Asymptotic behavior of stochastic currents under large deviation scaling with mean field interaction and vanishing noise

AMARJIT BUDHIRAJA AND MICHAEL CONROY

Abstract. We study the large deviation behavior of a system of diffusing particles with a mean field interaction, described through a collection of stochastic differential equations, in which each particle is driven by a vanishing independent Brownian noise. An important object in the description of the asymptotic behavior, as the number of particles approaches infinity and the noise intensity approaches zero, is the stochastic current associated with the interacting particle system in the sense of Flandoli *et al.* (2005). We establish a joint large deviation principle (LDP) for the path empirical measure for the particle system and the associated stochastic currents in the simultaneous large particle and small noise limit. Our work extends recent results of Orrieri (2018), in which the diffusion coefficient is taken to be the identity, to a setting of a state dependent and possibly degenerate noise with the mean field interaction influencing both the drift and diffusion coefficients, and allowing for a stronger topology on the space of stochastic currents in the LDP. Proof techniques differ from those used by Orrieri (2018) and rely on methods from stochastic control, theory of weak convergence, and representation formulas for Laplace functionals of Brownian motions.

Mathematics Subject Classification (2020): 60F10 (primary); 60K35, 60B10, 60H05, 60H10, 93E20 (secondary).

1. Introduction

Consider the interacting particle system described through a collection of stochastic differential equations (SDEs) on $\mathbb{R}^d$ given as

$$
\begin{aligned}
dX_j^N(t) = {} & b\left(X_j^N(t), V^N(t)\right) dt \\
& + \varepsilon_N \sigma\left(X_j^N(t), V^N(t)\right) dW_j(t), \quad 1 \le j \le N, \quad N \in \mathbb{N},
\end{aligned}
\tag{1.1}
$$

on some finite time horizon $0 \leq t \leq T$, where $\varepsilon_N \downarrow 0$ as $N \to \infty$ and $\{W_j, j \in \mathbb{N}\}$ are independent m-dimentional Brownian motions on $[0, T]$. Here $V^N(t)$ is the empirical measure of the particle states at time t, namely

$$V^N(t) = \frac{1}{N} \sum_{j=1}^{N} \delta_{X_j^N(t)}, \qquad 0 \leq t \leq T,$$

and thus the interaction among the particles is of the mean field type and influences both the drift and diffusion coefficients of each particle. The law of large numbers (LLN) and fluctuation results for such mean field systems have been widely studied, see for instance [4, 9, 22, 25–27, 31]. In particular, when $N \to \infty$, under conditions on the coefficients and the initial data, $\{V^N(t), 0 \leq t \leq T\}$ converges to the solution of the Vlasov equation

$$\frac{\partial}{\partial t} V + \nabla \cdot b(\cdot, V)V = 0,$$

which can be formally written as

$$\frac{\partial}{\partial t} V + \nabla \cdot \mathcal{J} = 0, \tag{1.2}$$

where $\mathcal{J} \doteq b(\cdot, V)V$ is the nonlinear current given as the limit of the *stochastic currents*

$$J^N(\varphi) = \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, X_j^N(t)\right) \circ dX_j^N(t), \tag{1.3}$$

defined for arbitrary smooth and compactly supported $\varphi : (0, T) \times \mathbb{R}^d \to \mathbb{R}^d$, where $\circ$ denotes the Stratonovich integral. Currents and their stochastic counterparts are key objects in geometric measure theory and play an important role in the theory of rough paths (*cf.* [14, 17, 18, 23]). In the current context they provide a convenient way to describe the asymptotics of the empirical measure process V^N.

In this work we are interested in studying the asymptotics of probabilities of significant deviations of the empirical measure V^N, for the N-particle microscopic stochastic evolution described by (1.1), from its macroscopic hydrodynamic limit described by the first order Vlasov equation in (1.2). A common approach to such a study is to establish a general large deviation principle (LDP) on an appropriate abstract space from which the information on probabilities of deviations for specific events involving the N-particle system (1.1) can be obtained by a suitable application of the contraction principle. In view of the representation of the hydrodynamic limit of V^N in terms of the nonlinear current functional $\mathcal{J}$, a natural candidate for an LDP are the pairs (V^N, J^N) regarded as random elements of an appropriate space. Under the conditions on the coefficients considered in this work (see Condition 2.1), V^N will take values in $\mathcal{V} \doteq \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$, namely, the space of continuous functions from $[0, T]$ to the space $\mathcal{P}_1(\mathbb{R}^d)$ of probability measures on

$\mathbb{R}^d$ with finite first moment, equipped with the Wasserstein-1 distance (see Section 2 for precise definitions). The identification of an appropriate space for J^N requires a bit more work (*cf.* [13, 29]). In particular, note that (1.3) describes an uncountably infinite collection of identities in which the right-hand side is defined in an almost sure sense for each fixed φ. Thus a basic problem is to provide a pathwise representation for the collection

$$\left\{ \varphi \mapsto \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, X_j^N(t)\right) \circ dX_j^N(t) \right\}, \qquad (1.4)$$

which defines a continuous, linear map on a suitable function space. This problem was studied in [13] (see also [29]) where it was shown that there is a random variable $\mathcal{J}^N$ with values in a certain negative Sobolev space $\mathbf{H}^{-\mathbf{s}}$ of distributions (see Subsection 2.2), which gives a pathwise representation for the collection in (1.4) in the sense that

$$\langle \mathcal{J}^N, \varphi \rangle = \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, X_j^N(t)\right) \circ dX_j^N(t) \qquad \text{a.s.,}$$

for every smooth φ with compact support. Thus the stochastic currents $\mathcal{J}^N$ can be viewed as random elements of the Hilbert space $\mathbf{H}^{-\mathbf{s}}$, and the basic problem of interest is then to establish a large deviation principle for $(V^N, \mathcal{J}^N)$ in $\mathcal{V} \times \mathbf{H}^{-\mathbf{s}}$.

This large deviation problem in the setting where $m = d$ and $\sigma = \mathrm{Id}$ was studied in [29] by direct change of measure arguments. Specifically, [29] treats the large deviation upper bound by first establishing an estimate for compact sets by considering an explicit tilt of the measure and then extends the estimate to all closed sets by establishing certain exponential tightness estimates. The lower bound is proved by exploiting connections between large deviations and Γ-convergence from [24], in particular the key idea is to construct a suitable 'recovery sequence' using results from [16]. One important aspect of the results and proof methods in [29] is that the LDP is established with the weak topology on the Hilbert space $\mathbf{H}^{-\mathbf{s}}$. Indeed, both the proofs of the upper and lower bounds rely on the use of the weak topology in important ways; *e.g.*, since bounded sets are relatively compact under the weak topology in $\mathbf{H}^{-\mathbf{s}}$, in proving exponential tightness it suffices to estimate the probability that $\mathcal{J}^N$ takes values in the complement of a bounded ball.

In the current work we take a different approach to the study of the large deviation principle that is based on methods from stochastic control, the theory of weak convergence of probability measures, and Laplace asymptotics. This approach allows us to avoid establishing exponential tightness estimates of the form in [29] and enables us to treat diffusion coefficients that are state dependent and possibly degenerate (see Subsection 2.1). In addition, since in this approach one needs to establish ordinary tightness rather than exponential tightness, by appealing to certain compact embedding results for Sobolev spaces, we are able to establish an LDP with the norm topology on $\mathbf{H}^{-\mathbf{s}}$ instead of the weak topology considered in [29].

In fact, we establish a somewhat more general large deviation principle than the one considered in [29] from which the LDP for $(V^N, \mathcal{J}^N)$ can be deduced by the contraction principle. Specifically, we consider path empirical measures μ^N associated with the interacting particle system in (1.1) defined as

$$\mu^N = \frac{1}{N} \sum_{j=1}^{N} \delta_{X_j^N}.$$

Under the conditions of this work it follows that μ^N is a random variable with values in $\mathcal{P}_1(\mathcal{C}([0, T], \mathbb{R}^d))$, namely the space of probability measures, on the Banach space of $\mathbb{R}^d$-valued continuous trajectories on $[0, T]$, with integrable norm (equipped with the Wasserstein-1 metric). Our main result, Theorem 2.10, gives an LDP for $(\mu^N, \mathcal{J}^N)$ in $\mathcal{P}_1(\mathcal{C}([0, T], \mathbb{R}^d)) \times \mathbf{H}^{-s}$. Using the continuity of the map $\nu \mapsto \{t \mapsto \nu \circ \pi_t^{-1}\}$ from $\mathcal{P}_1(\mathcal{C}([0, T], \mathbb{R}^d))$ into $\mathcal{V}$, where π_t is the projection map on $\mathcal{C}([0, T], \mathbb{R}^d)$ giving the evaluation at time t, we then deduce an LDP for the sequence $(V^N, \mathcal{J}^N)$ in $\mathcal{V} \times \mathbf{H}^{-s}$ in Corollary 2.11. The rate function, in the general setting of a state dependent diffusion coefficient, is given as a value function of a certain deterministic mean field control problem with a quadratic cost (see (2.11) and (2.16)). In Proposition 2.12 we show that in the special case where $\sigma = \mathrm{Id}$, this representation of the rate function simplifies to a more explicit form given in terms of certain controlled Vlasov equations (see (2.17)) which was obtained in [29].

