \). Section 4 connects the obtained Laplace principle with other classical works in the literature, i.e. the LDP in the \( \delta = 1 \) case of \cite{16} and the LDP in the \( N = 1 \) case of \cite{20}. Section 5 includes an example with a bistable confining potential motivated by the classical work of Dawson in \cite{15} and an alternative variational form of the rate function. In Section 6 we discuss the limiting behavior of the controlled particle system, proving tightness and identifying the limit. In Sections 7 and 8 we prove the Laplace’s principle (which is equivalent to the large deviations principle) lower and upper bounds respectively. Compactness of the level sets of the rate function is proven in Section 9. In Section 11 we discuss how the assumptions on the coefficients can be relaxed. In Appendices A, C and D we discuss technical preliminary results that are used in various places of the paper. For purposes of self containment and for the reader’s convenience, Appendix B reviews the necessary material from Lions differentiation. Lastly, an equivalent variational representation for the rate function offered in Theorem 3.4 is discussed in Appendix E. Section 11 has our conclusions and directions for future work.

2. Notation, Assumptions, and the Controlled McKean-Vlasov Process

Consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with filtration \(\mathcal{F}_t\) that satisfies the usual conditions. For \( S \) a Polish space, we will use \( C([0,1];S) \) to denote the space of continuous functions from \([0,1]\) to \( S \), equipped with the topology of uniform convergence and the sup norm. We will use \( D_S[0,1] \) to denote \( B : [0,1] \to S \) such that \( B \) is right-continuous and has left-hand limits, equipped with the Skorohod topology. A useful fact is that both \( D_{b,S}[0,1] \) and \( C([0,1];\mathbb{R}^d) \) with the previously described topologies are Polish spaces (see \cite{21} Theorem A.6.5). We will use \( C_c(S) \) to denote the space of continuous, bounded functions \( B : S \to \mathbb{R} \), and let \( ||B||_{\infty} := \sup_{x \in S} |B(x)| \). We use \( C_b^k(S) \) to denote the space of continuous, bounded functions \( B : S \to \mathbb{R} \) with \( k \) continuous, bounded derivatives. We use \( C_c^k(S) \) to denote the space of continuous, infinitely differentiable functions \( B : S \to \mathbb{R} \) with compact support. \( L^2(S, \mu; \mathbb{R}^k) \), where \( \mu \) is a measure on \( S \) will denote the class
of functions $B : S \to \mathbb{R}^k$ such that $\|B\|_{L^2(\mu)} := \left( \int_S |B(x)|^2 \mu(dx) \right)^{1/2} < \infty$. We may omit the codomain in this notation when convenient. $\mathcal{P}(S)$ will denote the space of probability measures on the Borel $\sigma$-field $\mathcal{B}(S)$, where open sets are induced by the metric on $S$. $\mathcal{P}(S)$ is given the topology of weak convergence and the Prokhorov metric, and is itself a Polish space ([12] p.360). Given a random variable $\eta$ the Prokhorov metric, and is itself a Polish space ([22] Theorem 3.1.7) $\mathcal{P}_2(E) \subset \mathcal{P}(E)$ will denote the set of square integrable measures on $E$. It is given the $L^2$-Wasserstein distance (see Definition [B.1] as its metric and is also a Polish space ([12] p.360). Given a random variable $\eta$, $\mathcal{L}(\eta)$ will denote the distribution of $\eta$.

Assume the following:

(A1) For $g = f, b, \sigma, g, \nabla_x g, \nabla_y g, \nabla_x \nabla_y g, \text{and } \nabla_x \nabla_y g$ exist, are uniformly bounded, and are continuous in $\mathbb{R}^d \times \mathbb{T}^d \times \mathcal{P}(\mathbb{R}^d)$.

(A2) For $A = \sigma \sigma^T$ there exists $\gamma > 0$ such that uniformly in $x \in \mathbb{R}^d, y \in \mathbb{T}^d, \mu \in \mathcal{P}_2^d$,

$$
\epsilon^T A(x, y, \mu) \epsilon \geq \gamma |\epsilon|^2, \forall \epsilon \in \mathbb{R}^d.
$$

(A3) For $g = f, b, \sigma$ and $x \in \mathbb{R}^d, y \in \mathbb{T}^d, g(x, y, \cdot), \nabla_x g(x, y, \cdot), \nabla_y g(x, y, \cdot),$ and $\nabla_y g(x, y, \cdot)$ are all Fully $C^2$ in the sense of Definition [B.2].

(A4) For $g = f, b$ or $\sigma, \partial_t g(x, y, \mu)(\cdot), \nabla_x \partial_t g(x, y, \mu)(\cdot), \nabla_y \partial_t g(x, y, \mu)(\cdot)$ are all bounded in $L^2(\mathbb{R}^d, \mu)$ uniformly in $x, y, \mu$.

(A5) Three exists $L \in (0, \infty)$ such that for $x_1, x_2 \in \mathbb{R}^d, y_1, y_2 \in \mathbb{T}^d, \mu_1, \mu_2 \in \mathcal{P}_2^d$,

$$
|g(x_1, y_1, \mu_1) - g(x_2, y_2, \mu_2)| \leq L \left( |x_1 - x_2| + |y_1 - y_2| + \mathbb{W}_2(\mu_1, \mu_2) \right),
$$

where $g = f, b, \sigma$ and $\mathbb{W}_2$ is the $L^2$-Wasserstein distance (see Appendix [B]).

(A6) For some $\nu_0 \in \mathcal{P}(\mathbb{R}^d), \frac{1}{N} \sum_{i=1}^N \delta_{x_i, N} \to \nu_0$ as $N \to \infty$ and $\sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=1}^N |x_i, N|^2 < \infty$.

(A7) All the first and second derivatives listed in (A1) and (A4) are Hölderian in $y$ uniformly in $x$ and $\mu$.

(A8) The centering condition:

$$
\int_{\mathbb{T}^d} f(x, y, \mu) \pi(dy|x, \mu) = 0, \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d)
$$

holds.

By Proposition [A.1] under Assumptions (A1)-(A8), this system has a unique strong solution for each $N \in \mathbb{N}$. It is also worth noting here that though the assumption (A8) is standard in the theory of averaging and homogenization for guaranteeing unique strong solutions to the Cell Problem [H] in this particular instance also has a significant effect on the identification of the limiting SDE given in Equation [12] (see [G1]).

As we shall see in Section 10, Assumptions (A1)-(A8) can be substantially relaxed in the expense of more technical estimates. We choose to present the main results with Assumptions (A1)-(A8) for the sake of readability.

An important object of study will be the operator $\mathcal{L}_{x, \mu}$, paramaterized by $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ which acts on $g \in C^2(\mathbb{T}^d)$ by

$$
\mathcal{L}_{x, \mu}^x g(y) := f(x, y, \mu) \cdot \nabla g(y) + \frac{1}{2} A(x, y, \mu) \cdot \nabla \nabla g(y).
$$

Related to this operator we consider the measure $\pi(\cdot|x, \mu) \in \mathcal{P}(\mathbb{T}^d)$, paramaterized by $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, whose density $\bar{\pi}(\cdot|x, \mu)$ satisfies the adjoint equation

$$
\left( \mathcal{L}_{x, \mu}^x \right)^* \bar{\pi}(y|x, \mu) = 0
$$

$$
\int_{\mathbb{T}^d} \bar{\pi}(y|x, \mu) dy = 1, \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d),
$$

and the function $\Phi : \mathbb{R}^d \times \mathbb{T}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$, $\Phi = (\Phi_1, ..., \Phi_d)$ solving

$$
\mathcal{L}_{x, \mu}^x \Phi_l(x, y, \mu) = -f_l(x, y, \mu)
$$

$$
\int_{\mathbb{T}^d} \Phi_l(x, y, \mu) \pi(dy|x, \mu) = 0,
$$

where both of these equations are given periodic boundary conditions. As we will see in Propositions [C.1] and [C.2], $\pi$ and $\Phi$ are uniquely defined and $\pi$ indeed admits a density $\bar{\pi}$. 
We wish to observe the behavior of the sequence of \( \mathcal{P}(C([0,1];\mathbb{R}^d)) \) random variables

\[
\mu^N(\omega) := \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i,N(\omega)}
\]

as \( N \to \infty \). Specifically, letting \( ev : C([0,1];\mathbb{R}^d) \to \mathbb{R}^d \) be the evaluation map at time \( t \), in Theorem 3.7 we see that, under assumptions (A1)-(A8), \( L(\mu^N) \to \delta_{\mu^*} \) in \( \mathcal{P}(C([0,1];\mathbb{R}^d)) \), where deterministic \( \mu^* \in \mathcal{P}(C([0,1];\mathbb{R}^d)) \) satisfies \( \mu^* \circ ev^{-1}(t) = L(X_t), t \in [0,1] \) for \( X \) solving the McKean-Vlasov SDE:

\[
\begin{align*}
\frac{1}{2} B(x,\mu) B(x,\mu)^\top &= \int_{\mathbb{R}^d} \left[ \nabla_y \phi(x,y,\mu) A(x,y,\mu) + f(x,y,\mu) \right] \pi(dy|x,\mu) \\
X_0 &\sim \nu_0
\end{align*}
\]

for \( W_t \) a \( d \)-dimensional \( \mathcal{F}_t^N \)-Brownian motion. Here we define

\[
A : \nabla_x \nabla_y \phi(x,y,\mu) := (A(x,y,\mu) : \nabla_x \nabla_y \phi_1(x,y,\mu), \ldots, A(x,y,\mu) : \nabla_x \nabla_y \phi_d(x,y,\mu))^\top.
\]

We seek to quantify the rate at which the convergence of the random measures given by Equation 5 to the law of the solution of Equation 6 occurs via deriving a large deviations principal for \( \{\mu^N\}_{N \in \mathbb{N}} \).

2.1. The Controlled Process. We start by constructing a controlled version of the system of mean-field SDEs which will then allow us to use the weak convergence approach to large deviations of [21].

For \( N \in \mathbb{N} \) let \( \mathcal{U}_N \) denote the space of \( \mathcal{F}_t \)-progressively measurable functions \( u : [0,1] \times \Omega \to \mathbb{R}^{N \times m} \) such that \( \mathbb{E}[\int_0^1 |u(t)|^2 dt] < \infty \), where \( \mathbb{E} \) denotes the expectation with respect to \( \mathbb{P} \) and \( |\cdot| \) the Euclidean norm. For \( u \in \mathcal{U}_N \), we write \( u = (u_1, \ldots, u_N) \) where \( u_i \in \mathbb{R}^m, i = 1, \ldots, N \).

Given \( u^N \in \mathcal{U}_N \), we consider the controlled system of SDEs

\[
\begin{align*}
d\bar{X}^i,N &= \left[ \frac{1}{\theta} f(\bar{X}^i,N, \bar{X}^i,N/\delta, \bar{\mu}^N) + b(\bar{X}^i,N, \bar{X}^i,N/\delta, \bar{\mu}^N) + \sigma(\bar{X}^i,N, \bar{X}^i,N/\delta, \bar{\mu}^N)u^N_i \right] dt \\
&\quad + \sigma(\bar{X}^i,N, \bar{X}^i,N/\delta, \bar{\mu}^N) dW^i_t \\
\bar{X}^0 = x^N
\end{align*}
\]

where \( \bar{\mu}^N(t) \) and \( \bar{\mu}^N \) are the empirical measures of \( X^i,N(t) \) and \( X^i,N \) respectively,

\[
\bar{\mu}^N(\omega) := \frac{1}{N} \sum_{i=1}^{N} \delta_{\bar{X}^i,N(\omega)}, \quad \bar{\mu}^N(\omega) := \frac{1}{N} \sum_{i=1}^{N} \delta_{\bar{X}^i,N(\omega)}.
\]

For notational convenience, we now introduce some spaces of interest. Let \( \mathcal{X} := C([0,1];\mathbb{R}^d), \mathcal{Y} := \mathcal{R}^1(\mathbb{R}^d), \mathcal{Z} = \mathcal{R}_+^1, \) and \( W = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \). Here

\[
\mathcal{R}_+^1 := \{ r : r \text{ is a positive Borel measure on } \mathbb{R}^m \times [0,\alpha], r(\mathbb{R}^m \times [0,t]) = t, \forall t \in [0,\alpha] \}
\]

and

\[
\int_{\mathbb{R}^m \times [0,\alpha]} |z| r(dz \times dt) < \infty.
\]

Note that we construct \( \mathcal{Y} \) and \( \mathcal{Z} \) this way to allow for extension of the results of this paper to bounded time intervals other than \([0,1]\). Also note that by Section 6.3 in [33] that \( \mathcal{Z} \) is a Polish space and that by Theorem A.3.3. in [21] that \( \mathcal{Y} \) is a Polish space.

Note that if \( u \in \mathcal{U}_N \) for any \( N \in \mathbb{N} \), then \( u \) induces a \( \mathcal{Z} \)-valued random variable \( r \) via

\[
r_\omega(D \times I) := \int_I \delta_{u(t,\omega)}(D) dt, \quad D \in \mathcal{B}(\mathbb{R}^m), I \in \mathcal{B}([0,1]), \omega \in \Omega.
\]
Since for $r \in \mathcal{Z}$, $t \mapsto r(B \times [0, t])$ for $B \in \mathcal{B}(\mathbb{R}^d)$ is absolutely continuous, there exists $r_t : [0, 1] \rightarrow \mathcal{P}(\mathbb{R}^m)$ such that $r(dzdt) = r_t(dz)dt$. Similarly, for $m \in \mathcal{Y}$, there exists $m_t : [0, 1] \rightarrow \mathcal{P}(\mathbb{T}^d)$ such that $m(dydt) = m_t(dy)dt$.

Consider the McKean-Vlasov SDE parameterized by $\nu \in C([0, 1]; \mathcal{P}(\mathbb{R}^d))$ given by:

\[
\begin{align*}
(10) \quad d\tilde{X}^\nu_t &= \int_{\mathbb{R}^d} \left\{ |\nabla_y \Phi(\tilde{X}^\nu_t, y, \nu(t)) + I| \left[ b(\tilde{X}^\nu_t, y, \nu(t)) + \sigma(\tilde{X}^\nu_t, y, \nu(t)) \right] z\nu_t(dz) \\
&+ \nabla_x \Phi(\tilde{X}^\nu_t, y, \nu(t)) f(\tilde{X}^\nu_t, y, \nu(t)) + A : \nabla_x \nabla_y \Phi(\tilde{X}^\nu_t, y, \nu(t)) \right\} m_t(dy)dt \\
&+ B(t, \tilde{X}^\nu_t, \nu(t))dW_t, \\
&= \frac{1}{2} B(t, x, \mu) B(t, x, \mu)^T = \int_{\mathbb{R}^d} \left[ \nabla_y \Phi(x, y, \mu) + \frac{1}{2} A(x, y, \mu) + f(x, y, \mu) \otimes \Phi(x, y, \mu) \right] m_t(dy)
\end{align*}
\]

for $\tilde{X}^\nu \in \mathcal{X}$, $m \in \mathcal{Y}$, and $\rho \in \mathcal{Z}$. $Q \in \mathcal{P}(\mathcal{W})$ corresponds to a weak solution of (10) if there exists a filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \{\tilde{\mathcal{F}}_t\})$ and an $m$-dimensional $\tilde{\mathcal{F}}_t$-Brownian motion $W$ such that $(\tilde{X}^\nu, m, \rho)$ is an $\tilde{\mathcal{F}}_t$-adapted $\mathcal{W}$-valued random variable that has distribution $Q$ under $\tilde{\mathbb{P}}$. Note that $\tilde{X}^\nu$, $m$, and $\rho$ are each random processes.

A useful remark making the point that the measures $\rho(\omega)$, $m(\omega)$ and the process $\tilde{X}^\nu(\omega)$ all dependent on each other follows.

Remark 2.1. While at a given $\omega$, the measure $\rho(\omega)$ appears separately from $m(\omega)$ and $\tilde{X}^\nu(\omega)$ in Equation (10) apriori $\tilde{X}^\nu$, $m$, and $\rho$ are not independent random variables (nor should we expect them to be).

We are interested in particular in weak solutions $Q$ to $\tilde{X}^{\nu_Q}$, where $\nu_Q(t) : [0, 1] \rightarrow \mathcal{P}(\mathbb{R}^d)$ is the Borel measurable mapping defined by

\[
\begin{align*}
(11) \quad \nu_Q(t) := Q(\{(\phi, n, r) : \omega \in \mathcal{W} : \phi(t) \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^d), t \in [0, 1].
\end{align*}
\]

(For a description of $\mathcal{B}(\mathcal{P}(\mathbb{R}^d))$ see [21] Lemma A.5.1).

Since $\nu_Q(t) = \mathcal{L}(\tilde{X}^{\nu_Q} \omega)$, we are thus interested in weak solutions to the limiting controlled McKean-Vlasov SDE

\[
\begin{align*}
(12) \quad d\tilde{X}_t &= \int_{\mathbb{R}^d} \left\{ |\nabla_y \Phi(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) + I| \left[ b(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) + \sigma(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) \right] z\nu_t(dz) \\
&+ \nabla_x \Phi(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) f(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) + A : \nabla_x \nabla_y \Phi(\tilde{X}_t, y, \mathcal{L}(\tilde{X}_t)) \right\} m_t(dy)dt \\
&+ B(t, \tilde{X}_t, \mathcal{L}(\tilde{X}_t))dW_t, \\
&= \frac{1}{2} B(t, x, \mu) B(t, x, \mu)^T = \int_{\mathbb{R}^d} \left[ \nabla_y \Phi(x, y, \mu) + \frac{1}{2} A(x, y, \mu) + f(x, y, \mu) \otimes \Phi(x, y, \mu) \right] m_t(dy)
\end{align*}
\]

Note that in the case $\int_{\mathbb{R}^m} z\nu_t(\omega)(dz) = 0$ and $m_t(dy) = \pi(dy|\tilde{X}_t, \mathcal{L}(\tilde{X}_t))$ for almost every $\omega \in \tilde{\Omega}$ and $t \in [0, 1]$, this agrees with Equation (6).

The process triple $(\tilde{X}, m, \rho)$ can be given explicitly as the coordinate process on the probability space $(\mathcal{W}, \mathcal{B}(\mathcal{W}), Q)$ endowed with the canonical filtration $\mathcal{G}_t := \sigma \left( (\tilde{X}_s, m(s), \rho(s)), 0 \leq s \leq t \right)$. Thus, for $\omega = (\phi, n, r) \in \mathcal{W}$,

\[
\begin{align*}
(13) \quad \tilde{X}_t(\omega) &= \phi(t), \quad m_t(\omega) = \pi|_{\mathcal{B}(\mathbb{R}^m \times [0, t])}, \quad \rho(t, \omega) = r|_{\mathcal{B}(\mathbb{R}^m \times [0, t])}.
\end{align*}
\]
Thus, for \( g : \mathbb{T}^d \to \mathbb{R} \), when we write \( \mathbb{E}^Q \left[ \int_{T^d \times [s,t]} g(y)m(t)(dyd\tau) \right] \), we mean

\[
\mathbb{E}^Q \left[ \int_{T^d \times [s,t]} g(y)m(t)(dyd\tau) \right] = \int_{\mathcal{W}} \int_{T^d \times [s,t]} g(y)m(t,\omega)(dyd\tau) Q(d\omega)
\]

\[
= \int_{X \times Y \times Z} \int_{T^d \times [s,t]} g(y)n(z)Q(dyd\tau)Q(d\omega) + \int_{X \times Y \times Z} \int_{T^d \times [s,t]} g(y)n_{\tau}(dy)Q(d\omega)d\tau.
\]

Throughout this paper we will only integrate \( m(t,\omega) \) against time intervals of the form \([s,t]\), so we will simply write \( \mathbb{E}^Q \left[ \int_{T^d \times [s,t]} g(y)m(dyd\tau) \right] \) in the place of \( \mathbb{E}^Q \left[ \int_{T^d \times [s,t]} g(y)m(t)(dyd\tau) \right] \) and \( n(dyd\tau) \) in the place of \( n_{\tau}(dy)Q(d\omega) \). The same applies to \( \rho(t,\omega) \) and \( \nu_{\tau}(dy)Q(d\omega) \).

We end this section with a discussion on well posedness of the limiting controlled McKean-Vlasov SDE [12].

**Definition 2.2.** We will say weak-sense uniqueness holds for Equation (12) if under the assumptions:

1. \( \Theta, \tilde{\Theta} \in \mathcal{V} \), where \( \mathcal{V} \) is defined in Definition 3.1.
2. \( \Theta_Z = \tilde{\Theta}_Z \)

we have \( \Theta_X = \tilde{\Theta}_X \).

**Proposition 2.3.** Under the assumptions [(A1)-(A8)], weak-sense uniqueness as defined in Definition 2.2 holds for Equation (12).

**Proof.** Since the product of bounded Lipshitz functions is Lipshitz (\( \Phi \) is Lipshitz in \( x, y \) because it has bounded first order derivatives and in \( \mu \) by Corollary [C.3]), the averaged coefficients are Lipshitz continuous by Section 6.2 of [74]. Then the requirements for Theorem 2.1 in [62] are satisfied by the limiting system. By a slight modification of the proof to include the control term (in particular the Yamada-Wantanabe Theorem used in Lemma 2.3 does not directly apply, but a modification of its proof is simple given we consider relaxed controls \( \rho \) such that \( \int_0^1 |z|^2 \rho(|z|d|z|dt < \infty) \) we get that for any fixed \( \rho \) there is a unique strong solution to Equation (10) under assumptions [(A1)-(A6)]. Then, noting that by Remark 6.7 that \( \Theta|_{B(X \times Y)} \) is completely determined by \( \Theta_X \), we get immediately that weak-sense uniqueness holds by the proof of Proposition C.2 in [30].

3. Statement of the Main Results

In order to state the main results of this paper, we need the following two definitions:

**Definition 3.1.** We will say \( \Theta \in \mathcal{P}(\mathcal{W}) \) is in \( \mathcal{V} \) if

(V1) \( \Theta \) corresponds to a weak solution \( \tilde{X} \) of (12).

(V2) \( \mathbb{E}^\Theta \left[ \int_{[0,1]} \mathbb{E}^{\tilde{X},\nu_\Theta} \left[ \int_{[0,1]} |z|^2 \rho(|z|d|z|dt \right] < \infty \right. \).

(V3) \( \nu_\Theta(0) = \nu_0 \) from (A6).

(V4) \( \forall t \in [0,1], g \in C_b^2(\mathbb{T}^d) \), \( \mathbb{E}^\Theta \left[ \int_{[0,1]} \mathbb{E}^{\tilde{X},\nu_\Theta(s)} \left[ \int_{[0,1]} g(y)m(dyds) \right] \right] = 0. \)

Where here we are using the notation for the coordinate process given in Equation (13).

Essentially, condition (V4) says that the coordinate process \( m_s(dy) = \pi(dy|\tilde{X}_s,\nu_\Theta(s)) \) is the invariant measure for the fast dynamics, for \( \Theta \)-a.e. \( \omega \in \mathcal{W} \). A detailed explanation on why this is true is in Section 6.2 and in particular in Remark 6.7.

**Definition 3.2.** A function \( I : \mathcal{P}(\mathcal{X}) \to [0,\infty] \) is called a (good) rate function if for each \( M < \infty \), the set \( \{ \theta \in \mathcal{P}(\mathcal{X}) : I(\theta) \leq M \} \) is compact. We say that a Laplace principal holds for the family \( \{\mu_N\}_{N \in \mathbb{N}} \) with rate function \( I \) if for any bounded, continuous \( F : \mathcal{P}(\mathcal{X}) \to \mathbb{R} \),

\[
\lim_{N \to \infty} -\frac{1}{N} \log \mathbb{E}\left[ \exp(-NF(\mu_N)) \right] = \inf_{\theta \in \mathcal{P}(\mathcal{X})} \{ F(\theta) + I(\theta) \}
\]
It is well known that in our setting the Laplace principal holds if and only if \( \{ \mu^N \}_{N \in \mathbb{N}} \) satisfies a LDP with rate function \( I \). See [21] Theorem 1.2.3.

In order to prove the Laplace Principal for \( \{ \mu^N \}_{N \in \mathbb{N}} \), we make use of the following proposition:

**Proposition 3.3.** The pre-limit expression in (14) can be written as

\[
\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] = \inf_{u^N \in \mathcal{U}_N} \left[ \frac{1}{2} \mathbb{E} \left[ \frac{1}{N} \int_0^1 |u^N(t)|^2 \, dt \right] + \mathbb{E}[F(\bar{\mu}^N)] \right]
\]

for any \( F \in C_0(\mathcal{P}(\mathcal{X})) \) where \( \bar{\mu}^N \) is given by (3) with \( u^N = (u^N_1, ..., u^N_N) \in \mathcal{U}_N \) the control in Equation 7.

**Proof.** By Proposition A.1 and [63] there is Borel measurable \( \psi^i,N \) such that

\[
\psi^i,N((x^{1:N}, ..., x^{N:N}), (W^1, ..., W^N)) = X_i^N,
\]

and by the characterization of \( B(\mathcal{P}(\mathcal{X})) \) given in Lemma A.5.1 of [21] \( p : C([0,1];\mathbb{R}^d)^N \to \mathcal{P}(C([0,1];\mathbb{R}^d)) \) given by \( p(\phi_1, ..., \phi_N) = \frac{1}{N} \sum_{i=1}^N \delta_{\phi_i} \) is Borel measurable. So

\[
\bar{\mu}^N = p(\psi^1,N((x^{1:N}, ..., x^{N:N}), (W^1, ..., W^N)), ..., \psi^N,N((x^{1:N}, ..., x^{N:N}), (W^1, ..., W^N)))
\]

and is thus a Borel-measurable function of the driving Wiener processes for each \( N \). Then Theorem 3.6 in [10] applies, giving us the desired result. \( \square \)

**Theorem 3.4.** Under assumptions \([A1],[A8]\) the sequence of \( \mathcal{P}(\mathcal{X}) \)-valued random variables \( \{ \mu^N \}_{N \in \mathbb{N}} \) as defined by Equation 3 satisfies the Laplace Principal with good rate function

\[
I(\theta) = \inf_{\Theta \in \mathcal{Y}; \Theta_X = \theta} \mathbb{E}^{\theta} \left[ \frac{1}{2} \int_{\mathbb{R}^\infty \times [0,1]} |z|^2 \rho(dz,dt) \right]
\]

where \( \inf(\emptyset) := +\infty \).

**Proof.** As is standard, in order to prove that \( \{ \mu^N \}_{N \in \mathbb{N}} \) satisfies the Laplace Principal with rate function \( I \), we prove the Laplace Principal lower bound [42] in Section 7 and the Laplace Principal upper bound [14] in Section 8. The main tool in these proofs is the Variational Representation Theorem for Functionals of Brownian Motion, given in Proposition 3.3. Once we identify the law of large numbers result for the controlled process in Section 6 the Laplace Principal lower bound, i.e.

\[
\lim \inf_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \geq \inf_{\Theta \in \mathcal{P}(\mathcal{X})} \{ F(\theta) + I(\theta) \}
\]

follows immediately from Fatou's lemma, as seen in Section 7. Then to prove the Laplace Principal upper bound,

\[
\lim \sup_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \leq \inf_{\Theta \in \mathcal{P}(\mathcal{X})} \{ F(\theta) + I(\theta) \},
\]

in Section 8 given \( \theta \in \mathcal{P}(\mathcal{X}) \), we construct a probability space and a sequence of iid controls whose law corresponds to \( \Theta_Z \) for \( \Theta \in \mathcal{Y}, \Theta_X = \theta \), that nearly reaches the infimum in the definition of \( I(\theta) \). Once we show these two bounds, we get

\[
\inf_{\Theta \in \mathcal{P}(\mathcal{X})} \{ F(\theta) + I(\theta) \} \leq \lim \inf_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \leq \lim \sup_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))],
\]

so that Equation 14 is satisfied.

In Section 9 we prove that the level sets of \( I \) are compact, so indeed \( I \) is a good rate function. \( \square \)
Remark 3.5. See Section 4 for a formal connection of this rate function to known rate functions in the literature which are not of variational form, and Theorem 5.2 for an alternative variational form of the rate function for a certain subclass of systems.