As noted previously, proof techniques here are quite different from [29]. The starting point of our analysis is a certain variational representation for exponential functionals of finite dimensional Brownian motions (see [2, 6]), using which the proof of the large deviation principle reduces to a study of tightness and convergence properties of certain controls and controlled analogues of the state processes $\{X_j^N, 1 \leq j \leq N\}$, state empirical measures V^N, path occupation measures μ^N, and stochastic currents $\mathcal{J}^N$, denoted as $\{\bar{X}_j^N, 1 \leq j \leq N\}$, $\bar{V}^N$, $\bar{\mu}^N$, and $\bar{\mathcal{J}}^N$, respectively. For the upper bound proof we introduce certain joint empirical measures, denoted as Q^N (see (3.7)), of particle trajectories and associated control processes. The main step in the proof of the upper bound is to establish the tightness of the sequence $\{(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N), N \in \mathbb{N}\}$ and to provide a suitable characterization of the weak limit points of this sequence. In particular, the tightness of the controlled stochastic currents $\{\bar{\mathcal{J}}^N\}$ is established with the norm topology on $\mathbf{H}^{-s}$ and relies on approximations of $\{\bar{\mathcal{J}}^N\}$ by distributions with compact support as well as certain compact embedding results for Sobolev spaces (see Lemma 4.5). The lower bound proof is constructive in that, given a near optimal measure μ on $\mathcal{C}([0, T], \mathbb{R}^d)$ and a near optimal current $\mathcal{J}$ in a certain variational problem associated with the rate function, we construct a sequence of controls and controlled variables $(\bar{\mu}^N, \bar{\mathcal{J}}^N)$ that converge to $(\mu, \mathcal{J})$ in a suitable manner. The key ingredients in the proof here are a weak uniqueness (*i.e.*, uniqueness in probability laws) property of certain equations associated with the controlled versions of the Vlasov equation (1.2) (see Lemma 3.5) and certain infinite product space constructions.

Large deviation principles for weakly interacting diffusions as in (1.1) with non-vanishing noise (*i.e.*, $\varepsilon_N = 1$) have been studied in [10]. A different approach, based on weak convergence methods of the form used in the current work, was taken in [8]. The latter paper, in contrast to [10], allowed for degenerate diffusion coefficients and for a mean field interaction in the diffusion coefficient. There have also been several works (in addition to the paper [29] discussed above) that have studied large deviation problems for weakly interacting diffusions with small noise. In particular, see [19, 30], and references therein, for large deviations results for McKean-Vlasov equations in the small noise limit; and see [20] for an analysis of interchanging of mean field limit with the small noise limit at the level of rate function convergence. In a related direction, the paper [5] studied large deviation properties of a system of interacting diffusions in which each particle is driven by an independent individual source of noise and also by a vanishing amount of noise that is common to all particles. Different levels of intensity of the small common noise lead to different types of large deviation behavior, and the paper [5] provided precise characterization of the various regimes.

1.1. Organization

The paper is organized as follows. In Section 2, we specify our model, describe the space on which the large deviation principle will hold, define the rate function, and present our main large deviation result. Section 3 provides the proof of this result, with the proofs of its key lemmas given in Section 4. The proofs of some auxiliary results are given in the Appendix.

1.2. Notation

The following notation will be used throughout. We use $\mathcal{C}(R, S)$, $\mathcal{C}_c(R, S)$, and $\mathcal{C}^k(R, S)$, $k \in \mathbb{N} \cup \{\infty\}$, to denote the spaces of continuous, continuous and compactly supported, and k-times continuously differentiable functions from R into S, respectively. Also, $\mathcal{C}_c^k(R, S) = \mathcal{C}_c(R, S) \cap \mathcal{C}^k(R, S)$ for $k \in \mathbb{N} \cup \{\infty\}$. We denote by $L^2(\mu, R, S)$ the space of μ-square integrable functions from R into S. When μ is the Lebesgue measure, we will occasionally suppress it in the notation and write $L^2(\mu, R, S)$ as $L^2(R, S)$. The evaluation of a distribution F on a test function φ will be denoted by $\langle F, \varphi \rangle$, and integration of a function f with respect to a measure μ will be denoted by $\langle \mu, f \rangle$. $\mathcal{B}(S)$ denotes the collection of all Borel sets on S. For a Polish space (S, d_S), $\mathcal{P}(S)$ denotes the space of probability measures on S, endowed with the topology of weak convergence. A convenient metric on this space is the bounded Lipschitz metric given as

$$d_{\mathrm{bl}}(\mu, \nu) \doteq \sup_{f \in \mathcal{L}_{\mathrm{b}}(S)} \left| \langle \mu, f \rangle - \langle \nu, f \rangle \right|, \qquad \mu, \nu \in \mathcal{P}(S), \qquad \text{where}$$

$$\mathcal{L}_{\mathrm{b}}(S) \doteq \left\{ f \in \mathcal{C}(S, \mathbb{R}) : \sup_{x \neq y} \frac{|f(x) - f(y)|}{d_S(x, y)} \leq 1, \ \sup_x |f(x)| \leq 1 \right\}.$$

When $\theta \in \mathcal{P}(S)$, the notation E_θ will be used to denote expectation on the probability space $(S, \mathcal{B}(S), \theta)$. For two spaces S_1 and S_2 and $\theta \in \mathcal{P}(S_1 \times S_2)$, $\theta_{(1)}$ and $\theta_{(2)}$ will denote the marginal distributions on S_1 and S_2, respectively. Similar notation will be used when more than two spaces are involved. Euclidean norms will be denoted by $|\cdot|$. For a Polish space (S, d_S), the space $\mathcal{C}([0, T], S)$ will be equipped with the metric

$$d(x, y) = \sup_{0 \le t \le T} d_S(x(t), y(t)),$$

under which it is a Polish space as well. On $\mathcal{C}([0, T], \mathbb{R}^d)$, we define the norm $\|x\|_\infty \doteq \sup_{0 \le t \le T} |x(t)|$, and the metric above becomes $d(x, y) = \|x - y\|_\infty$. We will use $\Rightarrow$ to denote convergence in distribution, and $\xrightarrow{P}$ to denote convergence in P-probability. Infimum over an empty set, by convention, is taken to be $+\infty$. For a metric space S, a function $I : S \to [0, \infty]$ is called a rate function if $\{x \in S : I(x) \le l\}$ is a compact set for every $l < \infty$.

2. Preliminaries and main result

Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}(t), 0 \le t \le T\})$ be a filtered probability space where the filtration satisfies the *usual conditions* (see [21, Definition 21.22]). Fix $m \in \mathbb{N}$, and let $\{W_j, j \in \mathbb{N}\}$ be a sequence of independent m-dimensional $\{\mathcal{F}(t)\}$-Brownian motions on the time horizon $0 \le t \le T$. For each $N \in \mathbb{N}$, we consider the following system of stochastic differential equations in $\mathbb{R}^d$:

$$
\begin{aligned}
X_j^N(t) = X_j^N(0) &+ \int_0^t b\left(X_j^N(s), V^N(s)\right) ds \\
&+ \varepsilon_N \int_0^t \sigma\left(X_j^N(s), V^N(s)\right) dW_j(s), \qquad 1 \le j \le N,
\end{aligned}
\tag{2.1}
$$

where $V^N(t)$ denotes the $\mathcal{P}(\mathbb{R}^d)$-valued empirical measure

$$
V^N(t) \doteq \frac{1}{N} \sum_{j=1}^N \delta_{X_j^N(t)}, \qquad 0 \le t \le T,
\tag{2.2}
$$

and $\{\varepsilon_N, N \in \mathbb{N}\}$ is some sequence in $\mathbb{R}_+$ such that $\varepsilon_N \downarrow 0$ as $N \to \infty$. Without loss of generality, we will assume that $\sup_N \varepsilon_N \le 1$ throughout. Denote $\mathcal{X} \doteq \mathcal{C}([0, T], \mathbb{R}^d)$, and define $\mathcal{P}(\mathcal{X})$-valued random variables, given as the empirical measure of $(X_1^N, \dots, X_N^N)$, by

$$
\mu^N \doteq \frac{1}{N} \sum_{j=1}^N \delta_{X_j^N}.
\tag{2.3}
$$

Note that the marginal of μ^N at time t is $V^N(t)$, that is, defining $\pi_t : \mathcal{C}([0,T],\mathbb{R}^d) \rightarrow \mathbb{R}^d$ as the projection map $\pi_t(x) = x(t)$, we have

$$\mu^N \circ \pi_t^{-1} = V^N(t), \qquad 0 \le t \le T.$$

We will view each μ^N as a random variable taking values in the Wasserstein-1 space which is defined as follows. For a Polish space (S, d_S), define the space $\mathcal{P}_1(S)$ by

$$\mathcal{P}_1(S) \doteq \left\{ \mu \in \mathcal{P}(S) : \int_S d_S(x, x_0) \mu(dx) < \infty \right\},$$

for some choice of $x_0 \in S$ (the space does not depend on the choice of x_0). Then $\mathcal{P}_1(S)$ is a Polish space under the Wassertstein-1 distance given by

$$
\begin{aligned}
d_1(\mu, \nu) &\doteq \sup_{f \in \mathcal{L}(S)} \left| \langle \mu, f \rangle - \langle \nu, f \rangle \right|, \\
\mathcal{L}(S) &\doteq \left\{ f \in \mathcal{C}(S, \mathbb{R}) : \sup_{x \neq y} \frac{|f(x) - f(y)|}{d_S(x, y)} \le 1 \right\}.
\end{aligned}
\tag{2.4}
$$

For further details on Wassertstein spaces, we refer to [32]. The particular cases of interest here are the spaces $\mathcal{P}_1(\mathbb{R}^d)$ and $\mathcal{P}_1(\mathcal{X})$, and the notation d_1 will be used for the metric on both spaces, with the distinction being clear from the context. Noting that (under Condition 2.1 given below)

$$\int_{\mathcal{X}} d_{\mathcal{X}}(x, 0) \mu^N(dx) = \int_{\mathcal{X}} \|x\|_\infty \mu^N(dx) = \frac{1}{N} \sum_{j=1}^N \left\| X_j^N \right\|_\infty < \infty \qquad \text{a.s.,}$$

we see that indeed μ^N is a $\mathcal{P}_1(\mathcal{X})$-valued random variable. Similarly, it can be checked that V^N is a $\mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$-valued random variable. Throughout, we will denote $\mathcal{V} \doteq \mathcal{C}([0, T], \mathcal{P}_1(\mathbb{R}^d))$.

2.1. Main conditions

The following is our main assumption on the coefficients.

Condition 2.1. There is some $L < \infty$ such that for all $x, y \in \mathbb{R}^d$ and $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$,

$$|b(x, \mu) - b(y, \nu)| + |\sigma(x, \mu) - \sigma(y, \nu)| \le L \left(|x - y| + d_1(\mu, \nu) \right)$$

and $|\sigma(x, \mu)| \le L$.

Note that the above condition implies in particular that for all $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$,

$$|b(x, \mu)| \le L \left(1 + |x| + \int_{\mathbb{R}^d} |y| \mu(dy) \right), \tag{2.5}$$

with possibly a larger choice of L than in Condition 2.1. By standard arguments, Condition 2.1 implies that there exists a unique pathwise solution to (2.1) for each $N \in \mathbb{N}$.

Remark 2.2. The boundedness of σ is used in an important way at several places in the proof. It is a key ingredient in the proof of Lemma 4.1 which in turn is key to Lemmas 4.6 and 3.3. The last two lemmas are used in both the upper and lower bound proofs. For the upper bound proof one can relax the assumption on the boundedness of σ by using localization arguments of the form used in [6] (see, e.g., [7, Theorem 8.4]), however these localization arguments do not work in a simple manner for the proof of the lower bound. Relaxing the condition on the boundedness of σ remains an interesting open problem.

We assume the following on the initial conditions of (2.1).

Condition 2.3. For each $N \in \mathbb{N}$ and $1 \leq j \leq N$, $X_j^N(0) = x_j^N \in \mathbb{R}^d$ is deterministic. The collection of initial conditions satisfies the following:

(i) There exists some $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ such that $d_{\mathrm{bl}}\left(V^N(0), \mu_0\right) \to 0$;

(ii) $\displaystyle \sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^{N} \left| x_j^N \right|^2 < \infty.$

Note that (i) and (ii) above imply that $\int_{\mathbb{R}^d} |x|^2 \, \mu_0(dx) < \infty$ from the observation

$$\int_{\mathbb{R}^d} \left(|x|^2 \wedge K \right) \mu_0(dx) = \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} \left(\left| x_j^N \right|^2 \wedge K \right) \leq \sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^{N} \left| x_j^N \right|^2$$

for any $K \in (0, \infty)$, and applying Fatou's lemma. The above condition also gives that, as $N \to \infty$,

$$d_1\left(V^N(0), \mu_0\right) \to 0.$$

In order to prove the Laplace lower bound, we will make a stronger assumption given below on the diffusion coefficient σ which says that it depends on the state of the system only through the empirical measure. We will also require the convergence of the initial data in a somewhat stronger sense.

Condition 2.4.

(i) For each $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, $\sigma(x, \mu) = \sigma(\mu)$;
(ii) For all μ_0-integrable $f : \mathbb{R}^d \to \mathbb{R}$,

$$\left\langle V^N(0), f \right\rangle \to \langle \mu_0, f \rangle \qquad \text{as } N \to \infty.$$

Remark 2.5. Part (i) of Condition 2.4 is used in the proof of the weak uniqueness result in Lemma 3.5. Relaxing this condition is a challenging open problem. The second part of Condition 2.4 is used in obtaining the convergence stated in (3.11).

We are interested in the large deviations behavior of μ^N and V^N as well as a collection of random linear functionals, referred to as *stochastic currents*, associated with the sequence of processes $\{X_j^N(t)\}$. We now introduce these objects. For each N and $\varphi \in \mathcal{C}_c^\infty([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ define

$$J^N(\varphi) \doteq \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi\left(t, X_j^N(t)\right) \circ dX_j^N(t), \qquad (2.6)$$

where the above is a Stratanovich stochastic integral. The relationship between Stratanovich and Itô integrals gives the following formula for $J^N(\varphi)$:

$$J^N(\varphi) = \frac{1}{N} \sum_{j=1}^N \left(\int_0^T \varphi\left(t, X_j^N(t)\right) \cdot dX_j^N(t) + \frac{1}{2} \left\langle \varphi\left(\cdot, X_j^N(\cdot)\right), X_j^N(\cdot)\right\rangle_T \right),$$

where $\langle Y, Z \rangle_t$ denotes the quadratic variation at time t of two continuous semi-martingales Y and Z. From results in [13], J^N can be viewed as a random linear functional on a suitable Sobolev space. We now briefly describe these results and make precise the space in which these random linear functionals take values.

2.2. Stochastic currents

Recall that for $k \in \mathbb{N}$, $H^k(\mathbb{R}^d, \mathbb{R}^d)$ is the Hilbert space of functions $f \in L^2(\mathbb{R}^d, \mathbb{R}^d)$ such that the distributional derivatives $D^\alpha f$ are also L^2 functions for all $|\alpha| \leq k$, where $\alpha = (\alpha_1, \ldots, \alpha_d)$ denotes a multi-index. More generally, for any $s \in \mathbb{R}_+$, $H^s(\mathbb{R}^d, \mathbb{R}^d)$ is defined as the space of functions $f \in L^2(\mathbb{R}^d, \mathbb{R}^d)$ such that

$$\|f\|_s^2 \doteq \int_{\mathbb{R}^d} \left|\hat{f}(\xi)\right|^2 (1 + |\xi|^2)^s \, d\xi < \infty, \qquad (2.7)$$

where $\hat{f}(\xi) = \int e^{-2\pi i \xi \cdot x} f(x) \, dx$ is the Fourier transform on $\mathbb{R}^d$. We refer the reader to [1, 15, 28] for details on these spaces.

In order to describe the linear space associated with the map $\varphi \mapsto J^N(\varphi)$, we will need to consider a suitable Sobolev space of functions of time and space. Following [3, 13, 29], a natural choice in this regard is the space

$$H^{s_1}\left((0, T), H^{s_2}\left(\mathbb{R}^d, \mathbb{R}^d\right)\right),$$

where $\mathbf{s} = (s_1, s_2) \in (\frac{1}{2}, 1) \times (\frac{d}{2} + 1, \infty)$ (see [29] for a precise description of the space). However in order to apply certain compact embedding results (see, *e.g.*, the

proof of Lemma 3.3) we will consider a slight modification of these spaces defined as follows.

Fix $a, b \in \mathbb{R}$ such that $a < 0 < T < b$ and define $U \doteq (a, b)$ and $\mathcal{O}_d \doteq (\frac{1}{2}, 1) \times (\frac{d}{2} + 1, \infty)$. Then define

$$\mathbf{H^s} \doteq H^{s_1}\left(U, H^{s_2}\left(\mathbb{R}^d, \mathbb{R}^d\right)\right), \qquad \mathbf{s} \in \mathcal{O}_d,$$

as the space of functions $f : U \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying

$$\|f\|_{\mathbf{s}}^2 \doteq \|f\|_{L^2(U, H^{s_2}(\mathbb{R}^d, \mathbb{R}^d))}^2 + [f]_{\mathbf{s}}^2$$
$$\doteq \int_U \|f(u, \cdot)\|_{s_2}^2 \, du + \int_U \int_U \frac{\|f(u, \cdot) - f(v, \cdot)\|_{s_2}^2}{|u - v|^{1 + 2s_1}} \, du \, dv < \infty, \tag{2.8}$$

where $\|\cdot\|_{s_2}$ is as in (2.7). The norm $\|\cdot\|_{\mathbf{s}}$ is usually referred to as a *Gagliardo norm*, and in fact corresponds to an inner product which makes $\mathbf{H^s}$ a separable Hilbert space (see [28, Section 3]). The topological dual of the Hilbert space $\mathbf{H^s}$ will be denoted as $\mathbf{H^{-s}}$, namely

$$\mathbf{H^{-s}} \doteq (\mathbf{H^s})'.$$

The norm on this space is given as

$$\|F\|_{-\mathbf{s}} \doteq \sup_{\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)} \frac{|\langle F, \varphi \rangle|}{\|\varphi\|_{\mathbf{s}}}.$$

For $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, abusing notation, we let

$$J^N(\varphi) \doteq \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi\big(t, X_j^N(t)\big) \circ dX_j^N(t).$$

Note that if φ_{res} denotes the restriction of φ to $[0, T] \times \mathbb{R}^d$, then $J^N(\varphi) = J^N(\varphi_{\text{res}})$. Also, any $\varphi \in \mathcal{C}_c^\infty([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ can be extended to a $\varphi_{\text{ext}} \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ where once more $J^N(\varphi) = J^N(\varphi_{\text{ext}})$. By a *pathwise realization* of the collection $\{\varphi \mapsto J^N(\varphi)\}$ on $\mathcal{C}_c^\infty([0, T] \times \mathbb{R}^d, \mathbb{R}^d)\}$, we mean a random variable $\mathcal{J}^N$ with values in $\mathbf{H^{-s}}$ such that for any $\varphi \in \mathcal{C}_c^\infty([0, T] \times \mathbb{R}^d, \mathbb{R}^d)$ and any extension φ_{ext} of φ in $\mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, $\langle \mathcal{J}^N, \varphi_{\text{ext}} \rangle = J^N(\varphi)$ a.s.