Remark 3.6. We can see that the Law of the control in the definition of $I$ depends on $\theta$ and the invariant measure $\pi$ directly as follows: decomposing $\Theta \in V$ into stochastic kernels as $\Theta(d\omega d\nu d\sigma) = \gamma(d\nu|n, \phi)\lambda(d\nu|\phi)\Theta_X(d\phi) = \gamma(d\nu|n, \phi)\lambda(d\nu|\phi)\Theta_X(d\phi)$, we have by Remark 6.7 that $\lambda(d\nu|\phi) = \delta_\pi(d\nu|\phi(s), \theta(s)) \|d\nu\|_\phi$, which we will abbreviate as $\delta_\pi(d\nu)$. Here by $\theta(s)$ we mean $\theta \circ ev^{-1}(s)$. Then, using the fact that the control only appears linearly in the dynamics of Equation 12

$$E^\Theta \left[ \frac{1}{2} \int_{\mathbb{R}^n \times [0,1]} |z|^2 \rho(dz dt) \right] = \int_\mathbb{R}^n \frac{1}{2} \int_{\mathbb{R}^n \times [0,1]} |z|^2 \rho(\omega)(dz dt) \Theta(d\omega)$$

$$= \int_X \int_{\mathbb{R}^n} \frac{1}{2} \int_0^1 \int_{\mathbb{R}^n} |z|^2 r_t(dz) dt \gamma(d\nu|n, \phi) \delta_\pi(d\nu) \theta(d\phi)$$

$$= \int_X \int_{\mathbb{R}^n} \frac{1}{2} \int_0^1 \int_{\mathbb{R}^n} |z|^2 r_t(dz) dt \gamma(d\nu|\pi, \phi) \theta(d\phi).$$

In order to solve the variational problem which defines the rate function $I$, one must find a minimizing stochastic kernel $\gamma$.

In proving Theorem 3.4 we will also have proved the following convergence result:

**Theorem 3.7.** Let $ev : X \to \mathbb{R}^d$ be the evaluation map at time $t$ and $\{\mu^N\}$ be as defined by Equation 4. Under assumptions $(A1), (A8) \mathcal{L}(\mu^N) \to \delta_{\mu^*}$ in $\mathcal{P}(\mathcal{P}(X))$, where deterministic $\mu^* \in \mathcal{P}(X)$ satisfies $\mu^* \circ ev^{-1}(t) = \mathcal{L}(X_t), t \in [0,1]$ for $X$ solving the McKean-Vlasov SDE 10.

**Proof.** This follows immediately from the proofs in Section 3 by taking $u^N \equiv 0$ for all $N \in \mathbb{N}$. \(\square\)

**Remark 3.8.** It is worth noting that via an integration-by-parts argument, the diffusion term in Equation 6 can also be written as

$$B(x, \mu)B(x, \mu)^\top = \int_{\mathbb{R}^n} [I + \nabla_y \Phi(x, y, \mu)] A(x, y, \mu)[I + \nabla_y \Phi(x, y, \mu)]^\top \pi(dy|x, \mu).$$

See 4 Chapter 3 Section 6.2.

4. **Formal Connection to Rate Functions in the Existing Literature**

The goal of this section is to connect, at least at a formal level, the rate function obtained in Theorem 3.4 to the ones obtained in the literature for $\delta = 1$, which is the result of 15, and for $\delta = 1$, which is the result of 20. As noted in Section 7.1 of 11, in the absence of multiple scales there is a heuristic connection between the variational form of the rate function of weakly interacting particles and that from the classical paper 16. In order to connect this to our case, we also must compare the LDP for small noise diffusions from 27 to the rate function for combined small noise and averaging from 20.

Assume $d = m, f \equiv 0, \sigma = I$, and the initial conditions of all the particles is 0. Then $\pi$ is Lebesgue measure and $\Phi = 0$. The limiting system given by Equation 12 is given by

$$d\dot{X}_t = \left\{ \int_{\mathbb{T}^n} \{ b(X_t, y, \mathcal{L}(X_t)) + \int_{\mathbb{R}^d} z \rho_t(dz) \} dy \right\} dt + dW_t.$$

Since in 20, the joint small noise and scale separation limit is being taken, we expect that we should be able to formally connect our rate function, given by Equation 15, to $I_{DG}$, the rate function from 16, and $I_{DS}$, the rate function from 20, by assuming that the analogous rate function to $I_{DS}$ which acts on $\mathcal{P}(X)$ instead of $X$ would be given through the same relationship as the small noise rate function from 27 to the mean-field rate function from 16.

The rate function in 16 for the empirical measure for strong solutions to

$$dX_t^{1:N} = b(X_t^{1:N}, \mu_t^{1:N}) + dW_t$$

is given by

$$I_{DG} = \int_{\mathbb{T}^n} \{ b(X_t, y, \mathcal{L}(X_t)) + \int_{\mathbb{R}^d} z \rho_t(dz) \} dy.$$

In the absence of multiple scales, the rate function $I_{DG}$ reduces to $I_{DS}$ above. The goal of 16 is to rigorously construct a rate function which connects $I_{DG}$ and $I_{DS}$ of 20. This is accomplished via

$$I_{DG} = \int_{\mathbb{T}^n} \{ b(X_t, y, \mathcal{L}(X_t)) + \int_{\mathbb{R}^d} z \rho_t(dz) \} dy.$$
acts on \( \theta \in C([0,1]; \mathcal{P}(\mathbb{R}^d)) \) by
\[
I_{DG}(\theta) = \frac{1}{2} \int_0^1 \sup_{g \in \mathcal{D}: \| \nabla g \|^2, \theta(t) \neq 0} \frac{|\langle g, \dot{\theta}(t) - \mathcal{L}^2(\theta(t))^* \theta(t) \rangle|^2}{\langle |\nabla g|^2, \theta(t) \rangle} dt
\]
if \( \theta : [0,1] \to \mathcal{P}(\mathbb{R}^d) \) is absolutely continuous in the distribution sense (see Definition 4.1 in [16]) and \( I_{DG}(\theta) = +\infty \) otherwise. Here \( \dot{\theta}(t) \) denotes the distributional derivative of \( \theta \) at time \( t \), \( \mathcal{D} \) is the Schwartz space of compactly supported smooth test functions \( g : \mathbb{R}^d \to \mathbb{R} \), and \( \mathcal{L}^2(\theta(t)) \) acts on \( g \in \mathcal{D} \) according to
\[
\mathcal{L}^2(\theta(t))g(x) := b(x, \theta(t)) \cdot \nabla g(x) + \frac{1}{2} \sum_{j,k=1}^d \frac{\partial^2 g}{\partial x_j \partial x_k}(x).
\]

Under the current assumptions, the rate function for \( \{X^t\}_{t>0} \), the strong solutions to
\[
dX_t = b(X_t) + \sqrt{c}dW_t,
\]
of [27] acts on \( \phi \in \mathcal{X} \) by:
\[
I_{DE}(\phi) = \frac{1}{2} \int_0^1 |\dot{\phi}(t) - b(\phi(t))|^2 dt
\]
for \( \phi \in H^1([0,1]; \mathbb{R}^d) \) and \(+\infty\) otherwise.

The rate function for \( \{X^\delta\}_{\delta>0} \), the strong solutions to
\[
dX^\delta_t = b(X^\delta_t, X^\delta_t/\delta) + \sqrt{\epsilon}dW_t,
\]
is given in [28] acts on \( \phi \in \mathcal{X} \) by
\[
I_{DS}(\phi) = \frac{1}{2} \int_0^1 |\dot{\phi}(t) - r(\phi(t))|^2 dt
\]
for \( \phi \) absolutely continuous and \(+\infty\) otherwise. Here \( \epsilon(\delta) \downarrow 0 \) as \( \delta \downarrow 0 \) such that \( \frac{\epsilon}{\delta} \to \infty \) as \( \delta \downarrow 0 \) and
\( r(x) := \int_{\mathbb{R}^d} b(x,y)dy \).

Comparing \( I_{DE} \) to \( I_{DG} \) and \( I_{DS} \), we see our rate function can be expected to be the same as \( I_{DG} \) but with \( b(x, \theta(t)) \) replaced by \( r(x, \theta(t)) := \int_{\mathbb{R}^d} b(x,y, \theta(t))dy \). In other words, in the notation of [16], we expect the rate function to have the representation
\[
S(\theta) = \frac{1}{2} \int_0^1 \sup_{g \in \mathcal{D}: \| \nabla g \|^2, \theta(t) \neq 0} \frac{|\langle g, \dot{\theta}(t) - \mathcal{L}^3(\theta(t))^* \theta(t) \rangle|^2}{\langle |\nabla g|^2, \theta(t) \rangle} dt
\]
if \( \theta : [0,1] \to \mathcal{P}(\mathbb{R}^d) \) is absolutely continuous in the distribution sense and \( S(\theta) = +\infty \) otherwise. Here \( \mathcal{L}^3(\theta(t)) \) acts on \( g \in \mathcal{D} \) according to
\[
\mathcal{L}^3(\theta(t))g(x) := \left( \int_{\mathbb{R}^d} b(x,y, \theta(t))dy \right) \cdot \nabla g(x) + \frac{1}{2} \sum_{j,k=1}^d \frac{\partial^2 g}{\partial x_j \partial x_k}(x).
\]

Formally we can see the connection between our rate function given by Equation [16] and \( S \) by denoting \( \mathcal{V} \) the class of \( \Theta \in \mathcal{V} \) such that \( \Theta \)-almost surely, \( \rho_t(dx) = \delta_{v(t,x)}(dx) \), for some \( v : [0,1] \times \mathbb{R}^d \to \mathbb{R} \) such that \( v(t,\cdot) \in \mathcal{D}, \forall t \in [0,1] \). Applying Itô’s formula, we get for \( g \in \mathcal{D} \) and \( \Theta \in \mathcal{V} \) with \( \Theta_X = \theta \) that
\[
g(\bar{X}_{t+h}) - g(\bar{X}_t) = \int_t^{t+h} \mathcal{L}^3(\theta(s))|g|(\bar{X}_s) ds + \int_t^{t+h} \nabla v(s, \bar{X}_s) \cdot \nabla g(\bar{X}_s) ds + \int_t^{t+h} \nabla \cdot g(\bar{X}_s) dW_s.
\]

Taking expectations, dividing by \( h \), and sending \( h \to 0 \), we get
\[
\langle g, \dot{\theta}(t) - \mathcal{L}^3(\theta(t))^* \theta(t) \rangle = \langle \nabla v(t,\cdot) \cdot \nabla g, \theta(t) \rangle, \quad \forall t \in [0,1].
\]

Then
\[
\sup_{g \in \mathcal{D}: \| \nabla g \|^2, \theta(t) \neq 0} \frac{|\langle g, \dot{\theta}(t) - \mathcal{L}^3(\theta(t))^* \theta(t) \rangle|^2}{\langle |\nabla g|^2, \theta(t) \rangle} = \sup_{g \in \mathcal{D}: \| \nabla g \|^2, \theta(t) \neq 0} \frac{|\langle \nabla v(t,\cdot) \cdot \nabla g, \theta(t) \rangle|^2}{\langle |\nabla g|^2, \theta(t) \rangle}.
\]
By Hölder’s inequality the expression on the right hand side is bounded above by \( \langle |\nabla v(t, \cdot)|^2, \theta(t) \rangle \), with equality reached when \( g(x) = v(t, x) \). So

\[
\sup_{g \in D: |\nabla g|, g(t) \neq 0} \frac{|\langle g, \dot{\theta}(t) - \mathcal{L}^1(\theta(t)) \mu(t) \rangle^2|}{\langle |\nabla g|^2, \theta(t) \rangle} = \langle |\nabla v(t, \cdot)|^2, \theta(t) \rangle.
\]

Since this holds for every \( \Theta \in \mathcal{V}^1 \), it then stands that it holds for \( \Theta \in \mathcal{V}^1 \) along which the infimum in

\[
I_{\mathcal{V}^1}(\Theta) := \inf_{\Theta \in \mathcal{V}^1: \Theta x = \theta} \mathbb{E}^{\Theta} \left[ \frac{1}{2} \int_0^1 \int_{\mathbb{R}^d} |z|^2 \mu_t(dz) dt \right]
\]

is attained, so \( I_{\mathcal{V}^1} = S \). Thus we see if indeed the infimum over \( \Theta \in \mathcal{V} \) in Equation 16 can be restricted to \( \mathcal{V}^1 \), then our rate function is equivalent to \( S \).

5. An example with a bistable confining potential

Even though we have chosen for presentation purposes to prove Theorem 3.4 under strong boundedness assumptions on the coefficients, the extended results of Section 10 show that Theorem 3.4 holds true under greater generality. In particular, if one can appropriately control the behavior of the solution to (11) then the results can be extended. As we explain in detail in Section 10, in the absence of such a general theory, one in principle would need to confirm this in a case by case situation. In this example, we do so for the popular interacting particle system considered in the classical work [15] where the drift is a bistable, confining potential.

Consider 1-D the system of weakly interacting particles from Dawson’s classical paper [15], where the interacting potential is modified from \( V(x) = x^4/4 - x^2/2 \) to \( V^\delta(x) = x^4/4 - x^2/2 - \frac{\epsilon}{2\pi} \cos(2\pi x/\delta) \) for some \( \epsilon > 0 \). Then the system of controlled SDE’s is given by

\[
d\bar{X}_t^{i,N} = \left[ -(\bar{X}_t^{i,N})^3 + \bar{X}_t^{i,N} + \sigma v_t^{i,N}(t) - \kappa (\bar{X}_t^{i,N} - v_t^{i,N}) - \frac{\epsilon}{\delta} \sin(2\pi \bar{X}_t^{i,N}/\delta) \right] dt + \sigma dW_t^i
\]

where \( \sigma, \kappa > 0, v_t^{i,N} := \langle \cdot, \bar{\mu}_t^{i,N} \rangle = \frac{1}{N} \sum_{i=1}^N \bar{X}_t^{i,N}, i \in \{1, ..., N\}, t \in [0, 1] \), and \( W_t^i \) are independent 1-D Brownian motions.

In order to confirm that the result of Theorem 3.4 holds in this case, we need to verify that Assumptions (A1'), (A12') of Section 10 hold (see Theorem 10.2). After doing so, we discuss how the typical events and large deviations principle look here, and provide an alternative variational form for the rate function.

5.1. Verification that the necessary assumptions hold. Let us start by assuming that

1. (A1') in Section 10 holds.
2. \( \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=1}^N |\bar{X}_t^{i,N}| < D, D \in (0, \infty) \).

To see that the uncontrolled equation has a unique strong solution, one can write the system as a 2N dimensional SDE in the same way as in the proof of Proposition A.1 and use a standard truncation argument for SDEs with one-sided Lipschitz drift and the fact that the solution is nonexplosive (see [15] p.37 and [17]). In fact for this particular case, Theorem 1 in [61] directly applies for each \( N \in \mathbb{N} \). Thus (A2') holds.

To see (A3') holds, we can apply standard PDE theory as in the bounded case. Further, since we are in the case \( d = m = 1 \), we can solve for the density \( \bar{\pi} \) of \( \pi \) explicitly. We get that \( \bar{\pi}(y) \) is the solution to

\[ -\left( -\epsilon \sin(2\pi y) \bar{\pi}(y) \right)' + \frac{1}{2} \sigma^2 \bar{\pi}''(y) = 0, \quad \bar{\pi} \text{ 1 - periodic, } \int_0^1 \bar{\pi}(y) dy = 1. \]

Solving, we get

\[ \bar{\pi}(y) = \frac{1}{Z} \exp \left( \frac{\epsilon}{\sigma^2 \pi} \cos(2\pi y) \right), \quad Z = \int_0^1 \exp \left( \frac{\epsilon}{\sigma^2 \pi} \cos(2\pi x) \right) dx. \]

The cell problem takes the form

\[ -\epsilon \sin(2\pi y) \Phi'(y) + \frac{1}{2} \sigma^2 \Phi''(y) = \epsilon \sin(2\pi y), \quad \Phi \text{ 1 - periodic, } \int_0^1 \Phi(y) \bar{\pi}(y) dy = 0. \]
and its unique solution is
\[ \Phi(y) = \frac{1}{Z} \int_0^y \exp\left(-\frac{\epsilon}{\sigma^2 \pi} \cos(2\pi x)\right)dx - y, \quad \hat{Z} = \int_0^1 \exp\left(-\frac{\epsilon}{\sigma^2 \pi} \cos(2\pi x)\right)dx. \]

So, keeping in mind that \( \Phi \) is 1-periodic and smooth, we have
\[ \sup_{y \in \mathbb{R}} \left\{ |\Phi(y)| + |\Phi'(y)| + |\Phi''(y)| \right\} \leq K \]
for some \( K > 0 \). Thus, \((A3')\) and \((A4')\) hold, and we attain some apriori bounds that will help us in proving \((A9')\). Since \( \Phi \) is independent of \( \mu \), \((A5')\) and \((A10')\) hold trivially, and by our explicit representation \((A6')\) holds.

The proof that \((A7')\) and \((A8')\) hold is given by Proposition 5.1 where we show that in this particular example we can take \( p = 4 \) and \( p_2 = 6 \).

**Proposition 5.1.** If \( \sup_{N \in \mathbb{N}} \mathbb{E} \left[ \sum_{i=1}^N \int_0^1 |u_i^N(t)|^2 dt \right] < B, B \in (0, \infty) \) then for each \( t \in [0, 1] \),
\[
\sup_{N \in \mathbb{N}} \mathbb{E} \left[ \sum_{i=1}^N \left| X_i^1 \right|^4 \right] + \sup_{N \in \mathbb{N}} \mathbb{E} \left[ \int_0^1 \left| \bar{X}_i^1 \right|^6 dt \right] \leq C(\kappa, \sigma, B, D, \epsilon).
\]

For presentation purposes we offer the proof of Proposition 5.1 in Appendix D. We next confirm that \((A9')\) is satisfied. We identify that
\[ b(x, y, \mu) = -x^3 + x - \kappa(x - \langle \cdot, \mu \rangle), \]
\[ f(x, y, \mu) = -\epsilon \sin(2\pi y), \sigma(x, y, \mu) = \sigma > 0, \]
\[ \Phi(x, y, \mu) = \Phi(y) \]

is bounded with bounded derivatives,

Thus, \( |b(x, y, \mu)| \leq C(1 + |x|^3 + \mu(|\cdot|)) \) and all the other terms are bounded uniformly in \( x \) and \( \mu \). This gives via the triangle and Cauchy-Schwarz inequalities that all the conditions in \((A9')\) are satisfied.

Lastly, we check that weak-sense uniqueness as given in Definition 2.2 holds, so that \((A12')\) is satisfied. For this, we first write down the limiting controlled system:
\[ d\bar{X}_t = \left\{ \int_\mathcal{Y} \Phi'(y) + 1 \left[ 1 - \hat{X}_t \right] \bar{X}_t_t + \kappa \left( \int_\mathcal{Y} x \mathcal{L}(\bar{X}_t)dx \right) \right\} dt + \sigma \sqrt{\int_\mathcal{Y} \Phi'(y)^2} m_t(dy) dW_t, \]
\[ (19) \]
where here we used Remark 3.8 to simplify the diffusion term. Since for all \( m \in \mathcal{Y} \) and all \( t \in [0, 1] \), \( c_1(t) = \int_\mathcal{Y} \Phi'(y) + 1 m_t(dy) \) and \( c_2(t) = \int_\mathcal{Y} 1 + \Phi'(y)^2 m_t(dy) \) are bounded and \( c_1(t) > 0 \), we quickly see that due to the assumed \( L^2 \) bound on the control in the conditions of Definition 2.2 that the proof goes through in the same way as in Appendix A of [15].

### 5.2. Form of the limiting theorems and equivalent formulations.

Now that we have confirmed that the necessary assumptions indeed hold for the system \([18]\) let us discuss what the law of large numbers and large deviations principle look like.

As already discussed, the limiting controlled system takes the form \((19)\). In particular, noting that
\[ \int_\mathcal{Y} \Phi'(y)^2 \pi(dy) = \int_\mathcal{Y} \left[ \Phi'(y) + 1 \right] \pi(dy) = \frac{1}{ZZ}, \]
we get that Theorem 3.7 (equivalently Theorem 10.3) holds and \( \mathcal{L}(\mu^N) \rightarrow \delta_{\mu^*} \) in \( \mathcal{P}(\mathcal{P}(\mathcal{X})) \), where deterministic \( \mu^* \in \mathcal{P}(\mathcal{X}) \) satisfies \( \mu^* \circ ev^{-1}(t) = \mathcal{L}(X_t), t \in [0, 1] \) for \( X \) solving the McKean-Vlasov SDE:
\[ dX_t = \frac{1}{ZZ} \left[ -X_t^3 + (1 - \kappa)X_t + \kappa \left( \int_\mathcal{Y} x \mathcal{L}(X_t)dx \right) \right] dt + \sigma \sqrt{\frac{1}{ZZ}} dW_t. \]
Analogously, Theorem 3.4 (equivalently Theorem 10.2) shows that the Large Deviations Principle holds with rate function given by (50) where \( \Theta \) corresponds to a weak solution of (19) according to Definition 10.1.

Unlike the system offered in Section 4 the nontrivial interaction of this system with the empirical measure appears linearly in the coefficient which blows up in the drift. For example, check that the required assumptions to satisfy Theorem 10.2 also hold for slightly more complicated cases.

In preparation for stating this result, let us set \( \hat{\mathcal{V}} = \mathcal{X} \times \hat{\mathcal{Y}} \) where \( \mathcal{X} = C([0,1]; \mathbb{R}^d) \) and \( \hat{\mathcal{Y}} = \{ \hat{\tau} \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^m \times [0,1]): \hat{\tau}(\mathbb{T}^d \times \mathbb{R}^m \times [0,t]) = t, \forall t \in [0,1] \text{ and } \int_{\mathbb{T}^d \times \mathbb{R}^m \times [0,1]} |z| \hat{\tau}(dy \times dz \times dt) < \infty \}. \)

We shall write that \( \hat{\Theta} \in \mathcal{P}(\hat{\mathcal{V}}) \) is in \( \hat{\mathcal{V}} \) if

\[
(\hat{V}1) \quad \hat{\Theta} \text{ corresponds to a weak solution } \hat{X} \text{ of (22)},
\]

\[
(\hat{V}2) \quad \mathbb{E}^\hat{\Theta} \left[ \int_{\mathbb{T}^d \times \mathbb{R}^m \times [0,1]} |z|^2 \hat{\rho}(dydzdt) + \int_0^1 |\hat{X}_t|^4 dt \right] < \infty.
\]

\[
(\hat{V}3) \quad \hat{\nu}_\Theta(0) = \nu_0, \text{ where } \hat{\nu}_\Theta \text{ is as in Equation (11) but acting on } \mathcal{P}(\hat{\mathcal{V}}).
\]

\[
(\hat{V}4) \quad \forall t \in [0,1], g \in C^2_b(\mathbb{T}^d), \quad \mathbb{E}^{\hat{\Theta}} \left[ \int_{\mathbb{T}^d \times \mathbb{R}^m \times [0,t]} \mathcal{L}_{\hat{X}_s, \nu_\Theta(s)} g(y) \hat{\rho}(dydzdt) \right] = 0.
\]

Now, we are ready to state Theorem 5.2.

**Theorem 5.2.** The rate function given by Theorem 10.2 (equivalently in Theorem 3.4) is equivalent to:

\[
\hat{I}(\theta) = \inf_{\hat{\Theta} \in \mathcal{P}(\hat{\mathcal{V}})} \mathbb{E}^\hat{\Theta} \left[ \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^m \times [0,1]} |z|^2 \hat{\rho}(dydzdt) \right]
\]

where \( \inf(\emptyset) := +\infty \). In particular, we have that \( I(\theta) = \hat{I}(\theta) \) for all \( \theta \in \mathcal{P}(\mathcal{X}) \).

The proof of Theorem 5.2 is deferred to Appendix 13. Note that the proof of this theorem extends to the more general case where \( \sigma(x,y,\mu) = \sigma(y,\mu), \quad f(x,y,\mu) = f(y,\mu) \), \( m = d \), \( \nabla_y \Phi(y,\mu) + I(\sigma(y,\mu)) \) is uniformly bounded in \( y \) and \( \mu \), and \( \int_{\mathbb{T}^d} |\nabla_y \Phi(y,\mu) + I(\sigma(y,\mu))| (dy|\mu) \) is invertible with uniformly bounded inverse for all \( \mu \), where as with Theorem 10.2, the exponent on \( |\hat{X}_t| \) in \( \hat{V}2 \) must be adapted to \( p \) in \( A7' \). For the sake of brevity, we leave the proof of this extension to the interested reader.

Lastly, even though we do not show this here, we do mention that along the same lines one is able to check that the required assumptions to satisfy Theorem 10.2 also hold for slightly more complicated cases where the empirical measure appears linearly in the coefficient which blows up in the drift. For example, one can consider a particle system of the form

\[
d\hat{X}_t^{i,N} = \left[ -(\hat{X}_t^{i,N})^3 + \hat{X}_t^{i,N} + \sigma u_t^{i,N}(t) - \kappa (\hat{X}_t^{i,N} - \nu_t^{i,N}) - \epsilon \cos(2\pi \hat{X}_t^{i,N}/\delta) \right] dt + \sigma dW_t^i.
\]

We leave the details of this example to the interested reader.

### 6. Limiting Behavior of the Controlled Empirical Measure

Our object of study in this section is the family of occupation measures \( \{Q^N\}_{N \in \mathbb{N}} \in \mathcal{P}(\hat{\mathcal{V}}) \) defined by:

\[
Q^N(A \times B \times C) = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{X}_t^{i,N}(\omega)}(A)\delta_{\nu_t^{i,N}(\omega)}(B)\delta_{\bar{\nu}_t^{i,N}(\omega)}(C)
\]
for $A \times B \times C \in \mathcal{B}(W)$, $\omega \in \Omega$, $\rho_i^N$ the ordinary relaxed control corresponding to $u_i^N$ via Equation 9, $X_i^N$ as in (7), and $m_i^N(\omega) \in \mathcal{Y}$ given by

$$m_i^N(\omega)(I \times D) := \int_I \delta_{(X_i^N(\omega)/\delta)}(D)dt$$

for $I \in \mathcal{B}([0,1])$ and $D \in \mathcal{B}(\mathbb{R}^d)$. We use the convention that for $s > 1$, $u_i^N(s) = 0, \forall i, N \in \mathbb{N}$.

**Remark 6.1.** For a discussion of another possible choice of occupation measures, see Remark 5.1 after the proof of Theorem 5.2 in Appendix E.

Assume that there exists $B > 0$ such that $\mathbb{P}$ almost-surely,

$$\sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=1}^{N} \int_0^1 |u_i^N(t)|^2 dt \leq B$$

We will prove the following two propositions:

**Proposition 6.2.** Under assumption (22), the sequence $\{\mathcal{L}(Q^N)\}_{N \in \mathbb{N}}$ is precompact in $\mathcal{P}(\mathcal{P}(W))$.

*Proof.* See Subsection 6.1 \hfill $\square$

**Proposition 6.3.** Under assumption (22), for $Q$ such that $\mathcal{L}(Q^N) \to \mathcal{L}(Q)$ in $\mathcal{P}(\mathcal{P}(W))$, $Q \in \mathcal{V}$ almost surely, where the class of measures $\mathcal{V}$ is described in Definition 3.4.

*Proof.* See Subsection 6.2 \hfill $\square$

### 6.1. Tightness of the Occupation Measures

We prove tightness of the occupation measures $\{Q^N\}_{N \in \mathbb{N}}$ defined in Equation (23) as $\mathcal{P}(W)$-valued random variables by proving tightness of each of the marginals, $Q^N_Z$, $Q^N_Y$, and $Q^N_X = \mu^N$ (as defined by Equation 8).