The following result, giving the existence of a pathwise realization, goes along the lines of [29] . The proof is an immediate consequence of Lemma 3.1 below (on taking $u_j^N = 0$ in the lemma), the proof of which is given in the Appendix.

Theorem 2.6. *Suppose Conditions 2.1 and 2.3 hold. Then, for each $N \in \mathbb{N}$ and $\mathbf{s} \in \mathcal{O}_d$, there is an $\mathbf{H^{-s}}$-valued random variable $\mathcal{J}^N$ on $(\Omega, \mathcal{F}, P)$ such that, for every $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, $\langle \mathcal{J}^N(\omega), \varphi \rangle = [J^N(\varphi)](\omega)$ for a.e. $\omega \in \Omega$. Namely, $\mathcal{J}^N$ is a pathwise realization of $\{\varphi \mapsto J^N(\varphi)\}$.*

Note that the pathwise realizations $\{\mathcal{J}^N\}$ are a.s. compactly supported in the first coordinate. Namely, if $U_0 \subset U$ is an open set such that $U_0 \cap [0,T] = \emptyset$, then for all φ with compact support in $U_0 \times \mathbb{R}^d$, $\langle \mathcal{J}^N, \varphi \rangle = 0$ a.s. In particular, $\mathcal{J}^N$ is a distribution a.s. supported in $[0,T] \times \mathbb{R}^d$.

In this work we will prove a large deviation principle for the pair $(\mu^N, \mathcal{J}^N)$ in the space $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$ for each $\mathbf{s} \in \mathcal{O}_d$, from which an LDP describing the asymptotics of V^N will follow by the contraction principle. We begin by introducing the rate function that will govern the large deviation behavior.

2.3. Rate function

Let $\mathcal{R}$ denote the set of positive measures r on $\mathcal{B}([0,T] \times \mathbb{R}^m)$ such that $r([0,t] \times \mathbb{R}^m) = t$ for all $0 \le t \le T$, and define

$$\mathcal{R}_1 \doteq \left\{ r \in \mathcal{R} : \int_{[0,T] \times \mathbb{R}^m} |y| \, r(dt, dy) < \infty \right\}.$$

The space $\mathcal{R}_1$ is a Polish space under the Wasserstein-1 metric (defined as in (2.4) with $S = [0,T] \times \mathbb{R}^m$). Each $r \in \mathcal{R}_1$ can be decomposed as $r(dt, dy) = r_t(dy) \, dt$, where $r_t \in \mathcal{P}(\mathbb{R}^m)$. For an $\mathcal{R}_1$-valued random variable ρ, consider the McKean-Vlasov equation

$$\begin{aligned}
dX(t) &= b(X(t), V(t)) \, dt + \int_{\mathbb{R}^m} \sigma(X(t), V(t)) y \, \rho_t(dy) \, dt, \\
V(t) &= P \circ X(t)^{-1}, \qquad V(0) = \mu_0,
\end{aligned} \tag{2.9}$$

where X is a stochastic process with sample paths in $\mathcal{X}$, $\rho(dt, dy) = \rho_t(dy) \, dt$ is the disintegration of ρ, and μ_0 is the measure in Condition 2.3(i). The distribution of a pair (X, ρ) that solves (2.9), which is a probability measure on $\mathcal{Z} \doteq \mathcal{X} \times \mathcal{R}_1$, is called a weak solution of (2.9). Let $\mathcal{S}(\mathcal{Z}) \subset \mathcal{P}(\mathcal{Z})$ denote the set of all such weak solutions. With an abuse of notation, we will denote the canonical coordinate maps on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ by (X, ρ) once more. That is,

$$X(\xi, r) = \xi, \qquad \rho(\xi, r) = r, \qquad (\xi, r) \in \mathcal{Z}.$$

Note that, if $\Theta \in \mathcal{S}(\mathcal{Z})$, then (X, ρ) satisfies (2.9) Θ-a.s. For each $\Theta \in \mathcal{P}(\mathcal{Z})$ and $0 \le t \le T$, define the measure

$$\nu_\Theta(t) \doteq \Theta \circ X(t)^{-1},$$

which is an element of $\mathcal{P}(\mathbb{R}^d)$. When $\Theta \in \mathcal{S}(\mathcal{Z})$, it is easy to check that Condition 2.1 and Gronwall's lemma imply that $E_\Theta [|X(t)|] < \infty$, and hence $\nu_\Theta(t) \in \mathcal{P}_1(\mathbb{R}^d)$ for each $0 \le t \le T$. Letting ν_Θ denote the map $t \mapsto \nu_\Theta(t)$, in fact we have that $\nu_\Theta \in \mathcal{V}$. For each $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, define the map $G_\varphi : \mathcal{S}(\mathcal{Z}) \to \mathbb{R}$

by

$$G_\varphi(\Theta) \doteq E_\Theta \left[\int_0^T \varphi\,(t, X(t)) \cdot dX(t) \right]$$

$$= E_\Theta \left[\int_0^T \varphi\,(t, X(t)) \cdot b(X(t), \nu_\Theta(t))dt \right] \qquad (2.10)$$

$$+ E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} \varphi\,(t, X(t)) \cdot \sigma(X(t), \nu_\Theta(t)) y\, \rho(dt, dy) \right].$$

Now let

$$\mathcal{P}_2(\mathcal{Z}) \doteq \left\{ \Theta \in \mathcal{P}(\mathcal{Z}) : E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] < \infty \right\},$$

and for $\mathcal{J} \in \mathbf{H}^{-\mathbf{s}}$, define

$$\mathcal{P}^*(\mathcal{J}) \doteq \left\{ \Theta \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z}) : \langle \mathcal{J}, \varphi \rangle = G_\varphi(\Theta) \text{ for all } \varphi \in \mathcal{C}_c^\infty \left(U \times \mathbb{R}^d, \mathbb{R}^d \right) \right\}.$$

Define $I : \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}} \to [0, \infty]$ as

$$I(\mu, \mathcal{J}) \doteq \inf \left\{ E_\Theta \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] : \Theta_{(1)} = \mu, \Theta \in \mathcal{P}^*(\mathcal{J}) \right\}, \quad (2.11)$$

where we recall that $\Theta_{(1)}$ denotes the marginal of Θ on $\mathcal{X}$.

Remark 2.7. Note that the domain of the function I depends on $\mathbf{s} \in \mathcal{O}_d$. However, it turns out (see Lemma 4.8) that if $I(\mu, \mathcal{J}) < \infty$ for some $\mathbf{s} \in \mathcal{O}_d$ and $(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$, then $\mathcal{J} \in \mathbf{H}^{-\mathbf{s}'}$ for all $\mathbf{s}' \in \mathcal{O}_d$, and the value of $I(\mu, \mathcal{J})$ is independent of $\mathbf{s}$.

2.4. Main results

In this subsection we present the main results. For each $N \in \mathbb{N}$, let μ^N, V^N and $\mathcal{J}^N$ be as in (2.3), (2.2), and Theorem 2.6 respectively. Our first main result is a law of large numbers for $(\mu^N, V^N, \mathcal{J}^N)$.

By using the Lipschitz property of b it can be checked that for μ_0 as in Condition 2.3 and any $\mathbb{R}^d$ valued random variable ξ_0 on $(\Omega, \mathcal{F}, P)$ with distribution μ_0, there is an a.s. unique solution ξ, with sample paths in $\mathcal{X}$, to the equation

$$\xi(t) = \xi_0 + \int_0^t b\left(\xi(s), V^*(s)\right) ds, \qquad V^*(t) = P \circ \xi(t)^{-1}, \ 0 \le t \le T. \quad (2.12)$$

Let

$$\mu^* = P \circ \xi^{-1}. \qquad (2.13)$$

Using the linear growth of b and Condition 2.3(ii) it can be checked that $\mu^* \in \mathcal{P}_1(\mathcal{X})$.

The following theorem gives the law of large numbers. Its proof is given in Subsection 3.6.

Theorem 2.8 (LLN). *Assume Conditions* 2.1 *and* 2.3 *hold and let* $\mathbf{s} \in \mathcal{O}_d$. *Then,*

$$\left(\mu^N, V^N, \mathcal{J}^N\right) \xrightarrow{P} \left(\mu^*, V^*, \mathcal{J}^*\right)$$

in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{V} \times \mathbf{H}^{-\mathbf{s}}$ *as* $N \to \infty$,, *where* V^* *and* μ^* *are as in* (2.12) *and* (2.13) *and* $\mathcal{J}^*$ *is characterized as*

$$\langle \mathcal{J}^*, \varphi \rangle = \int_0^T \left\langle V^*(t), \varphi(t, \cdot) \cdot b\left(\cdot, V^*(t)\right) \right\rangle dt \tag{2.14}$$

for $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$.