#### 6.1.1. Tightness of $Q^N_Z$

This will follow analogously to p.88-89 of [11]. We recall here the arguments for the readers convenience. Observe that

$$g(r) := \int_{\mathbb{R}^m \times [0,1]} |z|^2 r(dz \times dt)$$

is a tightness function on $\mathcal{R}_1$. Namely, it is bounded from below and has relatively compact level sets. Indeed, boundedness from below is obvious and in order to confirm the second property, for $c \in (0, \infty)$ let us set $R_c := \{r \in \mathcal{R}_1 : g(r) \leq c\}$. Chebyshev’s inequality for $M > 0$ gives that

$$\sup_{r \in R_c} r(\{|z| \in \mathcal{R}_1 \times \{[0,1]\} \leq \frac{c}{M^2}.$$

Therefore, $R_c$ is tight and thus relatively compact as a subset of $\mathcal{R}$. Let $\{r_n\}_{n \in \mathbb{N}} \subset R_c$ be such that $\{r_n\}_{n \in \mathbb{N}}$ converges weakly to $r_* \in \mathcal{R}$. We need to show that $r_*$ has finite first moment and that first moments of $\{r_n\}_{n \in \mathbb{N}}$ converge to the first moment of $r_*$. By Jensen’s inequality and Fatou’s lemma (Thm A.3.12 in [21]),

$$\sqrt{c} \geq \liminf_{n \to \infty} \sqrt{g(r_n)} \geq \liminf_{n \to \infty} \int_{\mathbb{R}^m \times [0,1]} |z|r_n(dz \times dt) \geq \int_{\mathbb{R}^m \times [0,1]} |z| r_*(dz \times dt).$$

Now letting $M > 0$, by Equation (26) and Hölder’s inequality, we have for all $r \in R_c$,

$$\int_{\mathbb{R}^m \times [0,1]} \mathbb{1}_{\{|z| \in \mathcal{R}_1 \times \{|z| > M\}} |z| r(dz \times dt) \leq \int_{\mathbb{R}^m \times [0,1]} |z|^2 r(dz \times dt) \int_{\mathbb{R}^m \times [0,1]} \mathbb{1}_{\{|z| \in \mathcal{R}_1 \times \{|z| > M\}} r(dz \times dt) \leq \sqrt{\frac{c}{M^2}} = \frac{c}{M}.$$

So by reverse Fatou’s Lemma we get

$$\limsup_{n \to \infty} \int_{\mathbb{R}^m \times [0,1]} |z|r_n(dz \times dt) \leq \frac{c}{M} \int_{\mathbb{R}^m \times [0,1]} \mathbb{1}_{\{|z| \in \mathcal{R}_1 \times \{|z| \leq M\}} |z| r_*(dz \times dt) \leq \frac{c}{M} \int_{\mathbb{R}^m \times [0,1]} |z| r_*(dz \times dt).$$
Given that $M$ may be taken to be arbitrarily large, we have

$$\lim_{n \to \infty} \int_{\mathbb{R}^m \times [0,1]} |z| r_n (dz \times dt) = \int_{\mathbb{R}^m \times [0,1]} |z| r_* (dz \times dt).$$

Thus we have $g$ is a tightness function on $R_1$. Now define $G : \mathcal{P}(Z) \to [0, \infty]$ by

$$G(\Theta) := \int_Z g(r) \Theta(dr).$$

Then $G$ is a tightness function on $\mathcal{P}(Z)$ (see Theorem A.3.17 in [21]). Thus in order to prove tightness of $\{Q^N\}_{N \in \mathbb{N}}$, it is enough to show that

$$\sup_{N \in \mathbb{N}} \mathbb{E}[G(Q^N)] < \infty.$$ 

But this follows immediately from assumption [25], since by definition of $G$ and $Q^N$,

$$\mathbb{E}[G(Q^N)] = \mathbb{E}\left[ \int_Z \int_{\mathbb{R}^m \times [0,1]} |z|^2 r(dzdt)Q^N_Z(dr) \right]$$

$$= \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho_i^{N}(dzdt) \right]$$

$$= \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \int_0^1 |u_i^{N}(t)|^2 dt \right]$$

$$< \infty.$$ 

6.1.2. Tightness of $Q^N$. First we note that $T^d \times [0, 1]$ is compact, so $\mathcal{M}(T^d \times [0, 1])$, where $\mathcal{M}(E)$ denotes the set of subprobability measures on $E$, is compact by Corollary A.3.16 in [21] (this also works for $\mathcal{M}(E)$, positive Borel measures $\mu$ on $E$ with $\mu(E) \leq \alpha$, for any $\alpha > 0$). Then by the proof of Lemma 3.3.1 in [21], $\mathcal{R}(T^d) \subset \mathcal{M}(T^d \times [0, 1])$ is closed in the topology of weak convergence (if a weakly converging sequence of measures on $T^d \times [0, 1]$ has the property that for each member of the sequence, its second marginal is Lebesgue measure, then this will also be true of the limiting measure), and hence $Y = R^1(T^d)$ is compact. Then $\mathcal{P}(Y)$ is compact, and hence $\mathcal{P}(\mathcal{P}(Y))$ is compact. Since $\{\mathcal{L}(Q^N)\}_{N \in \mathbb{N}} \subset \mathcal{P}(\mathcal{P}(Y))$, and on a metrizable space compactness implies sequential compactness, we immediately get $Q^N \subset \mathcal{P}(Y)$-random variable.

6.1.3. Tightness of $Q^N$. Let $E = \mathcal{P}(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$. We will prove tightness of $\{Q^N\}_{N \in \mathbb{N}} = \{\mu^N\}_{N \in \mathbb{N}}$ as $D(E[0,1])$-valued random variables, where $D(E[0,1])$ is the space of maps from $[0,1]$ to $E$ which are right continuous and have left-hand limits. Noting that $\mu^N(\omega) \in \mathcal{P}(X)$ for each $\omega \in \Omega, N \in \mathbb{N}$, we can treat $\{\mu^N\}_{N \in \mathbb{N}}$ as a sequence of $C([0,1]; \mathcal{P}(\mathbb{R}^d))$-valued random variables, so that indeed $\{\mu^N(\omega)\}_{N \in \mathbb{N}} \subset D(E[0,1])$, $\forall \omega \in \Omega, N \in \mathbb{N}$. Proving tightness of $\{\mu^N\}_{N \in \mathbb{N}}$ as $D(E[0,1])$-valued random variables will imply tightness as $C((0,1]; \mathcal{P}(\mathbb{R}^d))$-valued random variables by Problem 25 on p.153 of [22].

We will use Theorem 3.8.6 in [22] together with Theorem 3.1 in [39] to show the tightness of $\{\mu^N\}_{N \in \mathbb{N}}$ as $D(E[0,1])$-valued random variables.

Note that by Proposition 3.4.4. and Theorem 3.4.5. in [22], the class of functions

$$\mathcal{G} := \{ g \in C_b(E) : g(\mu) = \langle h, \mu \rangle, h \in C_c^\infty(\mathbb{R}^d) \}$$

separates points in $E$ and is closed under addition. Thus, by Theorem 3.1 in [39], if we show Lemma 6.3 holds, then to show tightness of $\{\mu^N\}_{N \in \mathbb{N}}$, it is enough to show tightness of $\{g \circ \mu^N\}_{N \in \mathbb{N}}$ as $D(E[0,1])$-valued random variables for each $g \in \mathcal{G}$.

**Lemma 6.4.** For each $\eta > 0$, there exists $K^\eta \subset \subset E$ such that

$$\mathbb{P}(\mu^N \notin \{ \mu \in D(E[0,1]) : \mu(t) \in K^\eta, \forall t \in [0,1] \}) < \eta.$$ 

Here by $A \subset \subset B$ we mean $A$ is a compact subset of $B$. Appealing to Theorem 8.6 and Remark 8.7 in [22], it will thus be sufficient to show:
Lemma 6.5. For \( q(x, y) := 1 \wedge |x - y| \) and \( 0 \leq t \leq t + \tau \leq 1, 0 \leq \tau \leq \gamma \), there exists a family \( \{\xi^N(\gamma)\}_{\gamma \in (0, 1), N \in \mathbb{N}} \) of non-negative random variables and \( r > 0 \) such that

\[
E \left[ q \left( g(\bar{\mu}_N(t + \tau)), g(\bar{\mu}_N(t)) \right) \mid \mathcal{F}_t^N \right] \leq E \left[ \xi^N(\gamma) \mid \mathcal{F}_t^N \right], \forall g \in \mathcal{G}
\]

and

\[
\lim_{\gamma \to 0} \limsup_{N \to \infty} E[\xi^N(\gamma)] = 0
\]

where here \( \mathcal{F}_t^N := \mathcal{F}_{\bar{\mu}_t}^N = \mathcal{F}_t \) since \( \bar{X}_{i,N}^t \) are strong solutions.

Note in particular that the following compact containment condition required by Theorem 3.8.6 in \[22\] as provided in the following corollary follows directly from Lemma 6.4 and its proof.

Corollary 6.6. For each \( \eta > 0 \) and \( t \in [0, 1] \), there exist \( K_t^\eta \subset \subset E \) such that:

\[
P(\bar{\mu}^N_t \not\in K_t^\eta) < \eta.
\]

We proceed with the proof of Lemma 6.4:

Proof of Lemma 6.4. Define \( K_L \subset \subset \mathbb{R}^d \) by \( K_L := \{ x \in \mathbb{R}^d : |x| \leq L \} \). Then

\[
E[\sup_{t \in [0, 1]} \bar{\mu}^N_t(\mathbb{R}^d \setminus K_L)] \leq \frac{1}{N} \sum_{i=1}^N \mathbb{P}(\sup_{t \in [0, 1]} |\bar{X}_{i}^t| > L)
\]

\[
\leq \frac{1}{N} \sum_{i=1}^N \frac{E[\sup_{t \in [0, 1]} |\bar{X}_{i}^t|^2]}{L^2}
\]

by Chebyshev’s inequality

\[
\leq \frac{C}{L^2}
\]

by Proposition A.2.

Now define

\[
K_L^t := \{ \nu \in \mathcal{P}(\mathbb{R}^d) : \nu(\mathbb{R}^d \setminus K_{(L+j)^2}) \leq \frac{1}{\sqrt{L+j}}, \forall j \in \mathbb{N} \}.
\]

By Prokhorov’s Theorem, \( K_L^t \subset \subset \mathcal{P}(\mathbb{R}^d) \) for each \( L \). Now we see that

\[
P(\bar{\mu}^N_t \not\in \{ \mu \in \mathcal{D}_E[0, 1] : \mu(t) \in K_L^t, \forall t \in [0, 1] \})
\]

\[
= P \left( \exists j \in \mathbb{N}, t \in [0, 1] : \sup_{t \in [0, 1]} |\bar{X}_{i}^t| > \frac{1}{\sqrt{L+j}} \right)
\]

\[
\leq \sum_{j=1}^\infty P \left( \sup_{t \in [0, 1]} |\bar{X}_{i}^t| > \frac{1}{\sqrt{L+j}} \right)
\]

\[
\leq \sum_{j=1}^\infty E \left[ \sup_{t \in [0, 1]} |\bar{X}_{i}^t| \right] \sqrt{L+j}
\]

by Chebyshev’s inequality

\[
\leq C \sum_{j=1}^\infty (L+j)^{-7/2}.
\]

This bound approaches 0 as \( L \to \infty \), so the lemma is proved.

Now we prove Lemma 6.5.
Proof of Lemma \ref{lemma:psi} We take \( r = 1 \). Now we observe that for all \( 0 \leq t \leq t + \tau \leq 1, 0 \leq \tau \leq \gamma \) and \( g \in \mathcal{G} \),

\[
\begin{align*}
\mathbb{E} \left[ q \left( g(\bar{\mu}^N(t + \tau)), g(\bar{\mu}^N(t)) \right) \right] & \leq \mathbb{E} \left[ \sum_{i=1}^{N} h(\bar{X}^i_{t+\tau}) - h(\bar{X}^i_{t}) \right] \\
& = \mathbb{E} \left[ \sum_{i=1}^{N} |h(\bar{X}^i_{t+\tau}) - h(\bar{X}^i_{t})| \right].
\end{align*}
\]

for some \( h \in C^\infty_c \).

Applying Itô’s formula to \( h \), we get (ignoring the arguments for notational convenience)

\[
h(\bar{X}^i_{t+\tau}) = h(x^i_{t}) + \int_{0}^{t} \left( \frac{1}{2} f + b + \sigma u^N_i \right) \cdot \nabla h + \frac{1}{2} A : \nabla \nabla h ds + \int_{0}^{t} \nabla h \cdot (\sigma dW^i_s).
\]

In order to control the term that blows up as \( \delta \to 0 \), we define \( \psi_l(x, y, \mu) := \Phi_l(x, y, \mu)h_{x_l}(x), l = 1, \ldots, d \), for \( \Phi \) as in Equation \ref{eq:Phi}. Then \( \psi_l \) solves

\[
L^1_{x,\mu} \psi_l(x, y, \mu) = -f_l(x, y, \mu)h_{x_l}(x).
\]

Now applying Itô’s formula (using Equations \ref{eq:psi} and \ref{eq:phi} in Proposition \ref{prop:psi} and the regularity of \( \Phi \) from Proposition \ref{prop:Phi}) to \( \psi_l \), we get

\[
\psi_l(\bar{X}^i_{t+\tau}, \bar{X}^i_{t}/\delta, \bar{\mu}^N_s) = \psi_l(x^i_{t}, x^i_{t}/\delta, \bar{\mu}^N_s) + \int_{0}^{t} \left( \frac{1}{2} f + b + \sigma u^N_i \right) \cdot \nabla x \psi_l + \frac{1}{2} A : \nabla \nabla x \psi_l + \frac{1}{\delta} \left( \frac{1}{\delta} f + b + \sigma u^N_i \right) \cdot \nabla y \psi_l + \frac{1}{2\delta^2} A : \nabla \nabla y \psi_l \right) ds + \int_{0}^{t} \nabla x \psi_l \cdot (\sigma dW^i_s) + \frac{1}{\delta} \int_{0}^{t} \nabla y \psi_l \cdot (\sigma dW^i_s)
\]

\[
+ \frac{1}{N} \sum_{j=1}^{N} \left[ \int_{0}^{t} \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s}) \cdot \right]
\]

\[
\left( \frac{1}{\delta} f(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) + b(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) + \sigma(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s)u^N_j(s) \right)
\]

\[
+ \frac{1}{2} A(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) : \partial_{\mu} \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s})
\]

\[
+ \frac{1}{2N} A(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) : \partial_{\mu} \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s)ds
\]

\[
+ \int_{0}^{t} \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s}) \cdot (\sigma(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s)dW^j_s)
\]

\[
+ \frac{1}{N} \int_{0}^{t} A : \nabla x \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s})ds + \frac{1}{N\delta} \int_{0}^{t} A : \nabla y \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s})ds
\]

\[
= \psi_l(x^i_{t}, x^i_{t}/\delta, \bar{\mu}^N_s) + \int_{0}^{t} \left( \frac{1}{2} f + b + \sigma u^N_i \right) \cdot \nabla x \psi_l + \frac{1}{2} A : \nabla \nabla x \psi_l + \frac{1}{\delta} \left( \frac{1}{\delta} f + b + \sigma u^N_i \right) \cdot \nabla y \psi_l
\]

\[
+ \frac{1}{\delta} A : \nabla \nabla y \psi_l ds + \int_{0}^{t} \nabla x \psi_l \cdot (\sigma dW^i_s) + \frac{1}{\delta} \int_{0}^{t} \nabla y \psi_l \cdot (\sigma dW^i_s) - \frac{1}{\delta^2} \int_{0}^{t} f_l h_{x_l} ds
\]

\[
+ \frac{1}{N} \sum_{j=1}^{N} \left[ \int_{0}^{t} \partial_{\mu} \psi_l(\bar{X}^i_{s}, \bar{X}^i_{s}/\delta, \bar{\mu}^N_s)(\bar{X}^j_{s}) \cdot \right]
\]

\[
\left( \frac{1}{\delta} f(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) + b(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s) + \sigma(\bar{X}^j_{s}, \bar{X}^j_{s}/\delta, \bar{\mu}^N_s)u^N_j(s) \right)
\]
\[
\begin{align*}
&+ \frac{1}{2} A(\tilde{X}_t^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N) \cdot \partial_x \partial_y \psi_I(\tilde{X}_s^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)(\tilde{X}_s^{j,N}) \\
&+ \frac{1}{2N} A(\tilde{X}_t^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N) \cdot \partial^2_x \psi_I(\tilde{X}_s^{i,j,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)(\tilde{X}_s^{j,N}, \tilde{X}_s^{j,N}) ds \\
&+ \int_0^t \partial_x \psi_I(\tilde{X}_s^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)(\tilde{X}_s^{j,N}) \cdot (\sigma(\tilde{X}_s^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N) dW_s^{j}) \\
&+ \frac{1}{N} \int_0^t A : \nabla_x \partial_y \psi_I(\tilde{X}_s^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)(\tilde{X}_s^{j,N}) ds + \frac{1}{N \delta} \int_0^t A : \nabla_y \partial_x \psi_I(\tilde{X}_s^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)(\tilde{X}_s^{j,N}) ds,
\end{align*}
\]

where in all coefficients where the argument is suppressed, the argument is \((\tilde{X}_t^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N)\). Solving for \(h(x_0, t)\) and plugging into our representation for \(h\), we get

\[(27)\]

\[
h(\tilde{X}_t^{i,N}) = h(x_0, t) + \sum_{k=1}^8 B_k^{i,N}(t)
\]

where

\[
B_1^{i,N}(t) = \int_0^t b \cdot \nabla h + \frac{1}{2} A : \nabla \nabla h + \sum_{l=1}^d \left\{ [f + \delta b] \cdot [\nabla_x \Phi_l h_{x_l} + \Phi_l \nabla_x h_{x_l}] \\
+ \frac{\delta}{2} A : [h_{x_l} \nabla_x \nabla_x \Phi_l + \nabla_x \nabla_x h_{x_l} \otimes \nabla_x \Phi_l] + b \cdot [h_{x_l} \nabla_y \Phi_l] \\
+ A : [h_{x_l} \nabla_x \nabla_y \Phi_l + \nabla_x h_{x_l} \otimes \nabla_y \Phi_l] \right\} ds
\]

\[
B_2^{i,N}(t) = \int_0^t \sigma u_t^{i,N} \cdot \nabla h + \sum_{l=1}^d \left\{ \delta[\sigma u_t^{i,N}] \cdot [\nabla_x \Phi_l h_{x_l} + \Phi_l \nabla_x h_{x_l}] + [\sigma u_t^{i,N}] \cdot [h_{x_l} \nabla_y \Phi_l] \right\} ds
\]

\[
B_3^{i,N}(t) = \int_0^t \nabla h \cdot (\sigma dW_s^i) + \int_0^t \sum_{l=1}^d \left\{ [\delta[\nabla_x \Phi_l h_{x_l} + \Phi_l \nabla_x h_{x_l}] + [h_{x_l} \nabla_y \Phi_l] \cdot (\sigma dW_s^i) \right\}
\]

\[
B_4^{i,N}(t) = \delta \sum_{l=1}^d \left\{ h_{x_l}(x_0^{i,N}) \Phi_l(x_0^{i,N}, x_0^{j,N}/\delta, \tilde{\mu}_0^N) - h_{x_l}(\tilde{X}_t^{i,N}) \Phi_l(\tilde{X}_t^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N) \right\}
\]

\[
= \delta \left\{ \nabla h(x_0^{i,N}) \cdot \Phi(x_0^{i,N}, x_0^{j,N}/\delta, \tilde{\mu}_0^N) - \nabla h(\tilde{X}_t^{i,N}) \cdot \Phi(\tilde{X}_t^{i,N}, \tilde{X}_s^{j,N}/\delta, \tilde{\mu}_s^N) \right\}
\]
\[ B_{5}^{i,N}(t) = \frac{\delta}{N} \sum_{l=1}^{d} \sum_{j=1}^{N} \left\{ \int_{0}^{t} [h_{x_{i}} \partial_{\mu} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N}(\bar{X}_{s}^{i,N})] \cdot \left( \frac{1}{\delta} f(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N}) + b(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N}) \right) + \frac{1}{2} A(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N}) : [h_{x_{i}} \partial_{\mu} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N})] \cdot ds \right\} + \frac{1}{2N} A(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N}) : [h_{x_{i}} \partial_{\mu}^{2} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N})] ds \right\} \]

\[ B_{6}^{i,N}(t) = \frac{\delta}{N} \sum_{l=1}^{d} \sum_{j=1}^{N} \left\{ \int_{0}^{t} \partial_{\mu} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N}) \cdot \left( \sigma(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})u_{j}^{N}(s) \right) ds \right\} \]

\[ B_{7}^{i,N}(t) = \frac{1}{N} \sum_{l=1}^{d} \left\{ \delta \int_{0}^{t} A : [\nabla_{x_{i}} h_{x_{i}} \otimes \partial_{\mu} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N})] \cdot ds \right\} + \frac{1}{N} \int_{0}^{t} A : [h_{x_{i}} \nabla_{x_{i}} \partial_{\mu} \Phi_{l}(\bar{X}_{s}^{i,N}, \bar{X}_{s}^{i,N}/\delta, \bar{\mu}_{s}^{N})(\bar{X}_{s}^{i,N})] ds \right\} \]

Thus we have

\[ E \left[ q(g(\mu^{N}(t + \tau)), g(\mu^{N}(t))) \middle| F_{t}^{N} \right] \leq E \left[ \frac{1}{N} \sum_{k=1}^{8} \sum_{i=1}^{N} B_{k}^{i,N}(t + \tau) - B_{k}^{i,N}(t) \middle| F_{t}^{N} \right]. \]

Firstly, we observe that

\[ \left| \frac{1}{N} \sum_{i=1}^{N} B_{4}^{i,N}(t + \tau) - B_{4}^{i,N}(t) \right| \leq 2\delta \| \nabla h \| \leq \frac{1}{N} \sum_{i=1}^{N} \sup_{t \in [0,1]} | \Phi(\bar{X}_{t}^{i,N}, \bar{X}_{t}^{i,N}/\delta, \bar{\mu}_{t}^{N}) | \leq \delta C \]

by Proposition C.2.
Next, we observe that

\[
\left| \frac{1}{N} \sum_{i=1}^{N} B_2^{i,N}(t + \tau) - B_2^{i,N}(t) \right|
\]

\[
\leq c_1(m, d) \| \nabla h \|_\infty \frac{1}{N} \sum_{i=1}^{N} \left( \int_{t}^{t+\tau} \| [I + \delta \nabla_x \Phi + \nabla_y \Phi] \sigma \|^2 ds \int_{0}^{1} |u_i^N(s)|^2 ds \right)^{1/2}
\]

\[
+ c_2(m, d) \| \nabla \nabla h \|_\infty \frac{\delta}{N} \sum_{i=1}^{N} \left( \int_{t}^{t+\tau} |\Phi|^2 \| \sigma \|^2 ds \int_{0}^{1} |u_i^N(s)|^2 ds \right)^{1/2}
\]

\[
\leq c_3(1 + \delta) \tau^{1/2} \frac{1}{N} \sum_{i=1}^{N} \left( \int_{0}^{1} |u_i^N(s)|^2 ds \right)^{1/2}
\]

by Hölder’s inequality, monotonicity, Assumption (A1), Proposition C.2, and Jensen’s inequality. Similarly, we have

\[
\left| \frac{1}{N} \sum_{i=1}^{N} B_6^{i,N}(t + \tau) - B_6^{i,N}(t) \right|
\]

\[
\leq c_4(d, m) \| \nabla h \|_2 \frac{\delta}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( \int_{t}^{t+\tau} |\partial_x \Phi(X_s^{i,N}, X_s^{i,N}/\delta, \bar{\mu}_s^N)(\bar{X}_s^{j,N}) \sigma(\bar{X}_s^{j,N}, \bar{X}_s^{j,N}/\delta, \bar{\mu}_s^N)|^2 ds \right)^{1/2}
\]

\[
\times \left( \int_{0}^{1} |u_j^N(s)|^2 ds \right)^{1/2}
\]

\[
\leq c_5(d, m) \| \nabla h \|_2 \frac{\delta}{N} \sum_{i=1}^{N} \left( \frac{1}{N} \sum_{j=1}^{N} \int_{t}^{t+\tau} |\partial_x \Phi(X_s^{i,N}, X_s^{i,N}/\delta, \bar{\mu}_s^N)(\bar{X}_s^{j,N})|^2 ds \right)^{1/2} \left( \frac{1}{N} \sum_{j=1}^{N} \int_{0}^{1} |u_j^N(s)|^2 ds \right)^{1/2}
\]

\[
\leq c_6 \delta \tau^{1/2} \left( \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{1} |u_i^N(s)|^2 ds \right)^{1/2}
\]

In addition,

\[
\mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^{N} \left( B_3^{i,N}(t + \tau) + B_7^{i,N}(t + \tau) - B_3^{i,N}(t) - B_7^{i,N}(t) \right) \right| \mathcal{F}_t^N \right]
\]

\[
\leq \left( \mathbb{E} \left[ \left| \frac{1}{N} \sum_{i=1}^{N} B_3^{i,N}(t + \tau) + B_7^{i,N}(t + \tau) - B_3^{i,N}(t) - B_7^{i,N}(t) \right|^2 \mathcal{F}_t^N \right] \right)^{1/2}
\]

by Jensen’s inequality

\[
\leq \tau^{1/2} c_7 \left( \frac{1}{N} + \frac{\delta}{N} \right)
\]

by Itô Isometry, Assumption (A1) and Proposition C.2.
Lastly, we observe that
\[
\left| \frac{1}{N} \sum_{i=1}^{N} \left( B_{1,N}^i (t + \tau) + B_{5,N}^i (t + \tau) + B_{8,N}^i (t + \tau) - B_{1,N}^i (t) - B_{5,N}^i (t) - B_{8,N}^i (t) \right) \right|
\leq \frac{1}{N} \sum_{i=1}^{N} \left\{ \| \nabla h \|_\infty \int_t^{t+\tau} \left| I + \delta \nabla_x \Phi + \nabla_y \Phi \right| b + \nabla_x \Phi \right. \\
+ A : \nabla_x \nabla_y \Phi + \frac{\delta}{2} \nabla_x \nabla_x \Phi + \frac{1}{N} \nabla_y \partial_x \Phi(\bar{X}_{s_{1,n}} - X_{s_{1,n}} - \delta, \tilde{\mu}_s^N)(\bar{X}_{s_{1,n}}) \\
+ \frac{\delta}{N} \nabla_x \partial_x \Phi(X_{s_{1,n}}^N, \bar{X}_{s_{1,n}}^N / \delta, \tilde{\mu}_s^N)(\bar{X}_{s_{1,n}}) \\
+ \left\{ \int_{\mathbb{R}^d} \partial_x \Phi(X_{s_{1,n}}^N, \bar{X}_{s_{1,n}}^N / \delta, \tilde{\mu}_s^N)(v) \left[ f(v, v / \delta, \tilde{\mu}_s^N) + \delta b(v, v / \delta, \tilde{\mu}_s^N) \right] \\
+ \frac{1}{2} A(v, v / \delta, \tilde{\mu}_s^N) \right\} \\
\left. : \delta \partial_x \partial_x \Phi(X_{s_{1,n}}^N, \bar{X}_{s_{1,n}}^N / \delta, \tilde{\mu}_s^N)(v) + \frac{\delta}{N} \partial_x \Phi(X_{s_{1,n}}^N, \bar{X}_{s_{1,n}}^N / \delta, \tilde{\mu}_s^N)(\bar{X}_{s_{1,n}}) \right] \left\| ds \right. \\
+ \| \nabla \nabla h \|_\infty \int_t^{t+\tau} \left[ \frac{1}{2} + \delta \nabla_x \Phi + \frac{\delta}{N} \partial_x \Phi(X_{s_{1,n}}^N, \bar{X}_{s_{1,n}}^N / \delta, \tilde{\mu}_s^N)(\bar{X}_{s_{1,n}}) \right] A + [f + \delta b] \otimes \Phi \right| ds \\
+ \delta c_{s_{1,m}} (d, m) \sup_{t \in \{1, \ldots, d\}} \left| \nabla \nabla h_{s_{1,t}} \right| \int_t^{t+\tau} | \Phi | | A | ds \right\} \\
\leq c_0 (1 + \delta + \frac{1}{N} + \frac{\delta}{N}) \tau
\]
by Assumption [A1] and Proposition [C.2]. So, letting
\[
\xi^N(\gamma) = C \left( \delta + (1 + \delta) \gamma^{1/2} \left( \frac{1}{N} \sum_{i=1}^{N} \int_0^{1} |u_i^N(s)|^2 ds \right)^{1/2} + (1 + \delta + \frac{1}{N} + \frac{\delta}{N}) \gamma + \left( \frac{1}{N} + \frac{\delta}{N} \right) \gamma^{1/2} \right)
\]
for \( C \) large enough to dominate all the constants in the previous bounds, we see that for each \( g \in \mathcal{G} \),
\[
\mathbb{E} \left[ q \left( g(\tilde{\mu}_N(t + \tau)), g(\tilde{\mu}_N(t)) \right) | \mathcal{F}_t^N \right] \leq \mathbb{E} \left[ \xi^N(\gamma) | \mathcal{F}_t^N \right].
\]
Now we must show that
\[
\lim_{\gamma \to 0} \limsup_{N \to \infty} \mathbb{E}[\xi^N(\gamma)] = 0.
\]
This follows immediately since
\[
\mathbb{E} \left[ \xi^N(\gamma) \right] \leq C \left( \delta + (1 + \delta) \gamma^{1/2} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^{N} \int_0^{1} |u_i^N(s)|^2 ds \right]^{1/2} + (1 + \delta + \frac{1}{N} + \frac{\delta}{N}) \gamma + \left( \frac{1}{N} + \frac{\delta}{N} \right) \gamma^{1/2} \right)
\]
by Jensen’s inequality
\[
\leq C \left( \delta + (1 + \delta) \gamma^{1/2} B^{1/2} + (1 + \delta + \frac{1}{N} + \frac{\delta}{N}) \gamma + \left( \frac{1}{N} + \frac{\delta}{N} \right) \gamma^{1/2} \right)
\]
by Assumption [C.2], so
\[
\limsup_{N \to \infty} \mathbb{E} \left[ \xi^N(\gamma) \right] \leq C \left( B^{1/2} \gamma^{1/2} + \gamma \right)
\]
and
\[
\lim_{\gamma \to 0} \limsup_{N \to \infty} \mathbb{E} \left[ \xi^N(\gamma) \right] \leq \lim_{\gamma \to 0} \left( C \left( B^{1/2} \gamma^{1/2} + \gamma \right) \right) = 0.
\]
6.2. Identification of the Limit. Extract a convergent subsequence from \(\{Q^N\}_{N \in \mathbb{N}}\) and relabel with new indexes so that \(\{Q^N\}_{N \in \mathbb{N}}\) converges to some \(Q\) weakly as a \(\mathcal{P}(\mathcal{W})\)-valued random variable. Let \((\Omega, \bar{\mathcal{F}}, \bar{\mathbb{P}})\) be the probability space on which \(Q\) lies.