Remark 2.9. The pair $(V^*, \mathcal{J}^*)$ can alternatively be characterized as the unique solution of the equation

$$\frac{\partial}{\partial t} V + \nabla \cdot b(\cdot, V)V = 0, \qquad \mathcal{J} = b(\cdot, V)V, \qquad V(0) = \mu_0, \tag{2.15}$$

in the distributional sense on $(0, T) \times \mathbb{R}^d$, by which we mean that for all $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$

$$\int_0^T \left\langle V(t), \frac{\partial}{\partial t} \varphi(t, \cdot) \right\rangle dt + \int_0^T \langle V(t), \nabla \varphi(t, \cdot) \cdot b(\cdot, V(t)) \rangle \, dt = 0,$$

and for all $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d \mathbb{R}^d)$

$$\langle \mathcal{J}, \varphi \rangle = \int_0^T \left\langle V(t), \varphi(t, \cdot) \cdot b(\cdot, V(t)) \right\rangle dt.$$

Recall the function I defined in (2.11), and for each $N \in \mathbb{N}$ let $a_N \doteq N/\varepsilon_N^2$. Our main large deviation result is as follows.

Theorem 2.10 (LDP). *Assume Conditions* 2.1 *and* 2.3 *hold. For each* $\mathbf{s} \in \mathcal{O}_d$, I *is a rate function on* $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$. *Furthermore, the following holds:*

(i) *The sequence* $\{(\mu^N, \mathcal{J}^N), N \in \mathbb{N}\}$ *satisfies the large deviation upper bound on* $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$ *with speed* a_N *and rate function* I. *Namely, for all closed sets* F *in* $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$,

$$\limsup_{N \to \infty} \frac{1}{a_N} \log P\left((\mu^N, \mathcal{J}^N) \in F\right) \leq - \inf_{(\mu, \mathcal{J}) \in F} I(\mu, \mathcal{J});$$

(ii) *If in addition Condition 2.4 holds, then $\{(\mu^N, \mathcal{J}^N), N \in \mathbb{N}\}$ satisfies the large deviation lower bound on $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$ with speed a_N and rate function I. Namely, for all open sets G in $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$,*

$$\liminf_{N \to \infty} \frac{1}{a_N} \log P\left((\mu^N, \mathcal{J}^N) \in G\right) \geq - \inf_{(\mu, \mathcal{J}) \in G} I(\mu, \mathcal{J}).$$

The proof of Theorem 2.10(i) is in Subsection 3.3, and the proof of Theorem 2.10(ii) is in Subsection 3.4. The rate function property of I is proved in Subsection 3.5. The proof of Theorem 2.8 is saved for Subsection 3.6, since it goes along the lines of the proof of the large deviation upper bound.

It is easy to verify that the map $\nu \mapsto \{t \mapsto \nu \circ \pi_t^{-1}\}$ is a continuous map from $\mathcal{P}_1(\mathcal{X})$ into $\mathcal{V}$, and recall from above that each $\Theta \in \mathcal{S}(\mathcal{Z})$ induces $\nu_\Theta \in \mathcal{V}$. From this and the contraction principle we immediately have a large deviation principle for $\{(\mu^N, V^N, \mathcal{J}^N)\}$. In particular, we have the following corollary. Define $\tilde{I}$: $\mathcal{V} \times \mathbf{H}^{-\mathbf{s}} \to [0, \infty]$ as

$$\tilde{I}(V, \mathcal{J}) \doteq \inf \left\{ E_\Theta \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] : \nu_\Theta = V, \Theta \in \mathcal{P}^*(\mathcal{J}) \right\}. \quad (2.16)$$

Corollary 2.11. *Assume Conditions 2.1 and 2.3 hold. For each $\mathbf{s} \in \mathcal{O}_d$, $\tilde{I}$ is a rate function on $\mathcal{V} \times \mathbf{H}^{-\mathbf{s}}$. Furthermore:*

(i) *The sequence $\{(V^N, \mathcal{J}^N), N \in \mathbb{N}\}$ satisfies the large deviation upper bound on $\mathcal{V} \times \mathbf{H}^{-\mathbf{s}}$ with speed a_N and rate function $\tilde{I}$;*
(ii) *If in addition Condition 2.4 holds, then $\{(V^N, \mathcal{J}^N), N \in \mathbb{N}\}$ satisfies the large deviation lower bound on $\mathcal{V} \times \mathbf{H}^{-\mathbf{s}}$ with speed a_N and rate function $\tilde{I}$.*

When $m = d$ and $\sigma(\mu)$ is invertible, one can give a more explicit representation for the rate function $\tilde{I}$ as follows. (A similar representation can be found in [29] for the case $\sigma = \text{Id}$.) For $\Theta \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z})$ and $V \in \mathcal{V}$ with $V = \nu_\Theta$, define

$$\eta_t^\Theta \doteq \Theta \circ (X(t), \sigma(V(t))v(t) + b(X(t), V(t)))^{-1},$$

where $v(t) = \int_{\mathbb{R}^d} y \, \rho_t(dy)$ and ρ_t is obtained from the disintegration of ρ as $\rho(dt, dy) = \rho_t(dy) \, dt$. Note that, since $V = \nu_\Theta$, η_t^Θ can be disintegrated as $\eta_t^\Theta(dx, dy) = \hat{\eta}_t^\Theta(x, dy) \, V(t, dx)$ for some $\hat{\eta}_t^\Theta$. Define $\tilde{I}_0 : \mathcal{V} \times \mathbf{H}^{-\mathbf{s}} \to [0, \infty]$ as

$$\tilde{I}_0(V, \mathcal{J}) \doteq \inf \left\{ \frac{1}{2} \int_0^T \left\langle V(t), \int_{\mathbb{R}^d} \left| \sigma^{-1}(V(t))(y - b(\cdot, V(t))) \right|^2 \hat{\eta}_t^\Theta(\cdot, dy) \right\rangle dt \right\}, \quad (2.17)$$

where the infimum is taken over all $\Theta \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z})$ such that $V = \nu_\Theta$ and with $h(t, x) = \int_{\mathbb{R}^d} y \, \hat{\eta}_t^\Theta(x, dy)$, $(V, \mathcal{J})$ is a distributional-sense solution of the equation

$$\frac{\partial}{\partial t} V + \nabla \cdot hV = 0, \qquad \mathcal{J} = hV, \qquad V(0) = \mu_0 \qquad (2.18)$$

on $(0, T) \times \mathbb{R}^d$. Namely, for all $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$

$$\int_0^T \left\langle V(t), \frac{\partial}{\partial t} \varphi(t, \cdot) \right\rangle dt + \int_0^T \langle V(t), \nabla \varphi(t, \cdot) \cdot h(t, \cdot) \rangle \, dt = 0,$$

and for all $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$

$$\langle \mathcal{J}, \varphi \rangle = \int_0^T \langle V(t), \varphi(t, \cdot) \cdot h(t, \cdot) \rangle \, dt. \tag{2.19}$$

The following result shows that $\tilde{I} = \tilde{I}_0$. The proof is given in Subsection 3.7.

Proposition 2.12. *Suppose that $m = d$, Conditions 2.1 and 2.4(i) are satisfied, and $\sigma(\mu)$ is invertible for every $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. Then $\tilde{I} = \tilde{I}_0$.*

3. Laplace asymptotics and variational representation

Using the well-known equivalence (*cf.* [7, 11]) between the large deviation upper bound (respectively lower bound) and the Laplace upper bound (respectively lower bound), we will prove Theorem 2.10 by establishing a Laplace principle on the space $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$. Specifically, Theorem 2.10(i) will follow from the upper bound

$$\liminf_{N \to \infty} -\frac{1}{a_N} \log E \left[e^{-a_N F(\mu^N, \mathcal{J}^N)} \right]$$
$$\geq \inf_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}} (F(\mu, \mathcal{J}) + I(\mu, \mathcal{J})), \tag{3.1}$$

and Theorem 2.10(ii) will follow from the lower bound

$$\limsup_{N \to \infty} -\frac{1}{a_N} \log E \left[e^{-a_N F(\mu^N, \mathcal{J}^N)} \right]$$
$$\leq \inf_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}} (F(\mu, \mathcal{J}) + I(\mu, \mathcal{J})), \tag{3.2}$$

where F is any bounded, continuous function on $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$.

The inequality (3.1) will be proved in Subsection 3.3 (under Conditions 2.1 and 2.3), and the inequality (3.2) will be proved in Subsection 3.4 (under Conditions 2.1, 2.3, and 2.4). The rate function property of I is shown in Subsection 3.5. The starting point for both upper and lower bounds is the following variational representation.

3.1. Variational representation

Let $\mathcal{A}_N$ denote the class of $\mathbb{R}^{Nm}$-valued $\mathcal{F}(t)$-progressively measurable processes u such that $E \left[\int_0^T |u(t)|^2 \, dt \right] < \infty$. For $u^N = (u_1^N, \ldots, u_N^N) \in \mathcal{A}_N$, with each

$u_j^N(t)$ taking values in $\mathbb{R}^m$, consider the controlled version of (2.1) given as

$$d\bar{X}_j^N(t) = b\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) dt + \varepsilon_N \sigma\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) dW_j(t)$$
$$+ \sigma\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) u_j^N(t) dt, \tag{3.3}$$

where $\bar{X}_j^N(0) = x_j^N$ and

$$\bar{V}^N(t) \doteq \frac{1}{N} \sum_{j=1}^{N} \delta_{\bar{X}_j^N(t)}, \qquad 0 \le t \le T.$$

Analogous to (2.3), $\bar{\mu}^N$ will denote the empirical measure of $(\bar{X}_1^N, \ldots, \bar{X}_N^N)$, so that $\bar{\mu}^N \circ \pi_t^{-1} = \bar{V}^N(t)$ for each $0 \le t \le T$. We will also need a controlled analogue of the stochastic current in Theorem 2.6. For $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, define

$$\bar{J}_j^N(\varphi) \doteq \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \circ d\bar{X}_j^N(t), \qquad \bar{J}^N(\varphi) \doteq \frac{1}{N} \sum_{j=1}^N \bar{J}_j^N(\varphi). \tag{3.4}$$

The proof of the following result, which is given in the Appendix, is similar to that of Theorem 2.6.

Lemma 3.1. *Suppose that Conditions 2.1 and 2.3 hold. Then, for each $N \in \mathbb{N}$, $1 \le j \le N$, and $\mathbf{s} \in \mathcal{O}_d$, there is a nonnegative square-integrable random variable $C_{j,\mathbf{s}}^N$ such that for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$*

$$\left|\bar{J}_j^N(\varphi)\right| \le C_{j,\mathbf{s}}^N \|\varphi\|_{\mathbf{s}} \qquad a.s.$$

In particular, the collection $\{\varphi \mapsto \bar{J}^N(\varphi)\}$ has a pathwise realization $\bar{\mathcal{J}}^N$ on $(\Omega, \mathcal{F}, P)$, namely $\bar{\mathcal{J}}^N$ is an $\mathbf{H}^{-\mathbf{s}}$-valued random variable such that $\langle \bar{\mathcal{J}}^N(\omega), \varphi \rangle = [\bar{J}^N(\varphi)](\omega)$ for a.e. $\omega \in \Omega$ and all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$. Furthermore, if

$$\sup_{N \ge 1} E\left[\frac{1}{N} \sum_{j=1}^N \int_0^T \left|u_j^N(t)\right|^2 dt\right] < \infty, \tag{3.5}$$

then $\sup_{N \ge 1} E\left[\frac{1}{N} \sum_{j=1}^N \left(C_{j,\mathbf{s}}^N\right)^2\right] < \infty$. In particular, if $C_{\mathbf{s}}^N \doteq \frac{1}{N} \sum_{j=1}^N C_{j,\mathbf{s}}^N$, then $\sup_{N \ge 1} E\left[\left(C_{\mathbf{s}}^N\right)^2\right] < \infty$.

The following variational representation follows from [2, 6] (see also [8]). Specifically, the case where $\{\mathcal{F}(t)\}$ is the filtration generated by the m-dimensional Brownian motions $\{W_j\}$ is covered in [2], while the setting of a general filtration is treated in [6]. Recall that $a_N = N/\varepsilon_N^2$.

Theorem 3.2 (Variational Representation). *Suppose that Conditions* 2.1 *and* 2.3 *hold. Let* $\mathbf{s} \in \mathcal{O}_d$ *and let* F *be a real-valued, bounded, continuous function on* $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$. *Then for each* $N \in \mathbb{N}$,

$$
\begin{aligned}
&-\frac{1}{a_N} \log E\left[e^{-a_N F(\mu^N, \mathcal{J}^N)}\right] \\
&= \inf_{u^N \in \mathcal{A}_N} E\left[\frac{1}{2N} \sum_{j=1}^{N} \int_0^T \left|u_j^N(t)\right|^2 dt + F\left(\bar{\mu}^N, \bar{\mathcal{J}}^N\right)\right].
\end{aligned}
\tag{3.6}
$$

3.2. Tightness properties

The following lemma gives a key tightness property that will be needed in the proofs of both upper and lower Laplace bounds. The proof is given in Subsection 4.1.

Lemma 3.3. *Suppose Conditions* 2.1 *and* 2.3 *hold. Fix* $\mathbf{s} \in \mathcal{O}_d$, *and let* $\{u^N, N \in \mathbb{N}\}$ *with* $u_N \in \mathcal{A}_N$ *for each* N *be such that*

$$
\sup_{N \geq 1} E\left[\frac{1}{N} \sum_{j=1}^{N} \int_0^T \left|u_j^N(t)\right|^2 dt\right] < \infty.
$$

Let $\bar{X}_j^N$, $\bar{\mu}^N$, *and* $\bar{\mathcal{J}}^N$ *be the controlled sequences corresponding to sequence of controls* $\{u^N\}$ *as defined in Subsection* 3.1. *For each* j *and* N, *let* ρ_j^N *be the* $\mathcal{R}_1$-*valued random variable given as*

$$
\rho_j^N(dt, dy) \doteq \delta_{u_j^N(t)}(dy)\, dt,
$$

and consider the sequence of $\mathcal{P}(\mathcal{Z})$-*valued random variables defined as*

$$
Q^N \doteq \frac{1}{N} \sum_{j=1}^{N} \delta_{\left(\bar{X}_j^N, \rho_j^N\right)}, \qquad N \in \mathbb{N}.
\tag{3.7}
$$

Then:

(i) *The sequence* $\{(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N), N \in \mathbb{N}\}$ *is tight in* $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$;
(ii) *If* $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N) \Rightarrow (\bar{\mu}, Q, \bar{\mathcal{J}})$ *as* $N \to \infty$ *in* $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$, *then* $Q_{(1)} = \bar{\mu}$ *and* $Q \in \mathcal{P}^*(\bar{\mathcal{J}})$ *a.s.*

3.3. Proof of the upper bound

In this subsection we prove part (i) of Theorem 2.10 by showing that (3.1) holds. Assume Conditions 2.1 and 2.3. Fix $\mathbf{s} = (s_1, s_2) \in \mathcal{O}_d$, and a real-valued,

bounded, continuous function F on $\mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-s}$. Let $\varepsilon \in (0, 1)$, and using Theorem 3.2 choose $\{u^N, N \in \mathbb{N}\}$ with $u_N \in \mathcal{A}_N$ for each N such that

$$
\begin{aligned}
&-\frac{1}{a_N} \log E\left[e^{-a_N F(\mu^N, \mathcal{J}^N)}\right] \\
&\geq E\left[\frac{1}{2N} \sum_{j=1}^{N} \int_0^T \left|u_j^N(t)\right|^2 dt + F\left(\bar{\mu}^N, \bar{\mathcal{J}}^N\right)\right] - \varepsilon,
\end{aligned}
\tag{3.8}
$$

where $(\bar{\mu}^N, \bar{\mathcal{J}}^N)$ are controlled variables corresponding to the control u^N as defined in Subsection 3.1. From the boundedness of F it follows that

$$
\sup_{N \geq 1} E\left[\frac{1}{2N} \sum_{j=1}^{N} \int_0^T \left|u_j^N(t)\right|^2 dt\right] \leq 2 \sup_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-s}} |F(\mu, \mathcal{J})| + 1 < \infty.
$$

By Lemma 3.3, $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N)$ is tight in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-s}$. Thus the sequence $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N)$ has a weak limit point $(\bar{\mu}, Q, \bar{\mathcal{J}})$ along some subsequence, and once again by Lemma 3.3, $Q \in \mathcal{P}^*(\mathcal{J})$ and $Q_{(1)} = \bar{\mu}$ a.s. Assume without loss of generality that $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N) \Rightarrow (\bar{\mu}, Q, \bar{\mathcal{J}})$ along the full sequence. Noting that $Q_{(1)}^N = \bar{\mu}^N$, we have, by (3.8),

$$
\begin{aligned}
&-\frac{1}{a_N} \log E\left[e^{-a_N F(\mu^N, \mathcal{J}^N)}\right] \\
&\geq E\left[\frac{1}{2} \int_{\mathcal{R}_1} \int_{[0,T] \times \mathbb{R}^m} |y|^2 r(dt, dy) \, Q_{(2)}^N(dr) + F\left(Q_{(1)}^N, \bar{\mathcal{J}}^N\right)\right] - \varepsilon.
\end{aligned}
$$

By Fatou's lemma and lower semicontinuity of the map $r \mapsto \int_{[0,T] \times \mathbb{R}^m} |y|^2 r(dt, dy)$ on $\mathcal{R}_1$,

$$
\begin{aligned}
&\liminf_{N \to \infty} -\frac{1}{a_N} \log E\left[e^{-a_N F(\mu^N, \mathcal{J}^N)}\right] \\
&\geq E\left[\frac{1}{2} \int_{\mathcal{R}_1} \int_{[0,T] \times \mathbb{R}^m} |y|^2 r(dt, dy) \, Q_{(2)}(dr) + F\left(Q_{(1)}, \bar{\mathcal{J}}\right)\right] - \varepsilon \\
&= E\left[E_Q\left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy)\right] + F\left(\bar{\mu}, \bar{\mathcal{J}}\right)\right] - \varepsilon \\
&\geq \inf_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-s}} (I(\mu, \mathcal{J}) + F(\mu, \mathcal{J})) - \varepsilon,
\end{aligned}
$$

where the last line follows on recalling the definition of I and the facts that $Q \in \mathcal{P}^*(\mathcal{J})$ and $Q_{(1)} = \bar{\mu}$ a.s. Since $\varepsilon \in (0, 1)$ is arbitrary, this completes the proof of the upper bound in (3.1) and thus that of Theorem 2.10(i).