We wish to identify the limit \(Q\) as a member of \(\mathcal{V}\) for \(\bar{\mathbb{P}}\) almost every \(\omega \in \bar{\Omega}\). Our main tool here is the associated martingale problem to weak solutions of \((11)\). An important element to the proof which is special to the joint limit as \(N \to \infty, \delta \downarrow 0\) is that we must first show that \((V4)\) holds before identifying the SDE associated to \(Q\) to prove \((V1)\). This is because, as proven in \((41)\), in the prelimit there exists a term which is \(O(1)\) in \(N\), but is in fact 0 in the limit due to the centering condition \((A8)\). As the centering condition is a statement involving the invariant measure \(\pi\) (see Equation \(3\)), it is necessary to the proof that we have already identified the \(\mathcal{Y}\) component of the limiting Coordinate Process \((13)\) as being concentrated on \(\pi\).

6.2.1. Proof of \((V4)\). As stated, we offer the proof of \((V4)\) first.

We want to show that for almost every \(\omega \in \bar{\Omega}\) and \(\forall t \in [0, 1], g \in C^2_b(\mathbb{R}^d)\),

\[
E^Q \left[ \int_{\mathbb{T}^d \times [0, t]} \mathcal{L}^1_{X^i, \nu_{Q(t)}} g(y) m(dyds) \right] = 0.
\]

Let \(g_t : \mathbb{R}^d \to \mathbb{R}, t \in \mathbb{N}\) be smooth and bounded with bounded derivatives and dense in \(C^2_b(\mathbb{R}^d)\). This set exists by using Stone–Weierstrass and taking rational coefficients. Let \(\bar{Y}^{i,N} = X^{i,N}/\delta\). Considering the operator which acts on \(g \in C^2_b(\mathbb{R}^d)\) by

\[
A_{x,z,\mu}[g](y) := \left[ \frac{1}{\delta^2} f(x, y, \mu) + \frac{1}{\delta} (b(x, y, \mu) + \sigma(x, y, \mu) z) \right] \cdot \nabla g(y) + \frac{1}{2\delta^2} A(x, y, \mu) : \nabla \nabla g(y).
\]

Note that by \((A1)\) for \(t \in [0, 1]\) and fixed \(N \in \mathbb{N}\),

\[
M_t^{i,N} := g_t(\bar{Y}^{i,N}_t) - g_t(x^{i,N}_0/\delta) - \int_0^t A_{X^{i,N}, u_{i}^{N}(s), \mu_{i}^{N}}[g_t](\bar{Y}^{i,N}_s)ds
\]

is an \(\mathcal{F}_t\)-martingale. By definition, for \(t \in [0, 1]\),

\[
\int_0^t A_{X^{i,N}, u_{i}^{N}(s), \mu_{i}^{N}} g_t(\bar{Y}^{i,N}_s)ds = \frac{1}{\delta^2} \int_0^t \mathcal{L}^1_{X^{i,N}, \mu_{i}^{N}} g_t(\bar{Y}^{i,N}_s)ds
\]

\[
+ \frac{1}{\delta} \int_0^t \left[ b(X^{i,N}_s, \bar{Y}^{i,N}_s, \mu_{i}^{N}) + \sigma(X^{i,N}_s, \bar{Y}^{i,N}_s, \mu_{i}^{N}) u_i^{N}(s) \right] \cdot \nabla g_t(\bar{Y}^{i,N}_s)ds
\]

Consider now the operator which acts on \(g \in C^2_b(\mathbb{R}^d)\) by

\[
B_{x,z,\mu}[g](y) := \left[ b(x, y, \mu) + \sigma(x, y, \mu) z \right] \cdot \nabla g(y).
\]

Then

\[
\frac{1}{N} \sum_{i=1}^N \left| \frac{\delta^2}{2} \left( -M_t^{i,N} + g_t(\bar{Y}^{i,N}_t) - g_t(x^{i,N}_0/\delta) \right) - \delta \int_0^t B_{X^{i,N}, u_{i}^{N}(s), \mu_{i}^{N}}[g_t](\bar{Y}^{i,N}_s)ds \right|
\]

\[
= \frac{1}{N} \sum_{i=1}^N \int_0^t \mathcal{L}^1_{X^{i,N}, \mu_{i}^{N}} g_t(\bar{Y}^{i,N}_s)ds.
\]

We will show the right hand side of Equation \(28\) converges in distribution to \(E^Q \left[ \int_{\mathbb{T}^d \times [0, t]} \mathcal{L}^1_{X^{i,N}, \nu_{Q(t)}} f(y) m(dyds) \right] \)

and the left hand side converges in distribution to 0, so by a density argument the result holds.

The proof that the right hand side of Equation \(28\) converges in distribution follows from the observation that

\[
\frac{1}{N} \sum_{i=1}^N \int_0^t \mathcal{L}^1_{X^{i,N}, \mu_{i}^{N}} g_t(\bar{Y}^{i,N}_s)ds = \int_{\mathbb{W}} \int_{\mathbb{T}^d \times [0, t]} \mathcal{L}^1_{\phi(s), v_{Q(s)}} g_t(y) m(dyds) Q^N(d\phi d\nu d\tau).
\]
We invoke Skorohod’s representation theorem (Theorem 3.1.8 in [22]) to assume the convergence of $Q^N \to Q$ holds with probability 1. Without making a distinction between the original probability space in the new one, we will prove

\begin{equation}
\mathbb{E}\left[ \left| \frac{1}{N} \sum_{i=1}^{N} \int_0^t \mathcal{L}^1_{X^i_{N},\nu_Q(s)} g_i(Y_{i,N}^s) ds \right| - \mathbb{E}^Q \left[ \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{X_{\nu_Q(s)},\nu_Q(s)} g_i(y) m(dyds) \right] \right] \to 0,
\end{equation}

so that by Chebyshev’s inequality the convergence holds in probability and hence in distribution.

First we prove that

\begin{equation}
\mathbb{E}\left[ \left| \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr)
- \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr) \right] \to 0.
\end{equation}

We have

\begin{align*}
\lim_{N \to \infty} \mathbb{E}\left[ \left| \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr)
- \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr) \right] \\
\leq \sup_{N \in \mathbb{N}} \lim_{J \to \infty} \mathbb{E}\left[ \left| \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr) \right] \\
\leq \sup_{N \in \mathbb{N}} \lim_{J \to \infty} \mathbb{E}\left[ \left| \int_{\mathbb{T}^d \times [0,t]} \left( \mathcal{L}^1_{\phi(s),\nu_Q(s)} - \mathcal{L}^1_{\phi(s),\nu_Q(s)} \right) g_i(y) n(dyds) \right| Q^N(d\phi dndr) \right].
\end{align*}

By Proposition A.3 and Assumption [A1], we get for fixed $\phi, n, s$ that

\begin{equation}
\lim_{J \to \infty} \int_{\mathbb{T}^d} \left( \mathcal{L}^1_{\phi(s),\nu_Q(s)} - \mathcal{L}^1_{\phi(s),\nu_Q(s)} \right) g_i(y) n_s(dy) = 0.
\end{equation}

By Assumption [A1], the bounded convergence theorem applies. So we can pass the limit through the integrals, and the results follows.

Now we will show

\begin{equation}
\mathbb{E}\left[ \left| \frac{1}{N} \sum_{i=1}^{N} \int_0^t \mathcal{L}^1_{X^i_{N},\nu_Q(s)} g_i(Y_{i,N}^s) ds \right| - \mathbb{E}^Q \left[ \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{X_{\nu_Q(s)},\nu_Q(s)} g_i(y) m(dyds) \right] \right] \to 0,
\end{equation}

which together with triangle inequality and Equation 30 proves Equation 29.

Noting that

\begin{align*}
\mathbb{E}\left[ \left| \frac{1}{N} \sum_{i=1}^{N} \int_0^t \mathcal{L}^1_{X^i_{N},\nu_Q(s)} g_i(Y_{i,N}^s) ds \right| - \mathbb{E}^Q \left[ \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{X_{\nu_Q(s)},\nu_Q(s)} g_i(y) m(dyds) \right] \right] \\
= \mathbb{E}\left[ \left| \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q^N(d\phi dndr) \\
- \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| Q(d\phi dndr) \right],
\end{align*}

so since $(\phi, n) \mapsto \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds)$ is continuous from $\mathcal{X} \times \mathcal{Y}$ to $\mathbb{R}$, and since we have that

\[ \sup_{(\phi, n) \in \mathcal{X} \times \mathcal{Y}} \left| \int_{\mathbb{T}^d \times [0,t]} \mathcal{L}^1_{\phi(s),\nu_Q(s)} g_i(y) n(dyds) \right| < \infty \text{ by assumption [A1]}, \]

the result follows immediately from the convergence of $Q^N \to Q$. 

To prove the left hand side of (28) converges to zero in distribution, we will show that

\[ E \left[ \frac{1}{N} \sum_{i=1}^{N} \delta^2 \left( -M_{t,N}^i + g_i(Y_{t,N}^i) - g_i(x_{0,N}^i/\delta) \right) - \delta \int_0^t \mathbb{B}_{\bar{X}_{s,N}^i, \bar{Y}_{s,N}^i} [g_i(Y_{s,N}^i)] ds \right] \rightarrow 0 \]

as \( N \rightarrow \infty \), so the result will follow by Chebyshev’s inequality.

We first note that

\[ \delta^2 \frac{1}{N} \sum_{i=1}^{N} E \left[ \left| M_{t,N}^i \right|^2 \right] \leq \delta^2 \frac{1}{N} \sum_{i=1}^{N} E \left[ \left( M_{t,N}^i \right)^2 \right]^{1/2} = 2 \delta \|

Also,

\[ \delta^2 \frac{1}{N} \sum_{i=1}^{N} E \left[ \left| g_i(Y_{t,N}^i) - g_i(x_{0,N}^i/\delta) \right| \right] \leq 2 \delta \|

Lastly,

\[ \delta \frac{1}{N} \sum_{i=1}^{N} E \left[ \left| \int_0^t \mathbb{B}_{\bar{X}_{s,N}^i, \bar{Y}_{s,N}^i} [g_i(Y_{s,N}^i)] ds \right| \right] \leq \delta \| \nabla g_i \| C(B) \]

by Hölder’s inequality, Assumption (25), and Assumption (A1).

Now we have that for each \( g_i \) and \( t \in [0,1] \), there exists a set \( N_{g_i,t} \) such that \( \hat{P}(N_{g_i,t}) = 0 \) and

\[ E_{\hat{\omega}} \left[ \int_{T^d \times [0,t]} \mathcal{L}_{X_{s},Y_{s}}^1 g(y) m(dyds) \right] = 0, \forall \hat{\omega} \in \hat{\Omega} \setminus N_{g_i,t}. \]

Taking a countable dense set \( D \subset [0,1] \) and letting \( N = \bigcup_{t \in D} N_{g_i,t} \), we have \( \hat{P}(N) = 0 \) and

\[ E_{\hat{\omega}} \left[ \int_{T^d \times [0,t]} \mathcal{L}_{X_{s},Y_{s}}^1 g(y) m(dyds) \right] = 0, \forall \hat{\omega} \in \hat{\Omega} \setminus N, \forall g \in C_b^2(T^d). \]

So (V4) holds for almost every \( \hat{\omega} \in \hat{\Omega} \). In Remark 4.7 we discuss implications of (V4).

Remark 6.7. \( \Theta \in \mathcal{P}(\mathcal{W}) \) satisfying condition (V4) implies that \( \forall t \in [0,1], g \in C_b^2(T^d), \)

\[ E_{\Theta} \left[ \int_{T^d \times [0,t]} \mathcal{L}_{X_{s},Y_{s}}^1 g(y) m(dyds) \right] = \int_{T^d} \int_{[0,t]} \mathcal{L}_{\phi(s),\nu_{\phi}(s)}^1 g(y) n(dyds) \Theta(d\phi d\nu) / \Theta(d\phi d\nu) = 0. \]

Note that we can write \( \Theta|_{B(X \times Y)}(d\phi d\nu) = \lambda(d\nu|\phi)\Theta_{X}(d\phi) \). Then (V4) can be written as:

\[ \forall t \in [0,1], g \in C_b^2(T^d), \int_{X} \int_{Y} \left| \mathcal{L}_{\phi(s),\nu_{\phi}(s)}^1 g(y) n(dyds) \right| \lambda(d\nu|\phi) \Theta_{X}(d\phi) = 0. \]

So

\[ \Rightarrow \text{for } \Theta_{X} \text{-a.e. } \phi \in X, \quad \lambda(\{ n \in Y : \mathcal{L}_{\phi(s),\nu_{\phi}(s)}^1 n_s = 0, \forall s \in [0,1] \})|\phi) = 1. \]
Lemma 6.10. The invariant measure associated with \( L^1_{\varphi(s), \nu_\Theta(s)} \) is unique for each \( \varphi, \Theta, \) and \( s \), so \( \lambda(\phi) \) is concentrated on the single measure \( \pi(dy|\phi(s), \nu_\Theta(s)) \otimes ds \) for \( \Theta_\lambda \)-a.e. \( \phi \). Thus we have

\[
\Theta \left( \left\{ (\phi, n, r) \in W : n_s(dy) = \pi(dy|\phi(s), \nu_\Theta(s)), \forall s \in [0,1] \right\} \right)
\]

so that the set we are measuring doesn’t depend on \( r \)

\[
\int A \left( \left\{ (\phi, n) \in X \times Y : n_s(dy) = \pi(dy|\phi(s), \nu_\Theta(s)), \forall s \in [0,1] \right\} \right) \Theta_X(d\phi)
\]

Since the measure \( Q \) is concentrated on the single measure \( \pi(dy|\phi(s), \nu_\Theta(s)) \otimes ds \) for \( \Theta_\lambda \)-a.e. \( \phi \), it suffices to show that for fixed \( h \)

\[
\omega \in \tilde{\Omega}.
\]

We proceed with the proof of Lemma 6.9. We note that it is enough to prove Theorem 6.8 for \( h \) such that \( \omega \in \tilde{\Omega} \).

Proof of (V1). We wish to prove that \( Q_{\omega} \) corresponds to \( \tilde{X} \), a weak solution of Equation (12) for \( \tilde{\Psi} \)-a.e. \( \tilde{\omega} \in \tilde{\Omega} \).

Given \( g \in C_b^2(\mathbb{R}^d) \) and \( \Theta \in \mathcal{P}(W) \), define a real-valued process \( \{ M_{\varphi}^\Theta(t) \}_{t \in [0,1]} \) on \( (W, \mathcal{B}(W), \Theta) \) given by

\[
M_{\varphi}^\Theta(t, (\phi, n, r)) = g(\phi(t)) - g(\phi(0)) - \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{T}^d} A[g](\phi(s), y, z, \nu_\Theta(s)) n_s(dy) r_s(dz) ds
\]

where

\[
A[g](x, y, z, \nu_\Theta(s)) := \left[ \nabla_y \Phi(x, y, \nu_\Theta(s)) + I |b(x, y, \nu_\Theta(s)) + \sigma(x, y, \nu_\Theta(s))z| + \nabla_x \Phi(x, y, \nu_\Theta(s)) f(x, y, \nu_\Theta(s)) + A : \nabla_x \nabla_y \Phi(x, y, \nu_\Theta(s)) \right] : \nabla g(x)
\]

\[
+ \left[ \nabla_y \Phi(x, y, \nu_\Theta(s)) + \frac{1}{2} A_3(x, y, \nu_\Theta(s)) + f(x, y, \nu_\Theta(s)) \otimes \Phi(x, y, \nu_\Theta(s)) \right] : \nabla \nabla g(x).
\]

We will say \( \Theta \in \mathcal{P}(W) \) solves the martingale problem associated to \( \tilde{X}_{\nu_\Theta} \) with initial distribution \( \nu_0 \) if for all \( 0 \leq s \leq t \leq 1 \) and \( g \in C_b^2(\mathbb{R}^d) \),

\[
\mathbb{E}^\Theta \left[ M_{\varphi}^\Theta(t) | G_s \right] = M_{\varphi}^\Theta(s), \quad \nu_\Theta(0) = \nu_0
\]

for \( \{ G_t \}_{t \in [0,1]} \) the canonical filtration on the coordinate process as defined in Equation (13). To identify the limit \( \tilde{X} \) as a weak solution to (10) by the density argument offered at the end of this subsection, it suffices to show that for fixed \( h \in C_c^\infty(\mathbb{R}^d) \), \( 0 \leq s \leq t \leq 1 \), and \( G_s \)-measurable \( \Psi \in C_b(W) \) that

**Theorem 6.8.**

\[
\mathbb{E}^{Q_{\omega}} \left[ \Psi(M_{\varphi}^{Q_{\omega}}(t) - M_{\varphi}^{Q_{\omega}}(s)) \right] = 0
\]

for almost every \( \tilde{\omega} \in \tilde{\Omega} \).

We note that it is enough to prove Theorem 6.8 for \( h \in C_c^\infty(\mathbb{R}^d) \), due to the fact that \( C_c^\infty(\mathbb{R}^d) \) is separating in the sense of Chapter 3 Section 4 in [22] (see [14] Definition 3.1 and Chapter 4 Section 8 in [22]). In order prove Theorem 6.8 we will prove Lemma 6.9 and Lemma 6.10.

**Lemma 6.9.**

\[
\mathbb{E}^{Q_N} \left[ \Psi(M_{\varphi}^{Q_N}(t) - M_{\varphi}^{Q_N}(s)) \right] \to \mathbb{E}^{Q} \left[ \Psi(M_{\varphi}^{Q}(t) - M_{\varphi}^{Q}(s)) \right] \quad \text{in distribution},
\]

**Lemma 6.10.**

\[
\mathbb{E}^{Q_N} \left[ \Psi(M_{\varphi}^{Q_N}(t) - M_{\varphi}^{Q_N}(s)) \right] \to 0 \quad \text{in distribution}.
\]

Then the conclusion follows.

We proceed with the proof of Lemma 6.9.
Proof of Lemma 6.9: Unpacking the notation in Equation 35, we see what we are trying to show is that

$$\int \Psi \left[ g(\phi(t)) - g(\phi(s)) - \int_{s}^{t} \int_{\mathbb{R}^{d}} A[h](\phi(\tau), y, z, \nu_{Q,N}(\tau))n_{\tau}(dy)r_{\tau}(dz)d\tau \right] Q^{N}(d\phi d\nu)$$

$$\rightarrow \int \Psi \left[ g(\phi(t)) - g(\phi(s)) - \int_{s}^{t} \int_{\mathbb{R}^{d}} A[h](\phi(\tau), y, z, \nu_{Q}(\tau))n_{\tau}(dy)r_{\tau}(dz)d\tau \right] Q(d\phi d\nu)$$

in distribution. We invoke Skorokhod’s representation theorem to assume the convergence of $Q^{N} \rightarrow Q$ occurs with probability 1, without making a distinction in the notation between the new probability space and the original one. Then we see Lemma 6.9 essentially follows from the definition of convergence of measures in the space $\mathcal{P}(\mathbb{W})$. The only caveats are that the integrand is also a function of the converging measures through its dependence on $\nu_{Q,N}$, and that the integrand is not apriori in $C_{b}(\mathbb{W})$, since it grows linearly in the control. Thus, we will show

$$E \left[ E^{Q^{N}} \left[ \sup_{t \in [0,1]} \left| M_{h}^{Q,N}(t) - M_{h}^{Q}(t) \right| \right] \right] \rightarrow 0$$

and

$$E \left[ E^{Q^{N}} \left[ \Psi(M_{h}^{Q,N}(t) - M_{h}^{Q}(t)) \right] - E^{Q} \left[ \Psi(M_{h}^{Q}(t) - M_{h}^{Q}(s)) \right] \right] \rightarrow 0.$$ 

Once these limits are established, by triangle inequality and Chebyshev’s inequality Lemma 6.9 will be proved.

To see Equation 37 holds we write

$$\sum_{i=1}^{n} \left| A^{1}(x, y, \nu_{\Theta}(s)) + A^{2}(x, y, \nu_{\Theta}(s))z \right| \cdot \nabla h(x) + A^{3}(x, y, \nu_{\Theta}(s)) : \nabla \nabla h(x)$$

where for $s \in [0,1], x \in \mathbb{R}^{d}, y \in \mathbb{T}^{d}$, and $\Theta \in \mathcal{P}(\mathbb{W})$

$$A^{1}(x, y, \nu_{\Theta}(s)) := [\nabla_{y} \Phi(x, y, \nu_{\Theta}(s)) + f(x, y, \nu_{\Theta}(s)) + \nabla_{x} \Phi(x, y, \nu_{\Theta}(s))]$$

$$A^{2}(x, y, \nu_{\Theta}(s)) := [\nabla_{y} \Phi(x, y, \nu_{\Theta}(s)) + f(x, y, \nu_{\Theta}(s)) + A \cdot \nabla_{y} \Phi(x, y, \nu_{\Theta}(s)) + f(x, y, \nu_{\Theta}(s))]$$

$$A^{3}(x, y, \nu_{\Theta}(s)) := [\nabla_{y} \Phi(x, y, \nu_{\Theta}(s)) + f(x, y, \nu_{\Theta}(s)) + \nabla_{x} \Phi(x, y, \nu_{\Theta}(s))].$$

Propositions A.2 and C.2 along with Assumption [A1] imply for fixed $z \in \mathbb{R}^{m}, h \in C_{c}^{\infty}(\mathbb{R}^{m}), x \in \mathbb{R}^{d}, y \in \mathbb{T}^{d}$, and $s \in [0,1]$ that

$$\lim_{N \rightarrow \infty} \left[ A^{1}(x, y, \nu_{Q,N}(s)) + A^{2}(x, y, \nu_{Q,N}(s))z \right] \cdot \nabla h(x) + A^{3}(x, y, \nu_{Q,N}(s)) : \nabla \nabla h(x)$$

$$- \left[ A^{1}(x, y, \nu_{Q}(s)) + A^{2}(x, y, \nu_{Q}(s))z \right] \cdot \nabla h(x) - A^{3}(x, y, \nu_{Q}(s)) : \nabla \nabla h(x) = 0$$

Thus if we are able to prove that we can pass the limit into the integrals in Equation 37 then we will be done. We observe that for $\Theta = Q^{N}, N \in \mathbb{N}$ or $Q$,

$$\left| A[h](x, y, \nu_{\Theta}(s)) \right| \leq \| \nabla h \|_{\infty} \left| A^{1}(x, y, \nu_{\Theta}(s)) \right| + C \| \nabla h \|_{\infty} \left( \left| A^{2}(x, y, \nu_{\Theta}(s)) \right|^{2} + |z|^{2} \right)$$

$$+ \| \nabla \nabla h \|_{\infty} \left| A^{3}(x, y, \nu_{\Theta}(s)) \right|$$

$$\leq C(1 + |z|^{2})$$

by Young’s inequality, Assumption [A1] and Proposition C.2

Then

$$\sup_{N \in \mathbb{N}} E \left[ E^{Q^{N}} \left[ \int_{0}^{1} \int_{\mathbb{R}^{d}} \int_{\mathbb{T}^{d}} C(1 + |z|^{2})n_{i}(dy)r_{i}(dz)dt \right] \right]$$

$$= \sup_{N \in \mathbb{N}} C(1 + E \left[ \int_{0}^{1} \frac{1}{N} \sum_{i=1}^{N} |w_{i}(t)|^{2} dt \right]) < \infty$$

by Assumption 25.
We have then that

$$\lim_{N \to \infty} \mathbb{E} \left[ \mathbb{E}^{Q_N} \left[ \sup_{t \in [0,1]} \left| M_h^{Q_N}(t, \phi, n, r) - M_h^{Q}(t, \phi, n, r) \right| \right] \right]$$

$$\leq \lim_{N \to \infty} \mathbb{E} \left[ \mathbb{E}^{Q_N} \left[ \int_0^1 \int_{\mathbb{R}^4} |A[h] \left( \phi(s), y, z, \nu_{Q_N}(s) \right) - A[h] \left( \phi(s), y, z, \nu_Q(s) \right) | n_s(dy) r_s(dz) ds \right] \right]$$

$$\leq \sup_{N \in \mathbb{N}} \lim_{j \to \infty} \mathbb{E} \left[ \mathbb{E}^{Q_N} \left[ \int_0^1 \int_{\mathbb{R}^4} \left| A[h](\phi(s), y, z, \nu_Q(s)) - A[h](\phi(s), y, z, \nu_{Q_N}(s)) \right| n_s(dy) r_s(dz) ds \right] \right]$$

$$= \sup_{N \in \mathbb{N}} \mathbb{E} \left[ \mathbb{E}^{Q_N} \left[ \int_0^1 \int_{\mathbb{R}^4} \lim_{j \to \infty} \left( A[h](\phi(s), y, z, \nu_Q(s)) - A[h](\phi(s), y, z, \nu_{Q_N}(s)) \right) n_s(dy) r_s(dz) ds \right] \right]$$

by Dominated Convergence Theorem

$$= 0 \text{ by Equation } 41$$

and hence Equation 37 is proved. Now we prove Equation 38. We have:

$$\mathbb{E} \left[ \left| \mathbb{E}^{Q_N} \left[ \Psi(M_h^{Q_N}(t) - M_h^{Q}(s)) \right] - \mathbb{E}^Q \left[ \Psi(M_h^{Q_N}(t) - M_h^{Q}(s)) \right] \right] \right]$$

$$= \mathbb{E} \left[ \mathbb{E}^{Q_N} \left[ h(\phi(t)) - h(\phi(s)) - \int_s^t \int_{\mathbb{R}^3} A[h](\phi(\tau), y, z, \nu_Q(\tau)) n_\tau(dy) r_\tau(dz) d\tau \right] \right]$$

Noting that $h(\phi(t)) - h(\phi(s))$ is bounded and writing $A[h]$ as in Equation 39, we see that only

$$\int_{\mathbb{R}^3} A^2(\phi(\tau), y, \nu_Q(s)) n_s(dy) r_s(dz)$$

exhibits growth in the control $r$. Since the desired convergence occurs immediately by boundedness of the integrand and almost-sure convergence of $Q^N \to Q$, we only show work to show the convergence of this term. Let

$$B^2(M) := \{ r \in \mathbb{Z} : \int_0^1 \int_{\mathbb{R}^m} |r| |dz| dt > M \},$$

$$\psi^1_M(r, t) := \begin{cases}
\int_{\mathbb{R}^m} zr_t(dz) \quad & \text{if } \int_0^1 \int_{\mathbb{R}^m} |z| |r| |dz| dt \leq M \\
\int_0^1 \int_{\mathbb{R}^m} |z| |r| |dz| dt 
\end{cases}$$

$$\psi^2_M(r, t) := \begin{cases}
0 \quad & \text{if } \int_0^1 \int_{\mathbb{R}^m} |z| |r| |dz| dt \leq M \\
\int_0^1 \int_{\mathbb{R}^m} zr_t(dz) \left( 1 - \frac{M}{\int_0^1 \int_{\mathbb{R}^m} |z| |r| |dz| dt} \right) 
\end{cases}$$

Then for all $r \in \mathbb{Z}, M > 0, t \in [0,1], \psi^1_M(r, t) + \psi^2_M(r, t) = \int_{\mathbb{R}^m} zr_t(dz), \int_0^1 |\psi^1_M(r, t)| dt \leq M,$ and

$$\int_0^1 |\psi^2_M(r, t)| dt < 1_{B^2(M)}(r) \int_0^1 \int_{\mathbb{R}^m} |z| |r| |dz| dt.$$
So
\[ E\left[ E^Q \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^2_M(r, \tau) n_\tau(dy) \right) d\tau \right] \]
\[ - E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) zn_\tau(dy) r_\tau(dz) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \]
\[ = E\left[ E^Q \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^2_M(r, \tau) n_\tau(dy) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \]
\[ + E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^2_M(r, \tau) n_\tau(dy) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \]
\[ - E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^1_M(r, \tau) n_\tau(dy) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \]
\[ - E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^1_M(r, \tau) n_\tau(dy) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \]
\[ \leq E\left[ E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^2_M(r, \tau) n_\tau(dy) \right) \cdot \nabla h(\phi(\tau)) d\tau \right] \right] \]
\[ + 2\|\nabla h\|_\infty \sup_{N \in \mathbb{N}} E\left[ E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} A^2(\phi(\tau), y, \nu_Q(\tau)) \psi^1_M(r, \tau) n_\tau(dy) \right) d\tau \right] \right] \]
by A.3.12 in [21].

By Bounded Convergence Theorem and convergence of $Q^N \to Q$, the first term vanishes as $N \to \infty$ (for continuity of the time integral of $\psi^2_M$ see Lemma 5.3.4/3.3.1 in [21]). To handle the second term, we have:
\[ \sup_{N \in \mathbb{N}} E\left[ E^Q \left[ \int_s^t \left( \int_{\mathbb{R}^d} \psi^2_M(r, \tau) n_\tau(dy) \right) d\tau \right] \right] \]
\[ \leq CE\left[ E^Q \left[ \int_0^1 \left| \psi^2_M(r, \tau) \right| d\tau \right] \right] \]
by Assumption (A1), Proposition C.2 and monotonicity.
\[ \leq CE\left[ E^Q \left[ \mathbb{1}_{B^2(M)}(r) \int_0^1 \int_{\mathbb{R}^d} |z|^2 r(dz) d\tau \right] \right] \]
\[ = C \sup_{N \in \mathbb{N}} E\left[ E^Q \left[ \mathbb{1}_{B^2(M)}(r) \frac{\int_0^1 \int_{\mathbb{R}^d} |z|^2 r(dz) d\tau}{\int_0^1 \int_{\mathbb{R}^d} |z|^2 r(dz) d\tau} \right] \right] \]
\[ \leq \frac{C}{M} \sup_{N \in \mathbb{N}} E\left[ E^Q \left[ \left( \int_0^1 \int_{\mathbb{R}^d} |z|^2 r(dz) d\tau \right)^2 \right] \right] \]
by the definition of $B^2(M)$.
\[ \leq \frac{C}{M} \sup_{N \in \mathbb{N}} E\left[ E^Q \left[ \int_0^1 \int_{\mathbb{R}^d} |z|^2 r(dz) d\tau \right] \right] \]
by Jensen’s inequality.
\[ = \frac{C}{M} \sup_{N \in \mathbb{N}} E\left[ \frac{1}{N} \sum_{i=1}^N \int_0^1 |u_i^N(\tau)|^2 d\tau \right] \]
by definition of $Q^N$.
\[ \leq \frac{CB}{M} \]
by Assumption (25).

Taking $N \to \infty$ then $M \to \infty$ the result follows.

Now we prove Lemma 6.10.

Proof of Lemma 6.10. Again, we invoke Skorohod’s representation theorem to assume the convergence of $Q^N \to Q$ occurs with probability 1.
We will show \( E \left[ \mathbb{E}^{Q^N} \left[ \Psi (M_h^{Q^N}(t) - M_h^{Q^N}(s)) \right] \right] \to 0 \) as \( N \to \infty \), and so the conclusion will follow via Chebyshev’s inequality.

From Equation (24) we get that

\[
\begin{align*}
&h(\bar{X}_t^{i,N}) - h(\bar{X}_s^{i,N}) = \int_s^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathcal{A}[h](\bar{X}_r^{i,N}, y, z, \mu_r^{i,N}) m^{i,N}_r(dy) \delta_{u^{i}_N}(\tau)(dz) d\tau + \sum_{k=1}^5 D_{k}^{i,N},
\end{align*}
\]

where

\[
D_{1}^{i,N} = \int_s^t \left[ \frac{1}{N} A : [\nabla_x \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(\bar{X}_r^{i,N})] 
+ \delta \left( \nabla_x \Phi b + \frac{1}{2} A : \nabla_x \nabla_x \Phi + \frac{1}{N} A : \nabla_x \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(\bar{X}_r^{i,N}) \right) 
+ \left\{ \int_{\mathbb{R}^d} \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(v) b(v, v/\delta, \bar{\mu}_r^{i,N}) + \frac{1}{2} A(v, v/\delta, \bar{\mu}_r^{i,N}) : \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(v, v) \right\} \right] 
\cdot \nabla h(\bar{X}_r^{i,N}) d\tau 
\right)

+ \delta \int_s^t \nabla_x \Phi A + b \otimes \Phi + \frac{1}{N} \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(\bar{X}_r^{i,N}) A : \nabla \nabla h(\bar{X}_r^{i,N}) d\tau 
\right)

+ \delta \int_s^t \sum_{i=1}^d \frac{1}{2} A : \nabla_x \nabla h_{x_i}(\bar{X}_r^{i,N}) \Phi_i d\tau 
\right)

D_{2}^{i,N} = \delta \int_s^t \left[ \nabla_x \Phi \sigma_{u^{i}_N}(\tau) + \frac{1}{N} \sum_{j=1}^N \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(\bar{X}_r^{i,N}) \sigma(\bar{X}_r^{j,N}, \bar{X}_r^{j,N}/\delta, \bar{\mu}_r^{j,N}) u^{i}_N(\tau) \right] 
\cdot \nabla h(\bar{X}_r^{i,N}) 
\right)

+ [\sigma_{u^{i}_N}(\tau)] \otimes \Phi : \nabla \nabla h(\bar{X}_r^{i,N}) d\tau 
\right)

D_{3}^{i,N} = \int_s^t \nabla h(\bar{X}_r^{i,N}) \cdot \left( [I + \delta \nabla_x \Phi + \nabla_y \Phi] \sigma dW_r^Y \right) + \int_s^t \delta \nabla \nabla h(\bar{X}_r^{i,N}) : \Phi \otimes (\sigma dW_r^Y) 
\right)

+ \int_s^t \nabla h(\bar{X}_r^{i,N}) \cdot \left[ \frac{\delta}{N} \sum_{j=1}^N \partial_{\mu} \Phi(\bar{X}_r^{j,N}, \bar{X}_r^{j,N}/\delta, \bar{\mu}_r^{j,N})(\bar{X}_r^{j,N}) \sigma(\bar{X}_r^{j,N}, \bar{X}_r^{j,N}/\delta, \bar{\mu}_r^{j,N}) dW_r^Y \right] 
\right)

D_{4}^{i,N} = \delta \left[ \nabla h(\bar{X}_s^{i,N}) \cdot \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N}) - \nabla h(\bar{X}_r^{i,N}) \cdot \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N}) \right] 
\right)

D_{5}^{i,N} = \int_s^t \left[ \frac{1}{N} \sum_{j=1}^N \partial_{\mu} \Phi(\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})(\bar{X}_r^{j,N}) f(\bar{X}_r^{j,N}, \bar{X}_r^{j,N}/\delta, \bar{\mu}_r^{j,N}) \right] \nabla h(\bar{X}_r^{i,N}) d\tau, 
\right)

m^{i,N} are defined as in Equation (24) and the arguments which are omitted are taken to be \((\bar{X}_r^{i,N}, \bar{X}_r^{i,N}/\delta, \bar{\mu}_r^{i,N})\).

Thus

\[
E^{Q^N} \left[ \Psi (M_h^{Q^N}(t) - M_h^{Q^N}(s)) \right] = \Psi \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^5 D_{k}^{i,N}. 
\]
By the same proofs as in tightness of $Q_X^N$ in Subsection 6.1.3 using Assumption (A1) and Proposition C.2, we get for large enough $N$ that
\[
\mathbb{E}\left[ \left| \frac{1}{N} \sum_{i=1}^{N} D_{k_i}^{i,N} \right| \right] \leq \max\{\delta, \frac{1}{N}\} C
\]
where $C$ depends only on the sup norms of $\Psi$ and $h$ and its first 3 derivatives. This vanishes as $N \to \infty$, so once we prove the following (41), Lemma 6.10 will be proved.

(41)
\[
\lim_{N \to \infty} \mathbb{E}\left[ |\frac{1}{N} \sum_{i=1}^{N} D_{5}^{i,N}| \right] = 0.
\]

Unlike the average of $D_{k_i}^{i,N}$, $k = 1, 2, 3, 4$, the term we wish to vanish in (41) is $O(1)$ in $N$. However, as we will see, the fact that $Q$ almost surely satisfies (V4) along with Remark 6.7 and the centering condition from Assumption (A8) will result in this term vanishing when we pass to the limit. We first observe that
\[
\mathbb{E}\left[ \left| \frac{1}{N} \sum_{i=1}^{N} D_{5}^{i,N} \right| \right] \leq \|\Psi\|_{\infty} \|\nabla h\|_{\infty} \mathbb{E}\left[ D^N \right]
\]
where
\[
D^N := \int_{s}^{t} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N} \sum_{j=1}^{N} \partial_{\mu} \Phi(\bar{X}_t^{i,N}, \bar{X}_t^{i,N}/\delta, \bar{\mu}_t^{N}) f(\bar{X}_t^{i,N}, \bar{X}_t^{i,N}/\delta, \bar{\mu}_t^{N}) |d\tau|
\]
We can rewrite $D^N$ in terms of the occupation measures defined in Equation 23 as:
\[
D^N = \int_{s}^{t} \int_{W} \int_{T^d} |\int_{T^d} \partial_{\mu} \Phi(\phi(\tau), y, \bar{\mu}_\tau^{N})(\psi(\tau)) f(\psi(\tau), \hat{y}, \bar{\mu}_\tau^{N}) n_\tau(dy) Q_N(dy) d\tau|
\]
\[
Q_N(d\phi d\mu d\nu) d\tau.
\]
And thus
\[
\mathbb{E}\left[ D^N \right] \leq \mathbb{E}\left[ \int_{s}^{t} \int_{W} \int_{T^d} |\int_{T^d} \partial_{\mu} \Phi(\phi(\tau), y, \bar{\mu}_\tau^{N})(\psi(\tau)) f(\psi(\tau), \hat{y}, \bar{\mu}_\tau^{N}) n_\tau(dy) Q_N(dy) d\tau| Q_N(d\phi d\mu d\nu) d\tau \right]
\]
\[
Q_N(d\phi d\mu d\nu) d\tau.
\]
First we show the first of these terms vanishes as $N \to \infty$. We have
\[
\lim_{N \to \infty} \mathbb{E}\left[ \int_{s}^{t} \int_{W} \int_{T^d} |\int_{T^d} \partial_{\mu} \Phi(\phi(\tau), y, \bar{\mu}_\tau^{N})(\psi(\tau)) f(\psi(\tau), \hat{y}, \bar{\mu}_\tau^{N}) n_\tau(dy) Q_N(dy) d\tau| Q_N(d\phi d\mu d\nu) d\tau \right]
\]
\[
\leq \lim_{j \to \infty} \sup_{N \in \mathbb{N}} \mathbb{E}\left[ \int_{s}^{t} \int_{W} \int_{T^d} |\int_{T^d} \partial_{\mu} \Phi(\phi(\tau), y, \hat{\mu}_\tau^{j})(\psi(\tau)) f(\psi(\tau), \hat{y}, \bar{\mu}_\tau^{j}) n_\tau(dy) Q_N(dy) d\tau| Q_N(d\phi d\mu d\nu) d\tau \right]
\]
We note now that by Propositions [A.3] and [C.2] along with Assumption [A1] and Bounded Convergence Theorem, for fixed \( \phi, \psi, n, m \) and \( \tau \)

\[
\lim_{j \to \infty} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \bar{\mu}^j) (\psi(\tau)) f(\psi(\tau), \bar{y}, \bar{\mu}^j) \right| \leq m_\tau(dy) n_\tau(d\bar{y}) m_\tau(dy) = 0.
\]

So by Assumption [A1] and Proposition [C.2] we have by Bounded Convergence Theorem that this term vanishes.

Now we have

\[
\lim_{N \to \infty} \mathbb{E}[D^N] \leq \lim_{N \to \infty} \mathbb{E} \left[ \int_{W} \int_{W} \int_{s} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right| m_\tau(dy) d\tau Q^N(d\psi d\mu d\rho) \right].
\]

By the same argument as above, we have that we can pass the limit through the expectation to get

\[
\lim_{N \to \infty} \mathbb{E}[D^N] \leq \mathbb{E} \left[ \lim_{N \to \infty} \int_{W} \int_{W} \int_{s} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right| m_\tau(dy) d\tau Q^N(d\psi d\mu d\rho) \right].
\]

Since \( Q^N \to Q \) in \( \mathcal{P}(W) \) almost surely and the integrand is bounded, we have via Proposition 4.6 on p.115 of [22] that

\[
\mathbb{E} \left[ \lim_{N \to \infty} \int_{W} \int_{W} \int_{s} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right| m_\tau(dy) d\tau Q^N(d\psi d\mu d\rho) \right]
\]

\[
= \mathbb{E} \left[ \int_{W} \int_{W} \int_{s} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right| m_\tau(dy) d\tau Q(d\psi d\mu d\rho) \right].
\]

Now by Hölder’s inequality and Tonelli’s Theorem,

\[
\mathbb{E} \left[ \int_{W} \int_{W} \int_{s} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) \left\{ \int_{T^d} f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right\} m_\tau(dy) d\tau \right| Q(d\psi d\mu d\rho) \right]
\]

\[
\leq \mathbb{E} \left[ \left( \int_{s} \left( \int_{W} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) \right|^2 m_\tau(dy) Q(d\psi d\mu d\rho) \right)^{1/2} d\tau \right)^2 \right]^{1/2}
\]

\[
\times \left( \int_{s} \left( \int_{W} \int_{T^d} \left| \int_{T^d} f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right|^2 Q(d\psi d\mu d\rho) d\tau \right)^{1/2} \right]^{1/2}
\]

\[
= \mathbb{E} \left[ \left( \int_{s} \left( \int_{W} \int_{T^d} \left| \int_{T^d} \partial_n \Phi(\phi(\tau), y, \nu_Q(\tau))(\psi(\tau)) \right|^2 m_\tau(dy) Q(d\psi d\mu d\rho) \right)^{1/2} d\tau \right)^2 \right]^{1/2}
\]

\[
\times \left( \int_{s} \left( \int_{W} \int_{T^d} \left| \int_{T^d} f(\psi(\tau), \bar{y}, \nu_Q(\tau)) n_\tau(d\bar{y}) \right|^2 Q(d\psi d\mu d\rho) d\tau \right)^{1/2} \right]^{1/2}.
\]
The first term in the product inside the expectation is bounded by Proposition C.2. For the second, we have that
\[
\int_{\mathcal{W}} \left| \int_{T^d} f(\psi(\tau), \hat{y}, \nu_Q(\tau)) n_\tau(dy) \right|^2 Q(d\psi) d\nu = \int_{X \times Y} \left| \int_{T^d} f(\psi(\tau), \hat{y}, \nu_Q(\tau)) n_\tau(dy) \right|^2 Q_{\mathcal{B}(X \times Y)}(d\psi) d\nu
\]
\[
= \int_X \int_Y \left| \int_{T^d} f(\psi(\tau), \hat{y}, \nu_Q(\tau)) n_\tau(dy) \right|^2 \lambda(d\psi) Q_X(d\psi)
\]
by writing \(Q_{\mathcal{B}(X \times Y)}(d\psi) = \lambda(d\psi)Q_X(d\psi)\). Since \(Q\) almost surely satisfies \((V4)\) we have by Remark 6.7 that \(\mathbb{P}\)-almost surely, for every \(\tau \in [0,1]\)
\[
\int_X \int_Y \left| \int_{T^d} f(\psi(\tau), \hat{y}, \nu_Q(\tau)) n_\tau(dy) \right|^2 \lambda(d\psi) Q_X(d\psi)
\]
\[
= \int_X \int_Y \left| \int_{T^d} f(\psi(\tau), \hat{y}, \nu_Q(\tau)) n_\tau(dy) \right|^2 Q_X(d\psi)
\]
\[
= 0 \text{ by Assumption } (A8)
\]

Thus \(\lim_{N \to \infty} E[D^N]\) is bounded by the expectation of a term which its almost surely bounded times a term which is almost-surely 0, and thus \((11)\) holds and the proof of Lemma 6.10 is complete.

\[\square\]

We have then that for each \((s, t, \psi, h) \in [0,1] \times [0,1] \times C_0(W) \times C^\infty_{\text{c}}(\mathbb{R}^d)\) there is a set \(Z_{(s,t,\psi,h)} \in \tilde{\mathcal{F}}\) such that \(\tilde{\mathbb{P}}(Z_{(s,t,\psi,h)}) = 0\) and
\[
E^{Q_\omega} \left[ \psi(M_h^{Q_\omega}(t) - M_h^{Q_\omega}(s)) \right] = 0, \forall \omega \in \tilde{\Omega} \setminus Z_{(s,t,\psi,h)}.
\]

Since there is a a countable collection of \(h \in C^\infty_{\text{c}}(\mathbb{R}^d)\) which is dense in \(C^\infty_{\text{c}}(\mathbb{R}^d)\), a countable collection \((s, t) \in [0,1]^2\) which is dense in \([0,1]^2\), and countably many \(\Phi \in C_b(W)\) generating each of the countably many sigma algebras \(\mathcal{G}_{n_i}\), letting \(Z\) be the union over all these countable collections of \(Z_{(s,t,\psi,h)}\), we have \(Z \in \tilde{\mathcal{F}}, \tilde{\mathbb{P}}(Z) = 0\), and
\[
E^{Q_\omega} \left[ \psi(M_h^{Q_\omega}(t) - M_h^{Q_\omega}(s)) \right] = 0, \forall \omega \in \tilde{\Omega} \setminus Z.
\]

So Theorem 6.8 is proved.

6.2.3. Proof of \((V2)\) By Skorohod’s representation theorem, we can invoke another probability space on which the convergence of \(Q^N \to Q\) occurs with probability 1. Without making a distinction in the notation between that probability space and our original one, we note that by Fatou’s lemma
\[
E^{Q} \left[ \int_{\mathbb{R}^m \times [0,1]} |\nu|^2 \rho(dzdt) \right] = E \left[ E^{Q} \left[ \int_{\mathbb{R}^m \times [0,1]} |\nu|^2 \rho(dzdt) \right] \right]
\]
\[
\leq \liminf_{N \to \infty} E \left[ \int_{\mathbb{W}} \left\{ \int_{\mathbb{R}^m \times [0,1]} |\nu|^2 \rho(dzdt) \right\} Q^N(d\phi) d\nu dt \right]
\]
\[
= \liminf_{N \to \infty} E \left[ \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{1} |u^N_i(s)|^2 ds \right]
\]
\[
< \infty \text{ by Assumption } (25),
\]
where in the first equality we use that through Theorem 6.8 we identified \(Q\) to \(\mathbb{P}\)-a.s. be the unique (by Proposition 2.3) deterministic measure in \(\mathcal{P}(W)\) with \(Z\)-marginal \(Q_Z\) solving the Martingale Problem 34.
6.2.4. Proof of (V3) This follows immediately from weak convergence, since by Proposition A.3 for \( f \in C_b(\mathbb{R}^d), \Theta \mapsto \int_{\mathbb{R}^d} f(x) \nu_0(0)(dx) \) is a continuous bounded map from \( \mathcal{P}(\mathcal{W}) \) to \( \mathbb{R} \). Thus,

\[
\int_{\mathbb{R}^d} f(x) \nu_0(dx) = \lim_{N \to \infty} \int_{\mathbb{R}^d} f(x) \left( \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i, N} \right)(dx)
\]

\[
= \lim_{N \to \infty} \mathbb{E} \left[ \int_{\mathbb{R}^d} f(x) \nu_{Q_N}(0)(dx) \right]
\]

\[
= \mathbb{E} \left[ \int_{\mathbb{R}^d} f(x) \nu_Q(0)(dx) \right]
\]

\[
= \int_{\mathbb{R}^d} f(x) \nu_Q(0)(dx),
\]

where again in the last step we use that through Theorem 6.8 we identified \( Q \) to \( \tilde{\mathbb{P}} \)-a.s. be the unique (by Proposition 2.3) deterministic measure in \( \mathcal{P}(\mathcal{W}) \) with \( \mathcal{Z} \)-marginal \( Q_Z \) solving the Martingale Problem 34. Thus we get that \( Q \) \( \tilde{\mathbb{P}} \)-a.s. satisfies (V3).

7. The Laplace Principal Lower Bound

We now proceed with proving the Laplace Principal Lower Bound:

\[
\liminf_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \geq \inf_{\theta \in \mathcal{P}(\mathcal{X})} \{F(\theta) + I(\theta)\}.
\]

It suffices to prove this bound along any subsequence such that the left hand side converges. Such a sequence exists since \(-\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \leq \|F\|_\infty\). Fix \( \eta > 0 \). By Proposition 3.3 for each \( N \in \mathbb{N} \), there exists \( v_N \in \mathcal{U}_N \) such that

\[
-\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \geq \frac{1}{N} \mathbb{E} \left[ \sum_{i=1}^{N} \int_{0}^{1} |v_i^N(t)|^2 dt \right] + \mathbb{E}[F(\bar{\mu}^N)] - \eta.
\]

Note also that for this choice of controls, we have for all \( N \in \mathbb{N} \),

\[
\mathbb{E} \left[ \sum_{i=1}^{N} \int_{0}^{1} |v_i^N(t)|^2 dt \right] \leq 4\|F\|_\infty + 2\eta.
\]

Thus, since (23) is satisfied, by the proof of Theorem 4.4 in [10], we also get that it is enough to assume that \( \mathbb{P} \) almost-surely,

\[
\frac{1}{N} \sum_{i=1}^{N} \int_{0}^{1} |v_i^N(t)|^2 dt \leq \frac{4\|F\|_\infty (4\|F\|_\infty + \eta)}{\eta},
\]

so the results of Section 6 apply with \( \{v^N\}_{N \in \mathbb{N}} \) as our choice of controls, and for \( \{Q^N\}_{N \in \mathbb{N}} \) as in Equation 23 with \( \mathcal{Z} \)-marginal determined by \( \{v^N\}_{N \in \mathbb{N}} \), \( \mathcal{L}(Q^N) \to \delta_Q \) in \( \mathcal{P}(\mathcal{P}(\mathcal{W})) \) such that \( Q \in \mathcal{V} \) almost-surely. So

\[
\liminf_{N \to \infty} -\frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] \geq \liminf_{N \to \infty} \left[ \frac{1}{2} \mathbb{E} \left[ \sum_{i=1}^{N} |v_i^N(t)|^2 dt \right] + \mathbb{E}[F(\bar{\mu}^N)] \right] - \eta
\]

\[
= \liminf_{N \to \infty} \left[ \mathbb{E} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \rho(dzd\tau)Q_{X}^\theta(d\tau) + \mathbb{E}[F(\bar{Q}_{X})] \right] \right] - \eta
\]

\[
\geq \frac{1}{2} \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \rho(dz d\tau) Q_{X}(d\tau) + F(\bar{Q}_{X}) - \eta
\]

by Fatou’s Lemma

\[
\geq \inf_{\theta \in \mathcal{P}(\mathcal{X})} \left\{ \inf_{\Theta \in \mathcal{P}(\mathcal{Z})} \mathbb{E}^{\Theta} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \rho(dz d\tau) \right] + F(\bar{\Theta}) \right\} - \eta
\]

\[
= \inf_{\theta \in \mathcal{P}(\mathcal{X})} \{I(\theta) + F(\theta)\} - \eta.
\]
Since η is arbitrary the lower bound (42) is proved.

8. The Laplace Principal Upper Bound

In order to close the proof of Theorem 3.4, we now need to show for F ∈ C_b(\mathcal{P}(\mathcal{X})) that

$$\limsup_{N \to \infty} -\frac{1}{N} \log \mathbb{E}\left[\exp(-NF(\mu^N))\right] \leq \inf_{\theta \in \mathcal{P}(\mathcal{X})} \left\{ I(\theta) + F(\theta) \right\}. \quad (44)$$

Given η > 0, take \( \theta \in \mathcal{P}(\mathcal{X}) \) such that

$$I(\theta) + F(\theta) \leq \inf_{\theta \in \mathcal{P}(\mathcal{X})} \left\{ I(\theta) + F(\theta) \right\} + \frac{\eta}{2}. \quad (45)$$

Since the bound given in Equation (44) is trivial if the right hand side is +∞, we may assume it is finite. By the definition of I, there exists \( \Theta \in \mathcal{V} \) such that \( \Theta_\mathcal{X} = \theta \). By merit of (V2) we get that letting \((X, m, \rho)\) be the canonical process on \( \mathcal{W} \) as defined by Equation (13) that there exists a \( m \)-dimensional \( \mathcal{G}_{t+}^\Theta \)-Brownian motion \( W \) such that

$$((\mathcal{W}, \mathcal{B}(\mathcal{W}), \Theta), \{\mathcal{G}_{t+}^\Theta\}, (X, m, \rho, W))$$

is a weak solution of (10). Note that we take the \( \Theta \)-augmentation and right limit of \( \mathcal{G}_t \) so that we have a filtration that satisfies the usual conditions. By (V1) we know that the martingale problem \((34)\) is satisfied by (45)

$$\limsup_{N \to \infty} -\frac{1}{N} \log \mathbb{E}\left[\exp(-NF(\mu^N))\right] \leq \inf_{\theta \in \mathcal{P}(\mathcal{X})} \left\{ I(\theta) + F(\theta) \right\}. \quad (46)$$

Then we see that defining \( \tilde{\rho}(\omega)(I \times D) = \int_I \delta_{\tilde{u}(\omega)}(D)dt \) for \( D \in \mathcal{B}(\mathbb{R}^m), I \in \mathcal{B}([0, 1]), \omega \in \mathcal{W} \), that

$$((\mathcal{W}, \mathcal{B}(\mathcal{W}), \tilde{\theta}), \{\mathcal{G}_{t+}^{\tilde{\theta}}\}, (\tilde{X}, m, \tilde{\rho}, W))$$

is still a weak solution to (11) and the cost of \( \tilde{\rho} \) does not exceed that of \( \rho \).