3.4. Proof of the lower bound

In this subsection we prove part (ii) of Theorem 2.10 by showing (3.2). Fix $\mathbf{s} = (s_1, s_2) \in \mathcal{O}_d$. We assume that Conditions 2.1, 2.3, and 2.4 hold. Let $\varepsilon \in (0, 1)$ and choose $(\Theta_0, \mathcal{J}_0) \in \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$ such that $\Theta_0 \in \mathcal{P}^*(\mathcal{J}_0)$ and

$$
\begin{aligned}
&E_{\Theta_0} \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] + F \left((\Theta_0)_{(1)}, \mathcal{J}_0 \right) \\
&\leq \inf_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}} (I(\mu, \mathcal{J}) + F(\mu, \mathcal{J})) + \varepsilon.
\end{aligned}
\tag{3.9}
$$

To prove the lower bound we will construct a sequence $\{u^N\}$ of controls on some filtered probability space such that $u^N \in \mathcal{A}_N$ for each N and

$$
\begin{aligned}
&\limsup_{N \to \infty} E \left[\frac{1}{2N} \sum_{j=1}^{N} \int_0^T \left| u_j^N(t) \right|^2 \, dt + F \left(\bar{\mu}^N, \bar{\mathcal{J}}^N \right) \right] \\
&\leq E_{\Theta_0} \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] + F \left((\Theta_0)_{(1)}, \mathcal{J}_0 \right),
\end{aligned}
\tag{3.10}
$$

where $\bar{\mu}^N$ and $\bar{\mathcal{J}}^N$ are the controlled processes corresponding to $\{u^N\}$. It will then follow by Theorem 3.2 and (3.9) that

$$
\begin{aligned}
&\limsup_{N \to \infty} -\frac{1}{a_N} \log E \left[e^{-a_N F(\mu^N, \mathcal{J}^N)} \right] \\
&\leq \limsup_{N \to \infty} E \left[\frac{1}{2N} \sum_{j=1}^{N} \int_0^T \left| u_j^N(t) \right|^2 \, dt + F \left(\bar{\mu}^N, \bar{\mathcal{J}}^N \right) \right] \\
&\leq \inf_{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}} (I(\mu, \mathcal{J}) + F(\mu, \mathcal{J})) + \varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ is arbitrary, the lower bound follows.

The construction of a sequence $\{u^N\}$ such that the inequality in (3.10) holds will need the following uniqueness property.

Definition 3.4. Let $\theta : \mathcal{Z} \to \mathbb{R}^d \times \mathcal{R}_1$ denote the map $\theta(\xi, r) = (\xi(0), r)$. We say that weak uniqueness of solutions of (2.9) holds if $\Theta_1, \Theta_2 \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z})$ and $\Theta_1 \circ \theta^{-1} = \Theta_2 \circ \theta^{-1}$ imply that $\Theta_1 = \Theta_2$.

The following lemma is key to the proof of the lower bound. The proof is provided in Subsection 4.3. Recall that in this subsection we assume that Conditions 2.1, 2.3 and 2.4 hold.

Lemma 3.5. *Weak uniqueness of solutions holds for* (2.9).

We now construct the sequence $\{u^N\}$ that satisfies (3.10). Because $\Theta_0 \in \mathcal{S}(\mathcal{Z})$, we can disintegrate

$$\Theta_0 \circ \theta^{-1}(dx\,dr) = \mu_0(dx)\,\Lambda_0(x, dr)$$

for some measurable map $\Lambda_0 : \mathbb{R}^d \to \mathcal{P}(\mathcal{R}_1)$. Let $\mathcal{W} \doteq \mathcal{C}([0, T], \mathbb{R}^m)$, and let γ be the standard Wiener measure on $\mathcal{W}$. Define a measurable map $\Lambda : \mathbb{R}^d \to \mathcal{P}(\mathcal{R}_1 \times \mathcal{W})$ as

$$\Lambda(x, dr, dw) \doteq \Lambda_0(x, dr) \otimes \gamma(dw), \qquad x \in \mathbb{R}^d.$$

Define the measurable space $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by

$$\tilde{\Omega} = (\mathcal{R}_1 \times \mathcal{W})^\infty, \quad \tilde{\mathcal{F}} = \mathcal{B}(\tilde{\Omega}),$$

where an element $(r, w) \in \tilde{\Omega}$ has the coordinates $r = (r_1, r_2, \ldots)$ and $w = (w_1, w_2, \ldots)$ with $r_j \in \mathcal{R}_1$ and $w_j \in \mathcal{W}$ for each j. Consider the canonical filtration $\{\tilde{\mathcal{F}}(t)\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ defined as

$$\tilde{\mathcal{F}}(t) \doteq \sigma\left(w_j(s), r_j([0, s] \times A), j \in \mathbb{N}, A \in \mathcal{B}(\mathbb{R}^m), s \leq t\right), \qquad 0 \leq t \leq T,$$

and define the sequence $\{P^N, N \in \mathbb{N}\}$ of probability measures on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by

$$P^N(dr, dw) = \bigotimes_{j \leq N} \Lambda\left(x_j^N, dr_j, dw_j\right) \bigotimes_{j > N} \left((\Theta_0)_{(2)} \otimes \gamma\right)(dr_j, dw_j),$$

where $\{x_j^N\}$ are as in Condition 2.3. Next define the sequence $\{\Lambda^N, N \in \mathbb{N}\}$ of $\mathcal{P}(\mathbb{R}^d \times \mathcal{R}_1)$-valued random variables on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by

$$\Lambda^N \doteq \frac{1}{N} \sum_{j=1}^{N} \delta_{\left(x_j^N, \rho_j\right)},$$

where for each $j \in \mathbb{N}$, ρ_j is the $\mathcal{R}_1$-valued random variable on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ defined as $\rho_j(r, w) = r_j$. Using Condition 2.4(ii), we see by a standard argument that

$$P^N \circ \left(\Lambda^N\right)^{-1} \to \delta_{\Theta_0 \circ \theta^{-1}} \tag{3.11}$$

in $\mathcal{P}(\mathcal{P}(\mathbb{R}^d \times \mathcal{R}_1))$ as $N \to \infty$.

Now, for each $j \in \mathbb{N}$, disintegrating ρ_j as $\rho_j(dt, dy) = (\rho_j)_t(dy)\,dt$, define

$$u_j(t) \doteq \int_{\mathbb{R}^m} y\,(\rho_j)_t(dy), \qquad 0 \leq t \leq T,$$

and define $u^N \doteq (u_1, \ldots, u_N)$ for each $N \in \mathbb{N}$. Furthermore, for each j and $(r, w) \in \tilde{\Omega}$, let

$$W_j(t, (r, w)) \doteq w_j(t), \qquad 0 \le t \le T.$$

Then, for each N, $W_1, \ldots, W_N$ are mutually independent $\{\tilde{\mathcal{F}}(t)\}$-Brownian motions on $(\tilde{\Omega}, \tilde{\mathcal{F}}, P^N)$. Recall that in this subsection we are assuming Condition 2.4, and so $\sigma(x, \nu) = \sigma(\nu)$ for $(x, \nu) \in \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$. Let $(\bar{X}_1^N, \ldots, \bar{X}_N^N)$ be the unique pathwise solution (which is guaranteed due to Conditions 2.1 and 2.3) on $(\tilde{\Omega}, \tilde{\mathcal{F}}, P^N)$ of the system

$$\bar{X}_j^N(t) = x_j^N + \int_0^t b\left(\bar{X}_j^N(s), \bar{V}^N(s)\right) ds + \varepsilon_N \int_0^t \sigma\left(\bar{V}^N(s)\right) dW_j(s)$$

$$+ \int_0^t \sigma\left(\bar{V}^N(s)\right) u_j(s) ds,$$

$$\bar{V}^N(t) = \frac{1}{N} \sum_{j=1}^N \delta_{\bar{X}_j^N(t)}, \qquad 0 \le t \le T, \qquad 1 \le j \le N.$$

Also let $\bar{\mu}^N = \frac{1}{N} \sum_{j=1}^N \delta_{\bar{X}_j^N}$. Now define the sequence $\{Q^N\}$ of $\mathcal{P}(\mathcal{Z})$-valued random variables as

$$Q^N \doteq \frac{1}{N} \sum_{j=1}^N \delta_{\left(\bar{X}_j^N, \rho_j\right)}, \qquad N \in \mathbb{N}.$$

Letting E^N denote expectation on $(\tilde{\Omega}, \tilde{\mathcal{F}}, P^N)$, we note that for a measurable $f : \mathcal{R}_1 \to \mathbb{R}_+$

$$\int_{\mathcal{R}_1} f(r) (\Theta_0)_{(2)}(dr) < \infty$$

$$\text{implies} \qquad E^N\left[\frac{1}{N} \sum_{j=1}^N f(\rho_j)\right] \to \int_{\mathcal{R}_1} f(r) (\Theta_0)_{(2)}(dr). \tag{3.12}$$

Indeed, if $g(x) = \int_{\mathcal{R}_1} f(r) \Lambda_0(x, dr)$ for $x \in \mathbb{R}^d$, then

$$E^N\left[\frac{1}{N} \sum_{j=1}^N f(\rho_j)\right] = \frac{1}{N} \sum_{j=1}^N \int_{\mathcal{R}_1} f(r) \Lambda_0\left(x_j^N, dr\right) = \frac{1}{N} \sum_{j=1}^N g\left(x_j^N\right),$$

and

$$\int_{\mathbb{R}^d} g(x) \mu_0(dx) = \int_{\mathbb{R}^d} \int_{\mathcal{R}_1} f(r) \Lambda_0(x, dr) \mu_0(dx)$$

$$= \int_{\mathbb{R}^d \times \mathcal{R}_1} f(r) \Theta_0 \circ \theta^{-1}(dx, dr) = \int_{\mathcal{R}_1} f(r) (\Theta_0)_{(2)}(dr) < \infty.$$