We thus can find \( \tilde{\Theta} \in \mathcal{V} \) such that \( \tilde{\Theta}_\mathcal{X} = \tilde{\theta} \), there exists a \( m \)-dimensional \( \mathcal{G}_{t+}^{\tilde{\theta}} \)-Brownian motion \( W \) such that

$$((\mathcal{W}, \mathcal{B}(\mathcal{W}), \tilde{\Theta}), \{\mathcal{G}_{t+}^{\tilde{\Theta}}\}, (\tilde{X}, m, \tilde{\rho}, W))$$

is a weak solution to (11) with \( \tilde{\rho}_t(\omega)(D) = \delta_{\tilde{u}(\omega)}(D) \) for \( D \in \mathcal{B}(\mathbb{R}^m) \), and \( \tilde{u} \) a \( \mathbb{R}^m \)-valued process on \( \mathcal{W} \) such that

$$\mathbb{E}^{\tilde{\Theta}} \left[ \frac{1}{2} \int_0^1 |\tilde{u}(t)|^2 dt \right] \geq \frac{1}{2} \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \tilde{\rho}(dzdt) \leq I(\tilde{\theta}) + \frac{\eta}{2}. \quad (47)$$

Now let us define a filtered probability space \((\Omega_\infty, \mathcal{F}_\infty, \mathbb{P}_\infty), \{\mathcal{F}_t^\infty\}\) by taking countably infinitely many products of

$$((\mathcal{W}, \mathcal{B}(\mathcal{W}), \tilde{\Theta}), \{\mathcal{G}_{t+}^{\tilde{\Theta}}\}).$$

For \( \omega = (\omega_1, \omega_2, ...) \in \Omega_\infty \), define (see also Section 6 of [11])

$$u^\infty_i(t, \omega) = \tilde{u}(t, \omega_i), \quad i \in \mathbb{N}, t \in [0, 1]. \quad (48)$$

Note that \{\( u^\infty_i \)\}_{i \in \mathbb{N}}\ are independent and identically-distributed and

$$\mathbb{E}^\infty \left[ \frac{1}{N} \sum_{i=1}^N \int_0^1 |u^\infty_i(t)|^2 dt \right] = \frac{1}{N} \sum_{i=1}^N \mathbb{E}^\infty \left[ \int_0^1 |u^\infty_i(t)|^2 dt \right] = \mathbb{E}^{\tilde{\Theta}} \left[ \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \tilde{\rho}(dzdt) \right] < \infty \quad (49)$$

by (V2).
Let \( \{W^i\}_{i \in \mathbb{N}} \) be independent \( m \)-dimensional \( \mathcal{G}_{t_i} \)-Brownian motions and \( \{\tilde{X}^{i,N}\}_{i \in \{1,...,N\}} \) be the unique solution to the system of SDEs on \((\Omega_\infty, \mathcal{F}_\infty, \mathbb{P}_\infty)\)

\[
\begin{aligned}
d\tilde{X}^{i,N}_t &= \left[ \frac{1}{\delta} f(\tilde{X}^{i,N}_t, \tilde{X}^{i,N}_t/\delta, \tilde{\mu}^{N}_t) + b(\tilde{X}^{i,N}_t, \tilde{X}^{i,N}_t/\delta, \tilde{\mu}^{N}_t) + \sigma(\tilde{X}^{i,N}_t, \tilde{X}^{i,N}_t/\delta, \tilde{\mu}^{N}_t)u^{i,\infty}(t) \right] dt \\
\mathbb{P}_0 \times \tilde{X}^{i,N} &= \tilde{x}^{i,N}
\end{aligned}
\]

for \( N \in \mathbb{N} \) and \( \tilde{\mu}^{N}_t \) the empirical measure of \( \tilde{X}^{1,N},...,\tilde{X}^{N,N} \) at time \( t \).

Now defining

\[
\tilde{Q}^N_{\omega}(A \times B \times C) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\tilde{X}^{i,N}_{\omega}(\omega)}(A) \delta_{\tilde{\mu}^{i,N}_{\omega}}(B) \delta_{\tilde{\rho}^{i,\infty}_{\omega}}(C)
\]

where \( \tilde{\rho}^{i,\infty}_{\omega}(I \times D) = \tilde{\rho}(I \times D) = \int_I \delta_{\tilde{\rho}(t,\omega)}(D) dt = \int_I \delta_{u^{i,\infty}(t,\omega)}(D) dt \) and \( \tilde{m}^{i,N}(\omega) \) is as in Equation \( 24 \) with \( \tilde{X}^{i,N} \) in the place of \( \bar{X}^{i,N} \). Since Equation \( 15 \) holds, \( \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=1}^{N} \int_0^1 |u^{i,\infty}_t(t)|^2 dt < \infty \) \( \mathbb{P}_\infty \) almost-surely, and the proof of tightness and the fact that \( \tilde{Q}^N \to \tilde{Q} \in \mathcal{P}(\mathcal{W}) \)-valued random variable along some subsequence holds by the proofs in Section 6.

We wish now to conclude that \( \tilde{Q}_\omega = \tilde{\Theta}_\omega \). By Proposition \( 2.3 \) weak-sense uniqueness as defined in Definition \( 2.2 \) holds, so it suffices to prove that \( \tilde{Q}_\omega = \tilde{\Theta}_\omega \). By the mapping theorem (Theorem 2.7 in \( \mathbf{5} \)) and continuity of the projection operator from \( \mathcal{W} \) to \( \mathcal{Z} \), we can simply show \( \tilde{Q}^N_Z \to \tilde{Q}_Z \) weakly as a \( \mathcal{P}(\mathcal{Z}) \)-valued random variable. Since \( \tilde{\rho}^{i,\infty}_{\omega} \) are independent and identically-distributed under \( \mathbb{P}_\infty \) with common distribution the same as that of \( \tilde{\rho} \) under \( \tilde{\Theta} \), for \( \tilde{Q} \) a limit point of \( \{Q^N\}_{N \in \mathbb{N}} \) defined on a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \), we have by Varadarajan’s theorem (\( \mathbf{15} \) p.399) that \( \bar{\mathbb{P}} \)-almost surely,

\[
\tilde{Q}_Z = \tilde{\Theta} \circ \tilde{\rho}^{-1} = \tilde{\Theta}_Z.
\]

Therefore \( \tilde{Q}_\omega = \tilde{\Theta}_\omega \), and we have, where the infimum in the first line is taken to be over all stochastic bases,

\[
\limsup_{N \to \infty} - \frac{1}{N} \log \mathbb{E}[\exp(-NF(\mu^N))] = \limsup_{N \to \infty} \inf_{u^N \in \mathcal{U}^N} \left\{ \frac{1}{2} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^{N} |u^{i,N}_t(t)|^2 dt \right] + \mathbb{E}[F(\mu^N)] \right\}
\]

\[
\leq \limsup_{N \to \infty} \left\{ \frac{1}{2} \mathbb{E}^{\tilde{\Theta}} \left[ \frac{1}{N} \sum_{i=1}^{N} \int_0^1 |u^{i,N}_t(t)|^2 dt \right] + \mathbb{E}^{\tilde{\Theta}}[F(\mu^N)] \right\}
\]

\[
= \mathbb{E}^{\tilde{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \tilde{\rho}(dz dt) \right] + \limsup_{N \to \infty} \mathbb{E}^{\tilde{\Theta}}[F(\tilde{Q}^N_Z)]
\]

\[
\leq \mathbb{E}^{\tilde{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \tilde{\rho}(dz dt) \right] + F(\tilde{\Theta}_Z)
\]

\[
\leq I(\theta) + F(\theta) + \frac{\eta}{2}
\]

Since \( \eta \) is arbitrary, Equation \( 44 \) is proved.

9. Compactness of Level Sets

Consider \( I \) as defined in Equation \( 16 \). We want to prove that for each \( s \in [0, \infty) \), the set

\[
I_s := \{ \theta \in \mathcal{P}(\mathcal{X}) : I(\theta) \leq s \}
\]

is a compact subset of \( \mathcal{P}(\mathcal{X}) \). This will imply that indeed \( I \) is a good rate function.

Since in this section we are dealing with sequences of measures all of which coincide with weak solutions of Equation \( 12 \) but with possibly different controls, we introduce a new notation for the coordinate process which allows us to keep track of which measure the \( \mathcal{X} \)-component of the coordinate process corresponds to. For this we use the parameterized version of the limiting Equation \( 10 \).
For $Q$ corresponding to a weak solution of Equation 12, $Q$ also corresponds to a solution of Equation 10 with $\nu_Q$ as defined in Equation 14 in the place of $\nu$. Thus we consider the process triple $(\tilde{X}^{\nu_Q}, m, \rho)$, which can be given explicitly as the coordinate process on the probability space $(W, B(W), Q)$ endowed with the canonical filtration $\mathcal{G}_t := \sigma\left(\tilde{X}^{\nu_Q}(s), m(s), \rho(s), 0 \leq s \leq t\right)$. Thus, for $\omega = (\phi, n, r) \in W$,

$$\tilde{X}^{\nu_Q}(\omega) = \phi(t), \quad m(t, \omega) = n|_{B(\mathbb{R}^m \times [0, t])}, \quad \rho(t, \omega) = r|_{B(\mathbb{R}^m \times [0, t])}.$$ (47)

**Lemma 9.1.** Fix $K < \infty$ and consider a sequence $\{Q^N\}_{N \in \mathbb{N}} \subset \mathcal{P}(W)$ such that for every $N \in \mathbb{N}$, $Q^N$ is viable and

$$\mathbb{E}^{Q^N}\left[ \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \rho(dzdt) \right] < K.$$

Then $\{Q^N\}_{N \in \mathbb{N}}$ is tight.

**Proof.** As in Subsection 6.1, it suffices to show tightness of each of the marginals. It is worth noting that where before we were proving tightness of $L(Q^N)$ in $\mathcal{P}(P(W))$, here we have that $Q^N$ are deterministic measures and we are proving tightness of the measures themselves in $\mathcal{P}(Q^N)$.

Tightness of the $Y$-marginals follows in essentially the same way as in Subsection 6.1.2. $\mathcal{P}(Y)$ is itself compact so $\{Q^N_Y\}_{N \in \mathbb{N}}$ is tight.

Tightness of the $Z$-marginals is also very similar to Subsection 6.1.1.

$$g(r) := \int_{\mathbb{R}^m \times [0, 1]} |z|^2 r(dzdt)$$

is a tightness function on $\mathcal{R}_1$, so since

$$\mathbb{E}^{Q^N}\left[ \int_{\mathbb{R}^m \times [0, 1]} |z|^2 \rho(dzdt) \right] < \infty,$$

$\{Q^N_Z\}_{N \in \mathbb{N}}$ is tight.

For the tightness of the $X$-marginals, we use that each $Q^N$ satisfies $[V1]$ that is, $Q^N_X = L(\tilde{X}^{\nu_Q})$. Via Theorem 2.4.10 in 11, it suffices to show that for every $\eta > 0$,

$$\limsup_{\rho_0 \downarrow 0, N \in \mathbb{N}} Q^N_X\left( \sup_{t_1-t_2 \leq \rho_0 \leq t_1 < t_2 \leq 1} |\tilde{X}^{\nu_Q}_{t_1} - \tilde{X}^{\nu_Q}_{t_2}| \geq \eta \right) = 0,$$

where here we are using the notation from Equation 17. We have that by Chebyshev’s inequality,

$$\limsup_{\rho_0 \downarrow 0, N \in \mathbb{N}} Q^N_X\left( \sup_{t_1-t_2 \leq \rho_0 \leq t_1 < t_2 \leq 1} |\tilde{X}^{\nu_Q}_{t_1} - \tilde{X}^{\nu_Q}_{t_2}| \geq \eta \right) \leq \lim_{\rho_0 \downarrow 0} \frac{1}{\eta} \sup_{N \in \mathbb{N}} \mathbb{E}^{Q^N}\left[ \sup_{t_1-t_2 \leq \rho_0 \leq t_1 < t_2 \leq 1} |\tilde{X}^{\nu_Q}_{t_1} - \tilde{X}^{\nu_Q}_{t_2}| \right].$$

Since

$$|\tilde{X}^{\nu_Q}_{t_1} - \tilde{X}^{\nu_Q}_{t_2}| = \left| \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left[ \nabla_y \Phi(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) + I ] b(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) + \sigma(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) \int_{\mathbb{R}^m} z \rho_t(dz) \right] + \nabla_x \Phi(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) f(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) + A : \nabla_x \nabla_y \Phi(\tilde{X}^{\nu_Q}_{t_1}, y, \nu_Q(t)) \right] m_t(dy)dt $$

$$+ \int_{t_1}^{t_2} B(t, \tilde{X}^{\nu_Q}_{t_1}, \nu(t))dW_t, $$

$$\frac{1}{2} B(t, x, \mu) \int_{\mathbb{T}^d} \left( \nabla_y \Phi(x, y, \mu) + \frac{1}{2} A(x, y, \mu) + f(x, y, \mu) \otimes \Phi(x, y, \mu) \right) m_t(dy),$$

we get via Hölder’s inequality, Itô isometry, Assumption $[A1]$ and Proposition C.2 that

$$|\tilde{X}^{\nu_Q}_{t_1} - \tilde{X}^{\nu_Q}_{t_2}| \leq C\left( (t_2 - t_1) + \sqrt{t_2 - t_1} \left( \int_{0}^{t_1} \int_{\mathbb{R}^m} |z|^2 \rho_t(dz)dt + 1 \right) \right).$$
Then we have by Young’s inequality that
\[
\sup_{|t_1-t_2|<\rho,0\leq t_1\leq t_2\leq 1} |\tilde{X}^{\nu^{Q_N}}_{t_1} - \tilde{X}^{\nu^{Q_N}}_{t_2}| \leq C \left( \frac{5}{2} + \frac{1}{2} \int_0^1 \int_{\mathbb{R}^m} |z|^2 \rho_t(dz) dt \right).
\]
Then
\[
\sup_{N \in \mathbb{N}} \mathbb{E}^{Q_N} \left[ \left( \frac{5}{2} + \frac{1}{2} \int_0^1 \int_{\mathbb{R}^m} |z|^2 \rho_t(dz) dt \right) \right] \leq C \left( \frac{5}{2} + \frac{1}{2} K \right) \text{ by assumption.}
\]
So by dominated convergence theorem, we have
\[
\limsup_{\rho \downarrow 0} \eta \in \mathbb{N} \mathbb{E}^{Q_N} \left[ \sup_{|t_1-t_2|<\rho,0\leq t_1\leq t_2\leq 1} |\tilde{X}^{\nu^{Q_N}}_{t_1} - \tilde{X}^{\nu^{Q_N}}_{t_2}| \geq \eta \right]
\leq \limsup_{\rho \downarrow 0} \eta \in \mathbb{N} \mathbb{E}^{Q_N} \left[ \sup_{|t_1-t_2|<\rho,0\leq t_1\leq t_2\leq 1} |\tilde{X}^{\nu^{Q_N}}_{t_1} - \tilde{X}^{\nu^{Q_N}}_{t_2}| \right]
\leq C \eta \sup_{N \in \mathbb{N}} \mathbb{E}^{Q_N} \left[ \lim_{\rho \downarrow 0} \sup_{|t_1-t_2|<\rho,0\leq t_1\leq t_2\leq 1} (t_2 - t_1) + \sqrt{t_2 - t_1} \left( \int_0^1 \int_{\mathbb{R}^m} \left| z \right|^2 \rho_t(dz) dt + 1 \right) \right]
= 0.
\]

\[\square\]

**Lemma 9.2.*** Fix $K < \infty$ and consider a convergent sequence $\{Q^N\}_{N \in \mathbb{N}} \subset \mathcal{P}(W)$ such that for every $N \in \mathbb{N}$, $Q^N$ is viable and
\[
\mathbb{E}^{Q_N} \left[ \int_{\mathbb{R}^m \times \{0,1\}} \left| z \right|^2 \rho(dz dt) \right] < K.
\]
Then for $Q$ such that $Q^N \to Q$, $Q$ is viable.

**Proof.** The fact that $Q$ satisfies (V2) follows immediately from Fatou’s lemma. Since by Proposition A.3 $\nu_0 = \lim_{N \to \infty} \nu^{Q_N}(0) = \nu_Q(0))$, (V3) is satisfied.

We now prove $Q$ satisfies (V1). As before, our tool here is the martingale problem associated to Equation 10. It suffices to show that for fixed $h \in C_c(\mathbb{R}^d)$, $0 \leq s \leq t \leq 1$, and $\mathcal{G}_s$-measurable $\Psi \in C_b(\mathcal{W})$ that
\[
\mathbb{E}^{Q} \left[ \Psi(M^Q_h(t) - M^Q_h(s)) \right] = 0
\]
where $M^Q_h$ is given in Equation 33. It suffices to show that
\[
\mathbb{E}^{Q_N} \left[ \Psi(M^Q_N(t) - M^Q_N(s)) \right] \to \mathbb{E}^{Q} \left[ \Psi(M^Q_h(t) - M^Q_h(s)) \right]
\]
since by (V1)
\[
\mathbb{E}^{Q_N} \left[ \Psi(M^Q_N(t) - M^Q_N(s)) \right] = 0.
\]

Unlike in the previous proof of (V1) here the convergence is as a sequence of real numbers and not in distribution, since $Q^N$ are deterministic.

So that we can keep track of which measure $m$ and $\rho$ correspond to in the Coordinate Process 47 on $W$ under $Q^N$, we relabel it $(\tilde{X}^{\nu^{Q_N}}, m^N, \rho_N)$. Under $Q$, we keep the notation $(\tilde{X}^{\nu_Q}, m, \rho)$. Invoking Skorohod’s representation theorem to find another probability space on which the convergence of the random variables $(\tilde{X}^{\nu^{Q_N}}, m^N, \rho^N) \to (\tilde{X}^{\nu_Q}, m, \rho)$ occurs for almost every $\omega \in \Omega$, we have
\[
\mathbb{E} \left[ \left| M^{Q_N}_h(t) - M^Q_h(t) + M^{Q}_N(s) - M^Q(s) \right| \right] \leq C(\|\Psi\|_\infty) \left( \mathbb{E} \left[ \left| M^{Q_N}_h(t) - M^Q_h(t) \right| \right] \right) + \mathbb{E} \left[ \left| M^{Q}_N(s) - M^Q(s) \right| \right]
\]
and

\[ E \left[ M_h^{Q_N}(t) - M_h^Q(t) \right] \\
= E \left[ g(\bar{X}_{t}^{\nu_Q}) - g(\bar{X}_{t}^{\nu_Q}) + \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(dz) ds \\
- \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(N) dz) ds \right]. \]

By continuity and boundedness of \( g \) and convergence of \( \bar{X}_{t}^{\nu_Q} \rightarrow \bar{X}_{t}^{\nu_Q} \) along with bounded convergence theorem,

\[ E \left[ g(\bar{X}_{t}^{\nu_Q}) - g(\bar{X}_{t}^{\nu_Q}) \right] \rightarrow 0 \text{ as } N \rightarrow \infty. \]

By Assumption \((A1)\) and Proposition \((C.2)\) along with bounded convergence theorem,

\[ \lim_{N \rightarrow \infty} E \left[ \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(dz) ds \\
- \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(N) dz) ds \right] \leq \lim_{N \rightarrow \infty} E \left[ \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(dz) ds \\
- \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(N) dz) ds \right] \\
= \lim_{N \rightarrow \infty} E \left[ \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(dz) ds \\
- \int_0^t \int_{\mathbb{R}^m} \int_{T_d} A[g](\bar{X}_{s}^{\nu_Q}, y, z, \nu_Q(s)) m_s(dy) \rho_s(N) dz) ds \right]. \]

By continuity of the coefficients in \( x \) and \( \mu \) from Assumption \((A1)\) and Proposition \((C.2)\) along with the assumed uniform \( L^2 \) bound on the control and with the fact that the growth in the control is linear, if we can show that \( \nu_{Q_N}(t) \rightarrow \nu_Q(t) \) in \( P(\mathbb{R}^d) \) for each \( t \in [0,1] \) then this term will vanish by the same triangle-inequality argument given in the proof of Lemma \((6.9)\). But this follows immediately by the assumption that \( Q^N \rightarrow Q \) almost surely and Proposition \((A.3)\) and by Chebyshev’s inequality and the same density argument as at the end of Subsection \((6.2.2)\) we have that \( Q \) satisfies \((V1)\).

Finally we prove that \( Q \) satisfies \((V4)\). Again invoking Skorohod’s representation theorem to find another probability space on which the convergence of the random variables \( (\bar{X}_{t}^{\nu_Q}, m^N, \rho^N) \rightarrow (\bar{X}_{t}^{\nu_Q}, m, \rho) \) occurs for almost every \( \omega \in \Omega \),

\[ E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m(dy, ds) \right] \]

\[ \leq \liminf_{N \rightarrow \infty} E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m^N(dy, ds) \right] \text{ by Fatou’s Lemma} \]

\[ \leq \liminf_{N \rightarrow \infty} E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m^N(dy, ds) \right] \]

\[ + \liminf_{N \rightarrow \infty} E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) - L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m^N(dy, ds) \right] \]

\[ = \liminf_{N \rightarrow \infty} E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) - L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m^N(dy, ds) \right] \text{ by \((V4)\)} \]

\[ \leq \sup_{N \in \mathbb{N}} \lim_{J \rightarrow \infty} E \left[ \int_{T_d \times [0,t]} L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) - L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} f(y) m^N(dy, ds) \right] \]

by Assumption \((A1)\) and Dominated Convergence Theorem

\[ = 0 \]

by continuity of \( L^1_{\bar{X}_{s}^{\nu_Q},\nu_Q(s)} \) in \( \mu \) via Assumption \((A1)\) and Proposition \((A.3)\).

So \( Q \) satisfies \((V4)\). □
Lemma 9.1 establishes precompactness of \( I_s \) defined in (49). Now we will use both Lemmas 9.1 and 9.2 to prove the level sets \( I_s \) are closed via showing lower-semicontinuity of \( I \).

**Lemma 9.3.** The functional \( I \) given in Equation (46) is lower semi-continuous.

*Proof.* Consider a sequence \( \{\theta^N\} \subset \mathcal{P}(X) \) with limit \( \theta \). We wish to show

\[
\lim_{N \to \infty} I(\theta^N) \geq I(\theta).
\]

It suffices to consider the case where the left hand side is finite, so there is \( M \in [0, \infty) \) such that \( \lim_{N \to \infty} I(\theta^N) \leq M \). Then, recalling that

\[
I(\theta^N) = \inf_{\Theta^N \in \mathcal{V} \cdot \Theta^N = \theta^N} \mathbb{E}^{\Theta^N} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right],
\]

by taking a subsequence of \( \{\theta^N\} \) if necessary, we can find measures \( \Theta^N \) such that \( \Theta^N = \theta^N \),

\[
(48) \quad \sup_{N \in \mathbb{N}} \mathbb{E}^{\Theta^N} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right] < M + 1,
\]

and

\[
I(\theta^N) \geq \mathbb{E}^{\Theta^N} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right] - \frac{1}{N}.
\]

Then by Lemma 9.1 we can consider a subsequence along which \( \{\Theta^N\} \) converges to some \( \Theta \). By Lemma 9.2 \( \Theta \) is viable. Hence by Fatou’s lemma,

\[
\liminf_{N \to \infty} I(\theta^N) \geq \liminf_{N \to \infty} \mathbb{E}^{\Theta^N} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right] - \frac{1}{N}
\]

\[
\geq \mathbb{E}^{\Theta} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right]
\]

\[
\geq \inf_{\Theta \in \mathcal{V} \cdot \Theta = \theta} \mathbb{E}^{\Theta} \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right]
\]

\[
= I(\theta),
\]

so lower semi-continuity of \( I \) is proved. \( \Box \)

10. A More General Regime

For presentation purposes we proved Theorem 3.4 under quite strong boundedness assumptions for the coefficients. The proofs however make it clear that one can relax such assumptions at the expense of requiring more precise information on the behavior of the solution to Equation (4) and its derivatives.

In particular, let us make the following set of assumptions:

(A1') For some \( \nu_0 \in \mathcal{P}_2(\mathbb{R}^d) \), \( \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i, N} \to \nu_0 \) as \( N \to \infty \).

(A2') The system of SDEs (1) admits a unique strong solution.

(A3') There exists a unique invariant measure \( \pi(dy|x, \mu) \) satisfying Equation (3) and the centering condition:

\[
\int_{\mathbb{R}^d} f(x, y, \mu) \pi(dy|x, \mu) = 0, \forall x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d)
\]

holds.

(A4') There exists a unique function \( \Phi : \mathbb{R}^d \times \mathbb{T}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \), \( \Phi = (\Phi_1, ..., \Phi_d) \) solving Equation (1).

(A5') For each \( x \in \mathbb{R}^d, y \in \mathbb{T}^d \), \( \Phi(x, y, \cdot) \) is fully \( C^2 \) in the sense of Definition 1.2.

(A6') The partial derivatives in \( x \) and \( y \) of \( \Phi, \partial_x \Phi, \) and \( \nabla_x \Phi, \nabla_y \Phi \) exist.

(A7') There exists \( p \geq 2 \) such that for \( \tilde{X}_{i, N} \) controlled by any \( u^N \in \mathcal{U}_N \) satisfying \( \sup_{N \in \mathbb{N}} \mathbb{E} \left[ \sum_{i=1}^{N} \int_{0}^{t} |u^N_i(t)|^2 dt \right] < \infty \), \( \sup_{0 \leq t \leq T} \sum_{i=1}^{N} |\tilde{X}_{i, N}^N|^p \) is finite. This assures that for all \( N \), there exists a modification of \( \tilde{\mu}^N \in C([0,1]; \mathcal{P}_2(\mathbb{R}^d)) \) so that \( \tilde{\mu}^N_t \) is in the domain of the coefficients and \( \Phi \) for all time.

38

Z.W. BIZEMJ, AND K. SPLICPOULOS
(A8') There exists $p_2 > p$ such that for $\tilde{X}^{i,N}$ controlled by any $u^{N} \in \mathcal{U}_{N}$ satisfying $\sup_{N \in \mathbb{N}} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^{N} \left| \tilde{X}^{i,N}_t \right|^p dt \right] < \infty$, $\sup_{N \in \mathbb{N}} \mathbb{E} \left[ \int_0^1 \frac{1}{N} \sum_{i=1}^{N} \left| \tilde{X}^{i,N}_t \right|^{p_2} dt \right] < \infty$.

(A9') For $(x, y, \nu) \in \mathbb{R}^d \times \mathbb{T}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and $\alpha > 0$ sufficiently small, we have the following terms are bounded by

$$C \left( 1 + |x|^p + \nu(|\cdot|^p) \right)$$

for some $C > 0$ and $p$ from Assumption (A7')

i.) $|f(x, y, \nu)| + |b(x, y, \nu)| + \sigma(x, y, \nu)|^2$

ii.) $|\nabla_x \Phi(x, y, \nu)\sigma(x, y, \nu)|^2 + |\Phi(x, y, \nu)|^2 |\sigma(x, y, \nu)|^2 + |I + \nabla_y \Phi(x, y, \nu)\sigma(x, y, \nu)|^2$

iii.) $\| \partial_\nu \Phi(x, y, \nu)\sigma(x, y, \nu) \|_{L^2(\nu, \mathbb{R}^d)} + \| \partial_\mu \Phi(x, y, \nu) \|_{L^2(\nu, \mathbb{R}^d)}$

iv.) $|I + \nabla_y \Phi(x, y, \nu)| b(x, y, \nu) + \nabla_x \Phi(x, y, \nu) f(x, y, \nu) + A : \nabla_x \nabla_y \Phi(x, y, \nu)$

v.) $|I + \nabla_y \Phi(x, y, \nu)| A(x, y, \nu) + f \otimes \Phi(x, y, \nu)$

vi.) $|\nabla_x \Phi(x, y, \nu) b(x, y, \nu) + A : \left[ \frac{1}{2} \nabla_x \nabla_x \Phi(x, y, \nu) + \nabla_y \partial_\mu \Phi(x, y, \nu) \right](x) + \alpha \nabla_x \partial_\mu \Phi(x, y, \nu)(x) + \int_{\mathbb{R}^d} \left[ \partial_\mu \Phi(x, y, \nu) \right](v) b(v, y, \nu)$

$\frac{1}{2} A(v, y, \nu) : \left[ \partial_\nu \partial_\mu \Phi(x, y, \nu) + \alpha \partial_\mu^2 \Phi(x, y, \nu) \right](v) \nu(dv) |$

vii.) $|\nabla_x \Phi(x, y, \nu) + \alpha \partial_\mu \Phi(x, y, \nu)(x)| A(x, y, \nu) + b \otimes \Phi(x, y, \nu)$

(A10') For fixed $x, \hat{x} \in \mathbb{R}^d, y \in \mathbb{T}^d$, we have $\nu \mapsto \partial_{\mu} \Phi(x, y, \nu)(\hat{x})$ and $\nu \mapsto \tilde{f}(x, y, \nu)$ for $\tilde{f} = b, f, \sigma, \Phi, \nabla_y \Phi, \nabla_x \Phi$, and $\nabla_x \nabla_y \Phi$ are continuous in $\mathcal{P}(\mathbb{R}^d)$ (or if $p > 2$ in (A7') it is sufficient to have continuity in $(\mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2)$).

(A11') For $x, \hat{x} \in \mathbb{R}^d, y, \hat{y} \in \mathbb{T}^d, \nu, \tilde{\nu} \in \mathcal{P}_2(\mathbb{R}^d)$ and $p$ from Assumption (A7') we assume $|\partial_{\nu} \Phi(x, y, \nu)(\hat{x}) f(\hat{x}, \hat{y}, \nu)| \leq C(1 + |x|^p + |\hat{x}|^p + \nu(|\cdot|^p))$.

(A12') Weak uniqueness in the sense of Definition 2.2 holds for the limiting Equation 12.

Admittedly these assumptions seem somewhat contrived to meet the requirements of our proofs, but the reason for this presentation is two-fold. The first reason is that these assumptions encapsulate two major classes of SDEs for the particles: ones where the coefficients are bounded (in which case we can drop assumption (A8')), and ones where $b$ is a confining potential and the growth of $f$ and $\sigma$ in $x$ and $\mu$ is not too fast (as in Section 5). The second is that, with the notable exception of [54], regularity of the Cell Problem 3 with respect to the parameter $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ is not yet well-studied in the existing literature. Thus this regularity must be studied on a case-to-case basis in applications, though sufficient regularity can be proved either through explicitly solving for $\Phi$ or through the representation offered by Equation 54 in the proof of Proposition 5.2.

The proofs go through in the same manner as under assumptions (A1) - (A8) with a few notable changes. The first is that the necessary results of Appendix A and Proposition 2.3 are assumed rather than proven. The second is that the bounds in assumption (A9') are needed for the proof of tightness in Subsection 6.1. Those bounds yield the moment dependence on $|\tilde{X}^{i,N}_t|^p$ which then due to Assumption (A7') - (A8') allow us to conclude tightness and then convergence. (A11') is needed for technical reasons in the proof of Lemma 6.10.

Lastly, the class of viable measures in the definition of the rate function 16 must be modified:
Definition 10.1. We will say \( \Theta \in \mathcal{P}(\mathcal{W}) \) is in \( \mathcal{V}^\prime \) if

1. \( \Theta \) satisfies \((V1)-(V3)\) and \((V4)\) from Definition 3.1
2. \( \Theta \) satisfies \((V2)\): \( \mathbb{E}^\Theta \left[ \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) + \int_0^1 |\dot{X}|^p dt \right] < \infty \), where \( p \) is from Assumption \((A7')\)

Due to this modification, the proofs in Section 9 have a few minor differences: Lemmas 9.1 and 9.2 should instead be proved under the assumption that there exists \( K > 0 \) such that

\[
\mathbb{E}^{Q^N} \left[ \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) + \int_0^1 |\dot{X}|^p dt \right] < K,
\]

where here we are using the modified coordinate process notation of Equation 47. The proofs follow in essentially the same manner using the uniform integrability granted by Assumption \((A8)\). Then, when proving Lemma 9.3, we see that in fact the bound offered in Equation 48 imply that Equation 49 holds for some \( K \) by using the same construction of a prelimit system for each \( Q^N \) given in Section 8 Fatou's Lemma, and Assumption \((A7')\).

Thus we have the following theorems:

Theorem 10.2. Under assumptions \((A1')-(A12')\), the sequence of \( \mathcal{P}(\mathcal{X}) \)-valued random variables \( \{\mu^N\} \) as defined by Equation 4 satisfies the Laplace Principal with good rate function

\[
I(\theta) = \inf_{\theta \in \mathcal{V}^\prime: \Theta = \theta} \mathbb{E}^\Theta \left[ \frac{1}{2} \int_{\mathbb{R}^m \times [0,1]} |z|^2 \rho(dzdt) \right]
\]

where \( \inf(\emptyset) := \infty \).

Theorem 10.3. Let \( ev: \mathcal{X} \to \mathbb{R}^d \) be the evaluation map at time \( t \) and \( \{\mu^N\} \) be as defined by Equation 2. Under assumptions \((A1')-(A12')\), \( \mathcal{L}(\mu^N) \to \delta_{\mu^*} \) in \( \mathcal{P}(\mathcal{P}(\mathcal{X})) \), where deterministic \( \mu^* \in \mathcal{P}(\mathcal{X}) \) satisfies \( \mu^* \circ ev^{-1}(t) = \mathcal{L}(X_i), t \in [0, 1] \) for \( X \) solving the McKean-Vlasov SDE 6.

11. Conclusions and Future Work

We have derived a large deviations principal and law of large numbers for the empirical measure of a system of weakly interacting particles in a two-scale environment in the joint many-particle and averaging limit. We use weak convergence methods, and obtain a variational form of the rate function. We saw that for the system 1, the two limiting procedures commute.

The results of this paper bring to light many interesting problems to be explored in future work. Of significant interest is making the discussion of Section 3 rigorous; in other words, proving the equivalence of rate functions in variational form, such as from Theorem 3.4 or 11, to rate functions of the form found in 10. A related problem is finding methods for solving constrained optimization problems over a space of measures where the constraint is determined by a class of joint laws for a controlled McKean-Vlasov SDE. This would allow for finding an explicit form of the rate function \( I \) in Theorem 3.4. Finally, an interesting extension of this work would be to proving a large deviations principal for systems whose coefficients depend on the “fast empirical measure” \( \mu_{N,\delta}^i := \frac{1}{N} \sum_{i=1}^N \delta_{X_i^N, N}/\delta \) as well. Such a result would then capture the system explored in 17, and would perhaps serve to give insight into the nature of bifurcations in the number of steady states for certain classes of McKean-Vlasov systems as originally investigated in 15.

Appendix A. Preliminary Results on the Prelimit System 7 and the Operator \( v_\theta(t) \)

Proposition A.1. Under assumptions \((A1)-(A8)\) the system of mean-field SDEs 1 admits a unique strong solution for each \( N \in \mathbb{N} \).

Proof. We observe that Equation 1 can be written as a standard 2dN-dimensional SDE via

\[
d\hat{X}^N_t = \left[ \frac{1}{\delta} f(\hat{X}^N_t) + \hat{b}(\hat{X}^N_t) \right] + \hat{\sigma}(\hat{X}^N_t)d\hat{W}^N_t
\]

where, letting \( Y^i_t = X^i_t/\delta, i \in \{1, \ldots, N\} \), and \( \hat{x} = (\hat{x}_1, \ldots, \hat{x}_{2N})^T, \hat{x}_i \in \mathbb{R}^d, i \in \{1, \ldots, 2N\} \), we have \( \hat{X}^N_t = (X^1_1, \ldots, X^N_1, Y^1_N, \ldots, Y^{N,N}_N) \).
Let $g$ play the role of $f, b$ or $\sigma$. For $g: \mathbb{R}^d \times \mathbb{T}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^j$, $j = d$ or $d \times m$, here we denote by

$$\hat{g}(x, y, \beta_1, \ldots, \beta_N) = g(x, y, \frac{1}{N} \sum_{i=1}^{N} \delta_{\beta_i}),$$

$$\forall x, \beta_1, \ldots, \beta_N \in \mathbb{R}^d, y \in \mathbb{T}^d.$$

One can then verify that (A5) then implies that for each $N \in \mathbb{N}, \exists C(N)$ such that for all $\hat{x}_1, \hat{x}_2 \in (\mathbb{R}^d)^{2N}$,

$$\frac{1}{N} f(\hat{x}_1) + b(\hat{x}_1) - \frac{1}{N} f(\hat{x}_2) - b(\hat{x}_2) + |\sigma(\hat{x}_1) - \sigma(\hat{x}_2)| \leq C(N)|\hat{x}_1 - \hat{x}_2|,$$

so that by standard existence and uniqueness results for SDE’s with globally Lipschitz coefficients, the proposition holds. See, for example, Theorem 5.2.1 in [23].

**Proposition A.2.** For $\bar{X}^{i,N}$ as in Equation (4) controlled by any $u^N \in \mathcal{U}_N$ satisfying almost surely for some $B > 0$ the bound $\sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{1} |u_i^N(t)|^2 dt < B$,

$$\sup_{N \in \mathbb{N}} \mathbb{E} \left[ \left( \sup_{0 \leq t \leq 1} \frac{1}{N} \sum_{i=1}^{N} |\bar{X}_t^{i,N}|^2 \right)^q \right] < \infty, \forall q > 0.$$

This ensures that for all $N$, there exists a modification of $\bar{\mu}^N \in C([0,1]; \mathcal{P}_2(\mathbb{R}^d))$ so that $\bar{\mu}^N_t$ is in the domain of the coefficients and $\Phi$ for all time.

**Proof.** It clearly suffices to prove the result assuming $q \geq 1$. We use the same computations as in the proof of tightness of $Q^N_X$ in Subsection 6.1.3 but taking $h(x) = |x|^2$.

Then, by Equation (27) in Subsection 6.1.3 we get

$$|\bar{X}^{i,N}|^p = |x^{i,N}|^p + \sum_{k=1}^{8} B_k^{i,N}(t).$$

Using that for any $p$,

$$|\nabla h(x)| = p|x|^{p-1}, \quad |\nabla \nabla h(x)| = p(p^2 - 2p + d)^{1/2}|x|^{p-2}$$

and when $p = 2$, we have $\nabla \nabla h_{x_i} = 0, l \in \{1, \ldots, d\}$, we apply Cauchy Schwarz to all the inner products with $\nabla h, \nabla \nabla h$, and $\nabla \nabla h_{x_i}$ in $B_k^{i,N}(t), k = 1, 2, 4, 5, 6, 8$ and, using Assumption [A1] and Proposition C.2.

$$\mathbb{E} \left[ \left( \sup_{0 \leq t \leq 1} \frac{1}{N} \sum_{i=1}^{N} |\bar{X}_t^{i,N}|^2 \right)^q \right]$$

$$\leq \mathbb{E} \left[ \sup_{t \in [0,1]} \left( \frac{1}{N} \sum_{i=1}^{N} \left( |x^{i,N}|^2 + C(d) |x^{i,N}| + |\bar{X}_t^{i,N}| + \int_{0}^{t} 1 + |\bar{X}_s^{i,N}| + |u_i^N(s)||\bar{X}_s^{i,N}| + |u_i^N(s)|ds \right) \right]^q$$

$$+ \frac{1}{N} \sum_{j=1}^{N} \int_{0}^{t} u_{j}^N(s)||\bar{X}_s^{i,N}| ds$$

$$+ \int_{0}^{t} \bar{X}_s^{i,N} \cdot \left[ \frac{1}{N} \sum_{j=1}^{N} \partial_{\Phi}(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N}) \sigma(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N})dw_s^j \right]$$

$$+ \int_{0}^{t} \bar{X}_s^{i,N} \cdot \left[ \left[ I + \delta \nabla_x \Phi(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N}) + \nabla \Phi(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N}) \right] \sigma(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N})dw_s^i \right]$$

$$+ \int_{0}^{t} \text{tr} \left( \Phi(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N}) \otimes \sigma(\bar{X}_s^{i,N}, \bar{X}_s^{i,N}/\delta, \bar{\mu}_s^{N})dw_s^i \right) \right)^q.$$
\[ \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^{N} |\bar{X}_t^{i,N}|^2 \right] \]

\[ \leq C_2(d, q) \left( 1 + \mathbb{E} \left[ \left( \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^{N} \left\{ \int_0^t 1 + |\bar{X}_s^{i,N}| + |u_s^{N}(s)| + |\bar{X}_s^{i,N}| + |u_i^{N}(s)|ds \right\} \right] \right) \]

\[ + \mathbb{E} \left[ \left( \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^{N} \left\{ \int_0^t |u_i^{N}(s)||\bar{X}_s^{i,N}|ds \right\} \right]^q \right] \]

\[ \leq C_3(d, q) \left( 1 + \mathbb{E} \left[ \left( \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^{N} \left\{ \int_0^t 1 + |\bar{X}_s^{i,N}| + |u_s^{N}(s)||\bar{X}_s^{i,N}| + |u_i^{N}(s)|ds \right\} \right]^{1/2} \right) \]

\[ + \mathbb{E} \left[ \left( \frac{1}{N^2} \sum_{i=1}^{N} \int_0^t |\bar{X}_s^{i,N}|^2ds \right)^{q/2} \right] \]

\[ \leq C_4(d, q) \left( 1 + \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^{N} \int_0^t |\bar{X}_s^{i,N}|^2dt \right)^{q/2} \right] \right) \]

\[ + \mathbb{E} \left[ \left( \frac{1}{N^2} \sum_{i=1}^{N} \int_0^t |\bar{X}_s^{i,N}|^2ds \right)^{q/2} \right] \]

\[ \leq C_5(d, q) \left( 1 + \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^{N} \int_0^t \sup_{s \in [0,t]} |\bar{X}_s^{i,N}|^2 dt \right)^{q/2} \right] \right) \]

\[ \leq C_6(d, q) \left( 1 + \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} |\bar{X}_s^{i,N}|^2 \right)^{q/2} \right] \right) \]

\[ \leq C_7(d, q) \left( 1 + \frac{1}{2} \int_0^t \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} |\bar{X}_s^{i,N}|^2 \right)^{q} \right] dt + \frac{1}{2} \left( 1 + B^{q/2} \right)^2 \]

by Young’s inequality, Tonelli’s Theorem, and Jensen’s inequality. By Grönwall’s inequality, we can conclude the proof of the proposition. \[ \square \]

We end this section with a proposition regarding the mapping defined in Equation \[11\]

**Proposition A.3.** For fixed \( t \in [0,1] \), \( Q \mapsto \nu_Q(t) \) is continuous.

**Proof.** Take \( \{Q^n\} \subset \mathcal{P}(W) \) such that \( Q^n \to Q \) and \( f \in C_b(\mathbb{R}^d) \). Then, since \( (\phi, n, r) \mapsto f(\phi(t)) \in C_b(W) \) we get

\[ \lim_{n \to \infty} \int_{\mathbb{R}^d} f(x)\nu_{Q^n}(t)(dx) = \lim_{n \to \infty} \int_{W} f(\phi(t))Q^n(d\phi d\nu d\mathcal{H}) = \int_{W} f(\phi(t))Q(d\phi d\mathcal{H}) = \int_{\mathbb{R}^d} f(x)\nu_Q(t)(dx). \]

\[ \square \]
Appendix B. On Lions Differentiation

We will need the following two definitions from [12]:

Definition B.1. Given a function \( u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \), we may define a lifting of it to \( \tilde{u} : L^2(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}; \mathbb{R}^d) \to \mathbb{R} \) via \( \tilde{u}(X) = u(\mathcal{L}(X)) \) for \( X \in L^2(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}; \mathbb{R}^d) \). Here we assume \( \hat{\Omega} \) is a Polish space, \( \hat{\mathcal{F}} \) its Borel \( \sigma \)-field, and \( \hat{\mathbb{P}} \) is an atomless probability measure (since \( \hat{\mathbb{P}} \) is Polish, this is equivalent to every singleton having zero measure).

Here, denoting by \( \mu(\cdot | \cdot) := \int_{\mathbb{R}^d} |x|^r \mu(dx) \) for \( r > 0 \),

\[
\mathcal{P}_2(\mathbb{R}^d) := \{ \mu \in \mathcal{P}(\mathbb{R}^d) : \mu(\cdot | \cdot)^{1/2} = \int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty \}.
\]

\( \mathcal{P}_2(\mathbb{R}^d) \) is a Polish space under the \( L^2 \)-Wasserstein distance

\[
\mathcal{W}_2(\mu_1, \mu_2) := \inf_{\pi \in \mathcal{C}_{\mu_1, \mu_2}} \left[ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) \right]^{1/2},
\]

where \( \mathcal{C}_{\mu_1, \mu_2} \) denotes the set of all couplings of \( \mu_1, \mu_2 \).

We say \( u \) is \textit{L-differentiable} or \textit{Lions-differentiable} at \( \mu_0 \in \mathcal{P}_2(\mathbb{R}^d) \) if there exists a random variable \( X_0 \) on some \( \left( \hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}} \right) \) satisfying the above assumptions such that \( \mathcal{L}(X_0) = \mu_0 \) and \( \tilde{u} \) is Fréchet differentiable at \( X_0 \).

The Fréchet derivative of \( \tilde{u} \) can be viewed as an element of \( L^2(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}; \mathbb{R}^d) \) by identifying \( L^2(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}; \mathbb{R}^d) \) and its dual. From this, one can find that if \( u \) is L-differentiable at \( \mu_0 \in \mathcal{P}_2(\mathbb{R}^d) \), there is a deterministic measurable function \( \xi : \mathbb{R}^d \to \mathbb{R}^d \) such that \( D\tilde{u}(X_0) = \xi(X_0) \), and that \( \xi \) is uniquely defined \( \mu_0 \)-almost everywhere on \( \mathbb{R}^d \). We denote this equivalence class of \( \xi \in L^2(\mathbb{R}^d, \mu_0; \mathbb{R}^d) \) by \( \partial_\mu u(\mu_0) \) and call \( \partial_\mu u(\mu_0)(\cdot) : \mathbb{R}^d \to \mathbb{R}^d \) the \textit{Lions derivative} of \( u \) at \( \mu_0 \). Note that this definition is independent of the choice of \( X_0 \) and \( \left( \hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}} \right) \). See [12] Section 5.2.

To avoid confusion when \( u \) depends on more variables than just \( \mu \), if \( \partial_\mu u(\mu_0) \) is differentiable at \( v_0 \in \mathbb{R}^d \), we denote its derivative at \( v_0 \) by \( \partial_\nu \partial_\mu u(\mu_0)(v_0) \).

Definition B.2. ([12] Definition 5.83) We say \( u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \) is \textit{Fully C}^2 if the following conditions are satisfied:

1. \( u \) is \( C^1 \) in the sense of L-differentiation, and its first derivative has a jointly continuous version \( \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \ni (\mu, v) \mapsto \partial_\mu u(\mu)(v) \in \mathbb{R}^d \).
2. For each fixed \( \mu \in \mathcal{P}_2(\mathbb{R}^d) \), the version of \( \mathbb{R}^d \ni v \mapsto \partial_\mu u(\mu)(v) \in \mathbb{R}^d \) from the first condition is differentiable on \( \mathbb{R}^d \) in the classical sense and its derivative is given by a jointly continuous function \( \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \ni (\mu, v) \mapsto \partial_{\nu, \mu} u(\mu)(v) \in \mathbb{R}^{d \times d} \), with \( \mathbb{L}(X) = \mu, \partial^2 u(\mu)(v)(X) \) gives the Fréchet derivative at \( X \) of \( L^2(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}; \mathbb{R}^d) \ni X' \mapsto \partial_\nu u(\mathcal{L}(X'))(v) \) for every \( v \in \mathbb{R}^d \). Denoting \( \partial^2_{\mu} u(\mu)(v)(v') \) by \( \partial^2_{\nu} u(\mu)(v, v') \), the map \( \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \ni (\mu, v, v') \mapsto \partial^2_{\nu} u(\mu)(v, v') \) is also assumed to be continuous in the product topology.

We recall now a useful connection between the Lion derivative as defined in [12] and the empirical measure.

Proposition B.3. For \( g : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \) which is Fully C^2 in the sense of definition B.2 we can define the empirical projection of \( g \), as \( g^N : (\mathbb{R}^d)^N \to \mathbb{R}^d \) given by

\[
g^N(\beta_1, ..., \beta_N) := g\left(\frac{1}{N} \sum_{i=1}^{N} \delta_{\beta_i}\right).
\]

Then \( g^N \) is twice differentiable on \( (\mathbb{R}^d)^N \), and for each \( \beta_1, ..., \beta_N \in \mathbb{R}^d \), \( (i, j) \in \{1, ..., N\}^2 \), \( l \in \{1, ..., d\} \)

\[
\nabla_{\beta_l} g^N(\beta_1, ..., \beta_N) = \frac{1}{N} \partial_{\nu} g_l(\frac{1}{N} \sum_{i=1}^{N} \delta_{\beta_i})(\beta_i)
\]

\[
(51)
\]
and
\begin{equation}
\nabla_{\beta_i} \nabla_{\beta_j} g^N(x, y, \beta_1, \ldots, \beta_N) = \frac{1}{N} \partial_{\beta_i} \partial_{\beta_j} g(\frac{1}{N} \sum_{i=1}^{N} \delta_{\beta_i}(\beta_i) \mathbb{1}_{i=j}) + \frac{1}{N^2} \partial_{\mu}^2 g(\frac{1}{N} \sum_{i=1}^{N} \delta_{\beta_i}(\beta_i, \beta_j)).
\end{equation}

In particular, this holds for \( \Phi(x, y, \cdot) \) for fixed \( x \in \mathbb{R}^d \) and \( y \in \mathbb{T}^d \).

**Proof.** This follows from Propositions 5.35 and 5.91 of \([12]\). Since by Proposition C.2 \( \Phi \) is fully \( C^2 \), it applies to \( \Phi(x, y, \cdot) \).

\[ \square \]

**APPENDIX C. ON THE OPERATOR \( \mathcal{L} \) AND RELATED PDEs**

**Proposition C.1.** Under assumptions \([A1] \) \([A8] \), the invariant measure \( \pi \) defined by Equation 1 is uniquely determined for each \( x \in \mathbb{R}^d \) and \( \mu \in \mathcal{P}_2(\mathbb{R}^d) \) and has a density \( \bar{\pi} \).

**Proof.** This follows immediately from Theorem 6.16 in \([52]\).

\[ \square \]

**Proposition C.2.** Under assumptions \([A1] \) \([A8] \), there is a unique strong solution \( \Phi \) to equation 2. Moreover, \( \Phi \) is fully \( C^2 \) in the sense of Definition 3.3 and \( \Phi \), all second order partial derivatives of \( \Phi \) in \( x \) and \( y \), and \( \partial_{\mu} \Phi(x, y, \mu)(\cdot), \partial_{\mu,\mu} \Phi(x, y, \mu)(\cdot), \nabla_{x,y} \partial_{\mu} \Phi(x, y, \mu)(\cdot), \nabla_{y} \partial_{\mu} \Phi(x, y, \mu)(\cdot) \) exist, are continuous with respect to all variables \( x, v, v' \in \mathbb{R}^d, y \in \mathbb{T}^d, \mu \in \mathcal{P}(\mathbb{R}^d) \), and are uniformly bounded \( L^2(\mathbb{R}^d, \mu) \) with respect to \( x \) and \( y \).

**Proof.** Existence and uniqueness follows directly from Theorems 6.16 and 7.9 in \([52]\).

Consider the frozen process on \( \mathbb{T}^d \) for fixed \( x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d) \), \( y \in \mathbb{T}^d \), given by
\begin{align}
dY_t^{x,y,\mu} &= f(x, Y_t^{x,y,\mu}, \mu)dt + \sigma(x, Y_t^{x,y,\mu}, \mu)d\tilde{W}_t, \\
y_0^{x,y,\mu} &= y
\end{align}

where \( \tilde{W}_t \) is a \( d \)-dimensional, \( \tilde{F}_t \)-adapted Brownian motion on some probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})\) satisfying the usual conditions.

As per Proposition 4.1 in \([54]\) and Section 11.6 in \([52]\), \( \Phi \) is given by
\begin{equation}
\Phi(x, y, \mu) = \int_0^\infty \mathbb{E}[f(x, Y_s^{x,y,\mu}, \mu)]ds.
\end{equation}

Then the fact that \( \Phi \) is fully \( C^2 \) and smooth in \( x \) and \( y \) and boundedness of \( \Phi \), along with regularity of \( \Phi \) of the same type given in \([A3] \) \([A4] \) follows from the unique representation of the cell problem given by Equation 53 and the regularity assumptions on the coefficients given by \([A3] \) \([A4] \) and \([A7] \) (see, for example \([50] \), \([51] \) for general results on Euclidean space with no measure dependence. \([4] \) Chapter 3 Section 6 for the case where the fast component is on the torus with no measure dependence, as well as \([54] \) for when \( \Phi \) depends on a measure). Differentiability in \( x \) or \( y \) can also be seen directly from standard interior regularity theory for elliptic PDEs (see \([36] \), \([28] \), and Proposition 5.1 in \([13] \) ). Though second differentiability in \( \mu \) isn’t addressed in \([54] \), it follows by applying the same direct differentiation to the transition density of the semigroup associated to Equation 33 twice. See \([13] \) and \([13] \) for an example of where estimates on the second derivative of the transition density in \( \mu \) are derived under different, weaker assumptions. In addition, in \([54] \) the fast component of the process is on Euclidean space, so they make use of a certain confining assumption on the coefficients to prove Lemma 3.6. In our case this assumption is not needed, as the analogous ergodicity result to Lemma 3.6 is automatically given on a compact space via Theorem 6.16 in \([52] \). For a proof see Chapter 3 Section 3 in \([4] \).  

\[ \square \]

**Corollary C.3.** For fixed \( x, v \in \mathbb{R}^d \) and \( y \in \mathbb{T}^d \), \( \mu \mapsto g(x, y, \mu) \) is Lipschitz continuous in \( (\mathcal{P}_2(\mathbb{R}^d), \mathbb{W}_2) \) for \( g = \Phi, \nabla_x \Phi, \nabla_y \Phi, \) or \( \nabla_x \nabla_y \Phi \).

**Proof.** From Proposition C.2 we have \( \partial_{\mu} g(x, y, \mu)(\cdot) \) is bounded in \( L^2(\mathbb{R}^d, \mu) \) for all \( x \in \mathbb{R}^d, y \in \mathbb{T}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d) \). This implies that \( g(x, y, \cdot) \) is Lipschitz continuous with respect to \( \mathbb{W}_2 \) for each \( x \in \mathbb{R}^d \) and \( y \in \mathbb{T}^d \) by Remark 5.27 in \([12] \).  

\[ \square \]
Here we present the proof of Proposition 5.1.

**Proof of Proposition 5.1.** Applying Itô’s formula, we get

\[
\frac{1}{N} \sum_{i=1}^{N} (\tilde{X}_t^{i,N})^4 = \frac{1}{N} \sum_{i=1}^{N} \left\{ (\tilde{X}_0^{i,N})^4 + \int_0^t \left[ -(\tilde{X}_s^{i,N})^3 + \tilde{X}_s^{i,N} + \sigma u_i^N(s) - \kappa(\tilde{X}_s^{i,N} - \nu_s^N) \right] \right. \\
- \frac{\epsilon}{\delta} \sin(2\pi \tilde{X}_t^{i,N}/\delta) \left. \right\} 4(\tilde{X}_s^{i,N})^3 + 6\sigma^2(\tilde{X}_s^{i,N})^2 ds + \int_0^t \sigma(\tilde{X}_s^{i,N})^3 dW_s \right\}.
\]

Now considering \( \psi(x, y) = x^3 \Phi(y) \), we have that \( \psi \) satisfies

\[-\epsilon \sin(2\pi y) \frac{\partial}{\partial y} \psi(x, y) + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial y^2} \psi(x, y) = \epsilon \sin(2\pi y)x^3.\]

Letting \( \tilde{Y}_t^{i,N} = \tilde{X}_t^{i,N} / \delta \) and applying Itô’s formula to \( \psi(\tilde{X}_t^{i,N}, \tilde{Y}_t^{i,N}) \), we get

\[
\psi(\tilde{X}_t^{i,N}, \tilde{Y}_t^{i,N}) = \psi(\tilde{X}_0^{i,N}, \tilde{Y}_0^{i,N}) + \int_0^t \left[ \left[ -(\tilde{X}_s^{i,N})^3 + \tilde{X}_s^{i,N} + \sigma u_i^N(s) - \kappa(\tilde{X}_s^{i,N} - \nu_s^N) \right] \left( \tilde{X}_s^{i,N})^3 \Phi'(\tilde{Y}_s^{i,N}) \right] \right. \\
+ \frac{1}{\delta} \left[-(\tilde{X}_s^{i,N})^3 + \tilde{X}_s^{i,N} + \sigma u_i^N(s) - \kappa(\tilde{X}_s^{i,N} - \nu_s^N) \right] \left( \tilde{X}_s^{i,N})^3 \right) ds + \int_0^t 3(\tilde{X}_s^{i,N})^2 \Phi'(\tilde{Y}_s^{i,N}) ds \\
+ \frac{\sigma^2}{\delta} \left(\tilde{X}_s^{i,N})^3 \right) dW_s + \frac{1}{\delta^2} \int_0^t \epsilon \sin(2\pi \tilde{Y}_t^{i,N})(\tilde{X}_s^{i,N})^3 ds.
\]

Solving for \( \frac{1}{\delta^2} \int_0^t \epsilon \sin(2\pi \tilde{Y}_t^{i,N})(\tilde{X}_s^{i,N})^3 ds \) and plugging this into our equation for \( \frac{1}{N} \sum_{i=1}^{N} (\tilde{X}_t^{i,N})^4 \), we get

\[
\frac{1}{N} \sum_{i=1}^{N} (\tilde{X}_t^{i,N})^4 = \frac{1}{N} \sum_{i=1}^{N} \left\{ (\tilde{X}_0^{i,N})^4 + 4\delta(\tilde{X}_0^{i,N})^3 \Phi(\tilde{Y}_0^{i,N}) - 4\delta(\tilde{X}_t^{i,N})^3 \Phi(\tilde{Y}_t^{i,N}) \right. \\
+ \int_0^t \left[ \left[ -(\tilde{X}_s^{i,N})^3 + \tilde{X}_s^{i,N} + \sigma u_i^N(s) - \kappa(\tilde{X}_s^{i,N} - \nu_s^N) \right] 4(\tilde{X}_s^{i,N})^3 + 6\sigma^2(\tilde{X}_s^{i,N})^2 \right] ds \right. \\
+ 12\delta \left[ -(\tilde{X}_s^{i,N})^3 + \tilde{X}_s^{i,N} + \sigma u_i^N(s) - \kappa(\tilde{X}_s^{i,N} - \nu_s^N) \right] \left( \tilde{X}_s^{i,N})^3 \right) ds \\
+ 12\delta \tilde{X}_s^{i,N} \Phi(\tilde{Y}_s^{i,N}) \sigma^2 + 12\sigma^2(\tilde{X}_s^{i,N})^2 \Phi'(\tilde{Y}_s^{i,N}) \right. ds \\
+ \left. \int_0^t \left( \sigma(1 + \Phi'(\tilde{Y}_s^{i,N}))(\tilde{X}_s^{i,N})^3 + 12\delta(\tilde{X}_s^{i,N})^2 \Phi(\tilde{Y}_s^{i,N}) \right) dW_s \right\}.
\]
First applying Young's inequality to $4\delta(\bar{X}_0^{i,N})^3\Phi(\bar{Y}_0^{i,N})$, we get $|4\delta(\bar{X}_0^{i,N})^3\Phi(\bar{Y}_0^{i,N})| \leq \frac{3}{4}(\bar{X}_0^{i,N})^4 + 4^3(\delta\Phi(\bar{Y}_0^{i,N}))^4$. Similarly, $|4\delta(\bar{X}_1^{i,N})^3\Phi(\bar{Y}_1^{i,N})| \leq \frac{3}{4}(\bar{X}_1^{i,N})^4 + 4^3(\delta\Phi(\bar{Y}_1^{i,N}))^4$. So

$$\frac{1}{4}E\left[\sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N (\bar{X}_t^{i,N})^4\right]$$

$$\leq C(D, K, \delta) + E\left[\sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N \int_0^t \left(\left(-\bar{X}_s^{i,N})^3 + \bar{X}_s^{i,N} + \sigma u_s^N(s) - \kappa(\bar{X}_s^{i,N} - \nu_s^N)\right)\right)\Phi(\bar{Y}_s^{i,N})^2\right] ds$$

Now we use the explicit form of $\Phi'(y)$ to get a crucial lower bound. We have

$$\Phi'(y) = -1 + \exp\left(-\frac{\epsilon}{\pi \sigma^2} \cos(2\pi y)\right) / Z > -1.$$
Now by Young’s and Jensen’s inequalities (in particular using repeatedly that $(\nu_t^N)^2 \leq \frac{1}{N} \sum_{k=1}^N (X_t^{k,N})^2$) as well as Burkholder-Davis Gundy inequality on the martingale terms,

\[
\frac{1}{4} \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N (X_t^{i,N})^4 \right] \\
\leq C(D, K, \delta) + \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N \left\{ \int_0^t \left( -4\gamma (\bar{X}_s^{i,N})^6 + 4(1 + K - \gamma \kappa) (\bar{X}_s^{i,N})^4 \right) \right\} \right] \\
+ \frac{3}{2\alpha} \frac{1}{16} \left[ (1 + K)^2 |\sigma u_t^{N}(s)|^2 + [(1 + K)\kappa] (\bar{X}_s^{i,N})^2 + 9(\delta (1 + \kappa) K)^2 \right] + \frac{\alpha}{2} |\bar{X}_s^{i,N}|^6 \\
+ \frac{5}{2\beta} \frac{1}{36} \left[ \sigma^4 + 4|\delta K|^2 |\sigma u_t^{N}(s)|^2 + 4|\kappa K|^2 (\bar{X}_s^{i,N})^2 + 4|\varepsilon K|^2 + 4|\sigma K|^2 \right] + \frac{\beta}{2} (\bar{X}_s^{i,N})^4 \\
+ 12\delta (\bar{X}_s^{i,N})^5 K + \frac{1}{2} (12\delta K \sigma^2)^2 + \frac{1}{2} (\bar{X}_s^{i,N})^2 ds + 4\sigma \left( \int_0^t (1 + K^2) (\bar{X}_s^{i,N})^6 ds \right)^{1/2} \\
+ 12\delta \sigma \left( \int_0^t K^2 (\bar{X}_s^{i,N})^4 ds \right)^{1/2} \right] \\
\leq C(D, K, \delta) + \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N \left\{ \int_0^t \left( -4\gamma (\bar{X}_s^{i,N})^6 + 4(1 + K - \gamma \kappa) (\bar{X}_s^{i,N})^4 \right) \right\} \right] \\
+ \frac{3}{2\alpha} \frac{1}{16} \left[ (1 + K)^2 |\sigma u_t^{N}(s)|^2 + [(1 + K)\kappa] (\bar{X}_s^{i,N})^2 + 9(\delta (1 + \kappa) K)^2 \right] + \frac{\alpha}{2} |\bar{X}_s^{i,N}|^6 \\
+ \frac{5}{2\beta} \frac{1}{36} \left[ \sigma^4 + 4|\delta K|^2 |\sigma u_t^{N}(s)|^2 + 4|\kappa K|^2 (\bar{X}_s^{i,N})^2 + 4|\varepsilon K|^2 + 4|\sigma K|^2 \right] + \frac{\beta}{2} (\bar{X}_s^{i,N})^4 \\
+ 12\delta (\bar{X}_s^{i,N})^5 K + \frac{1}{2} (12\delta K \sigma^2)^2 + \frac{1}{2} (\bar{X}_s^{i,N})^2 ds + \frac{4\sigma^2}{2\alpha} + \frac{\alpha}{2} \int_0^t (1 + K^2) (\bar{X}_s^{i,N})^6 ds \\
+ \frac{12\delta \sigma^2}{2\beta} + \frac{\beta}{2} \int_0^t K^2 (\bar{X}_s^{i,N})^4 ds \right\},
\]

where $\alpha, \beta > 0$ are to be chosen. To simplify things, since $\delta \downarrow 0$ as $N \to \infty$ and all terms are increasing in $\delta$ (including $C(D, K, \delta)$), we may assume $N$ is large enough that $\delta \leq 1$. This, along with our bound on $\mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \int_0^t |u_t^{N}(t)|^2 dt \right]$ gives

\[
\frac{1}{4} \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N (X_t^{i,N})^4 \right] \\
\leq C(D, K) + \mathbb{E} \left[ \sup_{t \in [0,1]} \frac{1}{N} \sum_{i=1}^N \left\{ \int_0^t \left( -4\gamma (\bar{X}_s^{i,N})^6 + 4(1 + K - \gamma \kappa) (\bar{X}_s^{i,N})^4 \right) \right\} \right] \\
+ \frac{3}{2\alpha} \frac{1}{16} \left[ (1 + K)^2 |\sigma u_t^{N}(s)|^2 + [(1 + K)\kappa] (\bar{X}_s^{i,N})^2 + 9(\delta (1 + \kappa) K)^2 \right] + \frac{\alpha}{2} |\bar{X}_s^{i,N}|^6 \\
+ \frac{5}{2\beta} \frac{1}{36} \left[ \sigma^4 + 4|\delta K|^2 |\sigma u_t^{N}(s)|^2 + 4|\kappa K|^2 (\bar{X}_s^{i,N})^2 + 4|\varepsilon K|^2 + 4|\sigma K|^2 \right] + \frac{\beta}{2} (\bar{X}_s^{i,N})^4 \\
+ 12\delta (\bar{X}_s^{i,N})^5 K + \frac{1}{2} (12\delta K \sigma^2)^2 + \frac{1}{2} (\bar{X}_s^{i,N})^2 ds + \frac{4\sigma^2}{2\alpha} + \frac{\alpha}{2} \int_0^t (1 + K^2) (\bar{X}_s^{i,N})^6 ds \\
+ \frac{12\delta \sigma^2}{2\beta} + \frac{\beta}{2} \int_0^t K^2 (\bar{X}_s^{i,N})^4 ds \right\].
\]
\[ = C(D, K) + \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} \left\{ \int_{0}^{t} \left( \alpha(1 + K^2/2) - 4\gamma\right)(\bar{X}_{i,s}^{i,N})^6 \right. \right. \]
\[ + \left. \left. (\beta(1 + K^2)/2 + 4(1 + K) - 4\gamma\kappa)(\bar{X}_{i,s}^{i,N})^4 \right. \right. \]
\[ + \frac{3}{2\alpha} \left[\left( (1 + K)\sigma^2 B + [(1 + K)\kappa]^2 (\bar{X}_{i,s}^{i,N})^2 + 9[(1 + \kappa)K]^2 \right] \right. \]
\[ + \frac{5}{2\beta} 36 \left[ \sigma^4 + 4[\sigma K] B + 4[\kappa K]^2 (\bar{X}_{i,s}^{i,N})^2 + 4[\epsilon K]^2 + 4[\sigma^2 K]^2 \right] \]
\[ + 12|\bar{X}_{i,s}^{i,N}|^5 K + \frac{1}{2}[12K\sigma^2] + \frac{1}{2}(\bar{X}_{i,s}^{i,N})^2 ds + \frac{1}{2\alpha} (4\sigma^2 + \frac{1}{2\beta} [12\sigma^2] ] \right\}. \]

Letting \( \alpha = 3\gamma/(1 + K^2/2) \), \( \beta = 8\gamma\kappa/(1 + K^2) \), we get
\[ \frac{1}{4} \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 \right] \leq C(D, K) + \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} \left\{ \int_{0}^{t} -\gamma(\bar{X}_{i,s}^{i,N})^6 + 4(1 + K)(\bar{X}_{i,s}^{i,N})^4 \right. \right. \]
\[ + C(K, \kappa)(\bar{X}_{i,s}^{i,N})^2 + 12K|\bar{X}_{i,s}^{i,N}|^5 ds \right\} + C(K, \sigma, \kappa, \epsilon, B, c, \gamma), \]
where we used \( t \leq 1 \) to pull the constant out of the integral. Once again by Young’s inequality, we have
\[ C(K, \kappa)x^2 \leq \frac{C(K, \kappa)^2}{2} + \frac{1}{2}x^4 \text{ and } 12K|x|^5 \leq (12K/\gamma)^6 + \frac{2}{3a}x^6 \text{ for any } \alpha > 0. \]
Then letting \( \alpha = (3\gamma/5)^{5/6} \), we get
\[ \frac{1}{4} \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 \right] \leq C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma). \]

Now we have
\[ \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 \right] \leq 4C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) + 4C_2(K, \kappa) \int_{0}^{t} \mathbb{E}\left[ \sup_{s \in [0, t]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 ds \right] \]
\[ \leq 4C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) + 4C_2(K, \kappa) \int_{0}^{t} \mathbb{E}\left[ \sup_{s \in [0, t]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 \right] dt, \]
so by Grönwall’s inequality,
\[ \mathbb{E}\left[ \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_{i,s}^{i,N})^4 \right] \leq 4C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) \exp \left( 4C_2(K, \kappa) \right) < \infty. \]

Also, by performing the exact same proof without the supremum over time in the expectation, (using Jensen’s inequality and Itô Isometry instead of Burkholder-Davis-Gundy inequality for the martingale terms), we get from Equation (55) that for each \( t \in [0, 1] \),
\[ \gamma \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} (\bar{X}_{i,s}^{i,N})^6 ds \right] \leq 2C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) + 2C_2(K, \kappa) \mathbb{E}\left[ \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} (\bar{X}_{i,s}^{i,N})^4 ds \right] \]
\[ \leq 2C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) + 8C_2(K, \kappa) C(D, K, \sigma, \kappa, \epsilon, B, c, \gamma) \exp \left( 4C_2(K, \kappa) \right) \text{ by Equation (56)} \]
\[ < \infty. \]

Recalling \( K \) and \( \gamma \) are just a fixed constants from the bounds on \( \Phi \) and its derivatives and \( c \) is the fixed constant from applying Burkholder-Davis-Gundy inequality, we are done. \( \square \)
We begin with the proof of Theorem 5.2.

**Proof.** As a consequence of (V.4) we get for $\Theta \in \hat{V}$,

\[
\hat{\Theta}\left(\left\{(\varphi, \hat{r}) \in \hat{W}: \hat{r}_t(dydz) = \gamma(dy|y, t)\pi(dy|\varphi(t), \hat{\nu}_\Theta(t)) \text{ for some stochastic kernel } \gamma, \forall t \in [0, 1]\right\}\right) = 1.
\]

This follows in essentially the same way as in Remark 6.7. Then we can write Equation 21 as:

\[
(57)
\]

As a consequence of (57) and (58) we can construct from $\Theta \in \dot{V}$ for any $\hat{\Theta}$ under $\hat{\Theta}$.

We begin with the proof of Theorem 5.2.

To prove the equivalence of $I$ given in Theorems 5.2 and 10.2 and $\hat{I}$ given in Theorem 5.2, we prove that given any $\Theta \in \dot{V}$, we can construct from $\Theta \in V$ some $\hat{\Theta} \in \hat{V}$ with the expression inside the infimum in the definition of $\hat{I}$ less than the expression inside the infimum in the definition of $I$, and vice versa. Then this construction holds for any $\Theta$ or $\hat{\Theta}$ where the infimum is attained, and the result will be proved.

Fix $\Theta \in \dot{V}$ such that $I(\Theta) < \infty$. Given $\Theta \in V$ such that $\Theta_X = \theta$, we construct $\hat{\Theta} \in \hat{V}$ with $\hat{\Theta}_X = \hat{\theta}$ such that

\[
E^\Theta\left[\frac{1}{2} \int_{x \times [0, 1]} |z|^2 \hat{\rho}(dydz)dt\right] \leq E^{\hat{\Theta}}\left[\frac{1}{2} \int_{x \times [0, 1]} |z|^2 \rho(dydz)dt\right].
\]

To do this, simply set

\[\hat{\Theta}_{\hat{\gamma}}(B) = \Theta_\gamma\left\{r \in \mathbb{Z} : \exists \hat{r} \in B \text{ where } \hat{r}(dydz) = r(dydz)\pi(dy)\right\}\]

for all $B \in \hat{V}$.

What we are really doing here is embedding $\mathbb{Z}$ and $\hat{V}$ into the product measure space $(\mathbb{Z} \times \hat{V}, \mathcal{B}(\mathbb{Z} \times \hat{V})$ and defining a probability measure $\hat{\Theta}$ on this space by

\[
\hat{\Theta}(E \times \mathcal{F}) = \int_E \int_F \delta_{p(r)}(dr)\Theta_{\mathbb{Z}}(dr)
\]

for $E \times \mathcal{F} \in \mathcal{B}(\mathbb{Z} \times \hat{V})$ and $p: \mathbb{Z} \to \hat{V}$ defined by $p(r(dydz)) = r(dydz)\otimes \pi(dy)$. Then, $\hat{\Theta}$ has first marginal $\Theta_{\mathbb{Z}}$ and we define its second marginal to be $\Theta_{\hat{\gamma}}$. Due to the fact that $p$ is injective, we can quickly see that $\hat{\Theta}$ is indeed a probability measure on the product space, so that this is a well-defined way of constructing $\hat{\Theta}_{\hat{\gamma}}$.

Then the equality in distribution

\[
\sigma \int_0^s \int_{T \times \mathbb{R}} z \left[\Phi(y) + 1\right] \hat{\rho}_t(dydz)dt = \sigma \int_0^s \int_T \left[\Phi(y) + 1\right] \pi(dy) \int_{\mathbb{R}} z \rho_t(dy)dt
\]

\[
= \frac{\sigma}{ZZ} \int_0^s \int_{\mathbb{R}} z \rho_t(dy)dt \text{ by Equation 20}
\]

holds for all $s \in [0, 1]$. In other words, the dynamics of Equation 58 under $\hat{\Theta}$ and Equation 59 under $\Theta$ are the same. The fact that $\hat{\Theta} \in \hat{V}$ then follows from the fact that $\Theta \in V$. In addition,

\[
E^{\hat{\Theta}}\left[\frac{1}{2} \int_{x \times [0, 1]} |z|^2 \hat{\rho}(dydz)dt\right] = E^{\hat{\Theta}}\left[\frac{1}{2} \int_T \int_{x \times [0, 1]} |z|^2 \rho(dydz)\pi(dy)\right]
\]

\[
= E^{\Theta}\left[\frac{1}{2} \int_{x \times [0, 1]} |z|^2 \rho(dydz)dt\right].
\]
Now fix $\theta \in \mathcal{P}(\mathcal{X})$ such that $\hat{I}(\theta) < \infty$. Given $\hat{\Theta} \in \hat{\mathcal{V}}$ such that $\hat{\Theta}_X = \theta$, we construct $\Theta \in \mathcal{V}$ with $\Theta_X = \theta$ such that
\[
\mathbb{E}^\Theta \left[ \frac{1}{2} \int_{\mathbb{R} \times [0,1]} |z|^2 \rho(dzdt) \right] \leq \mathbb{E}^{\hat{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{T} \times \mathbb{R} \times [0,1]} |z|^2 \hat{\rho}(dyd\tilde{z}dt) \right].
\]

The fact that $\Theta \in \mathcal{V}$ already fixes $\Theta_{X,Y}$. Then we set
\[
\Theta_{Z}(C) = \hat{\Theta}_{\hat{\mathcal{X}}}(\{ \hat{r} \in \hat{\mathcal{Y}} : \exists r \in C \text{ where } r_t(dz) = ZZ \ast \delta_{f_{r_X}[\Phi'(y)+1]|\hat{\tilde{r}}_t(dyd\tilde{z})}(dz), \text{ for a.e. } t \in [0,1]\})
\]
for all $C \in \mathcal{Z}$. Notice that in our definition of $Z$ we do not require $r_t$ to be a probability measure (the proof of tightness in Subsection 6.1.3 still holds in the larger space of positive Borel measures on $\mathbb{R}^m \times [0,1]$ with finite first moment rather than just probability measures with finite first moment since $\{Q_N^x\}_{N \in \mathbb{N}}$ is uniformly bounded in the total variation norm, see for example Theorem 8.6.2 in [8]), and that since $\Phi'(y)$ is bounded,
\[
\int_0^1 \int_{\mathbb{R}} ZZ |z| \delta_{f_{r_X}[\Phi'(y)+1]|\tilde{r}_t(dyd\tilde{z})}(dz)dt = ZZ \int_0^1 \left| \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) \right| dt \leq ZZ \int_0^1 \left| \tilde{r}_t(dyd\tilde{z}) dt \right| < \infty \text{ for } \hat{r} \in \hat{\mathcal{Y}}.
\]

So indeed $\hat{r} \in \hat{\mathcal{Y}}$, $r(dzdt) = ZZ \ast \delta_{f_{r_X}[\Phi'(y)+1]|\tilde{r}_t(dyd\tilde{z})}(dz)dt$ is in $Z$. The equality in distribution
\[
\frac{\sigma}{ZZ} \int_0^s \int_{\mathbb{R}} z\rho_t(dz)dt = \sigma \int_0^s \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) dt
\]
holds for all $s \in [0,1]$. In other words, the dynamics of Equation 58 under $\hat{\Theta}$ and Equation 59 under $\Theta$ are the same. The fact that $\Theta \in \mathcal{V}$ then follows from the fact that $\Theta \in \mathcal{V}$. In addition,
\[
\mathbb{E}^\Theta \left[ \frac{1}{2} \int_{\mathbb{R} \times [0,1]} |z|^2 \rho(dzdt) \right] = \mathbb{E}^{\hat{\Theta}} \left[ \frac{Z}{2} \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) \right] \left| \frac{\sigma}{ZZ} \int_0^s \int_{\mathbb{R}} z\rho_t(dz)dt \right|^2 dt
\]
\[
= \mathbb{E}^{\hat{\Theta}} \left[ \frac{Z}{2} \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) \right] \left| \frac{\sigma}{ZZ} \int_0^s \int_{\mathbb{R}} z\rho_t(dz)dt \right|^2 dt \text{ by Equation 57}
\]
\[
\leq \mathbb{E}^{\hat{\Theta}} \left[ \frac{Z}{2} \int_{\mathbb{T} \times \mathbb{R}} \left| \int_{\mathbb{T}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) \right|^2 \left( \int_{\mathbb{T} \times \mathbb{R}} z^2|\gamma(dz|y,t)\pi(dy) \right) dt \right] \text{ by H"{o}lder's inequality}
\]
\[
= \mathbb{E}^{\hat{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{T} \times \mathbb{R}} \int_{\mathbb{T}} z^2|\gamma(dz|y,t)\pi(dy) dt \right] \text{ by Equation 20}
\]
\[
\leq \mathbb{E}^{\hat{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{T} \times \mathbb{R}} \int_{\mathbb{T}} z^2|\gamma(dz|y,t)\pi(dy) dt \right] \text{ by Jensen's inequality}
\]
\[
= \mathbb{E}^{\hat{\Theta}} \left[ \frac{1}{2} \int_{\mathbb{T} \times \mathbb{R} \times [0,1]} |z|^2 \hat{\rho}(dyd\tilde{z}dt) \right].
\]

The rigorous construction of $\Theta_{Z}$ in this direction follows in the same way as for previous one, but this time $p : \hat{\mathcal{Y}} \rightarrow \mathcal{Z}$ is given by $p(\hat{r}) = ZZ \ast \delta_{f_{\hat{r}_X}[\Phi'(y)+1]|\tilde{r}_t(dyd\tilde{z})}(dz)dt$. The issue here is that for two different $\hat{r}^1, \hat{r}^2 \in \hat{\mathcal{Y}}$, we could have $\int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) = \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z})$, so $p$ is not injective. However, this can be overcome by defining the equivalence relation on $\hat{\mathcal{Y}}$ by $\hat{r}^1 \sim \hat{r}^2$ if $\int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) = \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z})$. Then we can take the quotient measure space $\hat{\mathcal{Y}}/\sim$ and corresponding measure $\hat{\Theta}_{\hat{\mathcal{Y}}/\sim}$ and define $\Theta_{Z}$ as above, but with $\hat{\Theta}_{\hat{\mathcal{Y}}/\sim}$ replaced by $\hat{\Theta}_{\hat{\mathcal{Y}}/\sim}$ (see, for example, Section 3 of [55]). Since $\hat{\rho}$ only shows up in the dynamics of Equation 55 as $\int_0^s \int_{\mathbb{T} \times \mathbb{R}} [\Phi'(y) + 1]|\tilde{r}_t(dyd\tilde{z}) dt$, $\Theta$ still satisfies (V1) by virtue of $\hat{\Theta}$ satisfying (V1) and the rest of the analysis still follows. \qed

\textbf{Remark E.1.} While reading Section 3 one might realize that, due to the linear interaction of the prelimit process with the second and third marginals of the occupation measures $\{Q_N^x\}_{N \in \mathbb{N}}$, there is an equally viable
choice of occupation measures which joins the relaxed controls and the prelimit invariant measure to lie on one space. Namely:

\[ \hat{Q}_N^i(A \times B) = \frac{1}{N} \sum_{i=1}^{N} \delta_{\hat{X}_i^{1,N}(\omega)}(A) \delta_{\hat{P}_i^{1,N}(\omega)}(B) \]

for \( A \times B \in B(\hat{W}), \omega \in \Omega, \hat{X}_i^{1,N}(\omega) \in \hat{Y} \) given by

\[ \hat{P}_i^{1,N}(\omega)(I \times C \times D) := \int I(\hat{X}_i^{1,N}(t,\omega)/\delta) \mod 1(C) \delta_{\hat{u}_i^N(t,\omega)}(D) dt \]

for \( I \in B([0,1]), C \in B(\mathbb{T}^d), D \in B(\mathbb{R}^m) \). Here \( \hat{W} \) is as in Theorem 5.2.

Indeed, with such a choice the proofs in Sections 6, 7, and 9 go through in essentially the same way. It is a well known result that the rate function for the Laplace Principal of a sequence of random variables on a given space is unique (see, for example, Theorem 1.3.1 in [21]). The proof of Theorem 5.2 shows that, at least in a certain subclass of our system, the rate function attained from the use of \( \{\hat{Q}^N\}_{N \in \mathbb{N}} \) and \( \{Q^N\}_{N \in \mathbb{N}} \) agree. It is thus possibly a viable option to use \( \{Q^N\}_{N \in \mathbb{N}} \) to attain a rate function of the form offered in Theorem 5.2 for the general system.

The reason why we instead make the choice of occupation measures described by Equation 23 is because it makes the proof of the upper bound in Section 8 much easier. We would not be able to exploit weak-sense uniqueness (see Definition 2.2) as we did if we had used \( \{Q^N\}_{N \in \mathbb{N}} \).

References

[1] A. Ansari, Mean first passage time solution of the Smoluchowski equation: Application of relaxation dynamics in myoglobin, Journal of Chemical Physics 112 (2000), no. 5, 2516–2522.
[2] P. Baldi, Large deviations for diffusions processes with homogenization and applications, Annals of Probability 19 (1991), no. 2, 509–524.
[3] J. Barré, C. Bernardin, R. Chétrite, Y. Chopra, and M. Mariani, Gamma convergence approach for the large deviations of the density in systems of interacting diffusion processes (2019), available at arXiv:1910.04020 [math.AP].
[4] A. Budhirja, J. L. Lions, and G. Papanicolaou, Asymptotic Analysis for Periodic Structures, North Holland, Amsterdam, 1978.
[5] P. Billingsley, Convergence of Probability Measures, Wiley, NY, 1999.
[6] J. Binney and S. Tremaine, Galactic Dynamics, Princeton University Press, Princeton, 2008.
[7] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Funnels, pathways and the energy landscape of protein folding: A synthesis, Proteins 21 (1995), no. 3, 167–195.
[8] V. I. Bogachev, Measure Theory, Vol. 2, Springer, NY, 2007.
[9] V. Borkar and V. Gaitsgory, Averaging of singularly perturbed controlled stochastic differential equations, Applied Mathematics and Optimization 56 (2007), no. 2, 169–209.
[10] A. Budhirja and P. Dupuis, A Variational representation for positive functionals of infinite dimensional brownian motion, Probab. Math. Statist. 20 (2001).
[11] A. Budhirja, P. Dupuis, and M. Fischer, Large deviation properties of weakly interacting particles via weak convergence methods, T.A. of Prob. 40 (2012), 74–100.
[12] R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications I, Springer, NY, 2018.
[13] P. E. Chaudru de Raynal and N. Frikha, From the backward Kolmogorov PDE on the Wasserstein space to propagation of chaos for McKean-Vlasov SDEs (2019), available at arXiv:1907.01410 [math.AP].
[14] __________, Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space (2018), available at arXiv:1811.06904 [math.CA].
[15] D. A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior, J. Stat. Phys. 31 (1983), 29–85.
[16] D. A. Dawson and J. Gärtner, Large deviations from the mckean-vlasov limit for weakly interacting diffusions, Stochastics 20 (1987), no. 4, 247–308.
[17] M. G. Delgadino, R. S. Gvalani, and G. A. Pavliotis, On the diffusive-mean field limit for weakly interacting diffusions exhibiting phase transitions (2020), available at arXiv:2001.03920 [math.AP].
[18] IR. M. Dudley, Real Analysis and Probability, Cambridge University Press, Cambridge, 2010.
[19] A. B. Duncan and G. A. Pavliotis, Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space (2016), available at arXiv:1605.05884 [math.ph].
[20] P. Dupuis and K. Spiliopoulos, Large deviations for multscale diffusion via weak convergence methods, Stochastic Processes and their Applications 122 (2012), no. 4, 1947–1987.
[21] P. Dupuis and R. S. Ellis, A Weak Convergence Approach to the Theory of Large Deviations, Wiley, NY, 1997.
[22] S. Ethier and T. Kurtz, Markov Processes: Characterization and Convergence, Wiley, NY, 1986.
[23] J. Feng, M. Forde, and J. P. Fouque, Short-maturity asymptotics for a fast mean-reverting heston stochastic volatility model, SIAM Journal on Financial Mathematics 1 (2010), no. 1, 126–141.
[64] R. Zwanzig, *Diffusion in a rough potential*, Proc. Natl. Acad. Sci. **85** (1988), 2029–2030.