Thus, from Condition 2.4(ii),

$$\lim_{N\to\infty} E^N \left[\frac{1}{N} \sum_{j=1}^N f(\rho_j) \right] = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N g\left(x_j^N\right)$$

$$= \int_{\mathbb{R}^d} g(x)\, \mu_0(dx) = \int_{\mathcal{R}_1} f(r)\, (\Theta_0)_{(2)}(dr), \tag{3.13}$$

which proves (3.12). Now, we have

$$\limsup_{N\to\infty} E^N \left[\frac{1}{N} \sum_{j=1}^N \int_0^T |u_j(t)|^2\, dt \right]$$

$$\leq \limsup_{N\to\infty} E^N \left[\frac{1}{N} \sum_{j=1}^N \int_{[0,T]\times\mathbb{R}^m} |y|^2\, \rho_j(dt, dy) \right] \tag{3.14}$$

$$= E_{\Theta_0} \left[\int_{[0,T]\times\mathbb{R}^m} |y|^2\, \rho(dt, dy) \right] < \infty,$$

where the convergence on the second line follows from (3.12) on observing that, since $\Theta_0 \in \mathcal{P}_2(\mathcal{Z})$,

$$f(r) = \int_{[0,T]\times\mathbb{R}^m} |y|^2\, r(dt, dy), \qquad r \in \mathcal{R}_1$$

satisfies

$$\int_{\mathcal{R}_1} f(r)\, (\Theta_0)_{(2)}(dr) = E_{\Theta_0} \left[\int_{[0,T]\times\mathbb{R}^m} |y|^2\, \rho(dt, dy) \right] < \infty.$$

Next, for each $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ define

$$\bar{J}^N(\varphi) \doteq \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \circ d\bar{X}_j^N(t), \qquad N \in \mathbb{N}.$$

From Lemma 3.1, the collection $\{\varphi \mapsto \bar{J}^N(\varphi)\}$ has a pathwise realization $\bar{\mathcal{J}}^N$ in $\mathbf{H}^{-s}$. Using Lemma 3.3 and the moment bound in (3.14), we now see that $\{(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N), N \in \mathbb{N}\}$ is tight in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-s}$. Suppose, without loss of generality, that $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N) \Rightarrow (\bar{\mu}, Q, \bar{\mathcal{J}})$ in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-s}$. By Lemma 3.3 again, $Q \in \mathcal{P}^*(\bar{\mathcal{J}})$ and $Q_{(1)} = \bar{\mu}$ a.s. Since $Q^N \circ \theta^{-1} = \Lambda^N$, (3.11) implies that $Q \circ \theta^{-1} = \Theta_0 \circ \theta^{-1}$ a.s., and hence by the weak uniqueness established in Lemma 3.5, $Q = \Theta_0$ a.s. Furthermore, from the definition of $\mathcal{P}^*(\bar{\mathcal{J}})$,

$$\langle \bar{\mathcal{J}}, \varphi \rangle = G_\varphi(Q) = G_\varphi(\Theta_0) = \langle \mathcal{J}_0, \varphi \rangle$$

for every φ, a.s., and hence $\bar{\mathcal{J}} = \mathcal{J}_0$ a.s. by separability of $\mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ and its denseness in $\mathbf{H}^{-\mathbf{s}}$.

It follows that $(Q^N, \bar{\mathcal{J}}^N) \Rightarrow (\Theta_0, \mathcal{J}_0)$. Finally,

$$\limsup_{N \to \infty} E^N \left[\frac{1}{2N} \sum_{j=1}^N \int_0^T |u_j(t)|^2 \, dt + F\left(\bar{\mu}^N, \bar{\mathcal{J}}^N\right) \right]$$

$$= \limsup_{N \to \infty} E^N \left[\frac{1}{2N} \sum_{j=1}^N \int_0^T |u_j(t)|^2 \, dt + F\left(Q_{(1)}^N, \bar{\mathcal{J}}^N\right) \right]$$

$$\leq E_{\Theta_0} \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] + F\left((\Theta_0)_{(1)}, \mathcal{J}_0\right),$$

where the last inequality is from (3.14) and since F is a bounded continuous function. This shows (3.10) and completes the proof of the lower bound in (3.2), and part (ii) of Theorem 2.10 follows.

3.5. Rate function property

In this subsection we show that the function $I : \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}} \to [0, \infty]$ defined in (2.11) has compact sublevel sets for every $\mathbf{s} \in \mathcal{O}_d$. Fix $\mathbf{s}$, and for each $l < \infty$ consider the level set $\Gamma_l \doteq \{(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}} : I(\mu, \mathcal{J}) \leq l\}$. The proof of the following lemma is given in Subsection 4.2.

Lemma 3.6. *Suppose Conditions* 2.1 *and* 2.3 *hold. Let* $\mathbf{s} \in \mathcal{O}_d$ *and let* $\{(\mu_k, \Theta_k, \mathcal{J}_k), k \in \mathbb{N}\}$ *be a sequence in* $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$ *such that, for each* k, $\Theta_k \in \mathcal{P}^*(\mathcal{J}_k)$, $(\Theta_k)_{(1)} = \mu_k$, *and*

$$\sup_{k \geq 1} E_{\Theta_k} \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] < \infty. \tag{3.15}$$

Then the sequence $\{(\mu_k, \Theta_k, \mathcal{J}_k), k \in \mathbb{N}\}$ *is relatively compact in* $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$.

Now we prove the compactness of Γ_l. Let $\{(\mu_k, \mathcal{J}_k), k \in \mathbb{N}\}$ be a sequence in Γ_l. From the definition of I, for each $k \in \mathbb{N}$ there is a $\Theta_k \in \mathcal{P}^*(\mathcal{J}_k)$ with $(\Theta_k)_{(1)} = \mu_k$ such that

$$E_{\Theta_k} \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] \leq l + \frac{1}{k}. \tag{3.16}$$

From Lemma 3.6, $\{(\mu_k, \Theta_k, \mathcal{J}_k)\}$ is relatively compact in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$. It is easily checked that if $(\mu, \Theta, \mathcal{J})$ is a limit point along some subsequence, then $\Theta_{(1)} = \mu$ and along the same subsequence $G_\varphi(\Theta_k) \to G_\varphi(\Theta)$ and $\langle \mathcal{J}_k, \varphi \rangle \to \langle \mathcal{J}, \varphi \rangle$ for every $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$. This shows that $\Theta \in \mathcal{P}^*(\mathcal{J})$. Sending $k \to \infty$ in

(3.16) and using lower semicontinuity of the map $r \mapsto \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, r(dt, dy)$ on $\mathcal{R}_1$, we obtain

$$E_\Theta \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \leq l,$$

and hence $(\mu, \mathcal{J})$ lies in Γ_l. Compactness of Γ_l follows.

3.6. Law of large numbers

Here we prove Theorem 2.8. The model (2.1) can be viewed as the controlled equation (3.3) with the controls taken to be $u_j^N \equiv 0$ for all $1 \leq j \leq N$ and $N \in \mathbb{N}$. From Lemma 3.3 it then follows that $(\mu^N, Q^N, \mathcal{J}^N)$ is tight in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-s}$. Suppose that along some subsequence $(\mu^N, Q^N, \mathcal{J}^N) \Rightarrow (\mu, Q, \mathcal{J})$. Then, once again from Lemma 3.3, $Q_{(1)} = \mu$ and $Q \in \mathcal{P}^*(\mathcal{J})$ a.s. Furthermore, since $u_j^N \equiv 0$ for all $1 \leq j \leq N$ and $N \in \mathbb{N}$, we see that the second coordinate variable on $\mathcal{Z}$ satisfies $Q(\rho = 0) = 1$ a.s., and thus, under Q, the first coordinate variable on $\mathcal{Z}$ satisfies

$$X(t) = X(0) + \int_0^t b(X(s), V(s)) \, ds, \qquad V(t) = Q \circ X(t)^{-1}, \qquad V(0) = \mu_0$$

for all $0 \leq t \leq T$. Then, from the unique solvability of (2.12), it follows that $\mu = \mu^*$ a.s., and hence we have that μ^N converges in probability in $\mathcal{P}_1(\mathcal{X})$ (along the full sequence) to μ^*. Since $V^N(t) = \mu^N \circ \pi_t^{-1}$ and $V^*(t) = \mu^* \circ \pi_t^{-1}$ for each $0 \leq t \leq T$, we also have that $V^N \to V^*$ in probability in $\mathcal{V}$. Finally, since $Q \in \mathcal{P}^*(\mathcal{J})$ a.s.,

$$G_\varphi(Q) = \langle \mathcal{J}, \varphi \rangle$$

for all $\varphi \in C_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, a.s., and note that

$$G_\varphi(Q) = E_Q \left[\int_0^T \varphi(t, X(t)) \cdot dX(t) \right]$$

$$= E_Q \left[\int_0^T \varphi(t, X(t)) \cdot b(X(t), V^*(t)) \, dt \right]$$

$$= \int_0^T \langle V^*(t), \varphi(t, \cdot) \cdot b(\cdot, V^*(t)) \rangle \, dt.$$

Thus $\langle \mathcal{J}, \varphi \rangle$ is (a.s.) uniquely characterized for all $\varphi \in C_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$. From the separability of $C_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ and its denseness in $\mathbf{H}^{-s}$ we now see that $\mathcal{J}^N$ converges (along the full sequence) in probability, in $\mathbf{H}^{-s}$, to the nonrandom limit $\mathcal{J}^*$ characterized as

$$\langle \mathcal{J}^*, \varphi \rangle = \int_0^T \langle V^*(t), \varphi(t, \cdot) \cdot b(\cdot, V^*(t)) \rangle \, dt.$$

The result follows.

3.7. Equivalent formulation of the rate function

In this subsection we give the proof of Proposition 2.12. Let $m = d$, and suppose that, for every $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, $\sigma(\mu)$ is invertible. We first argue that $\tilde{I}_0 \leq \tilde{I}$. Fix $(V, \mathcal{J}) \in \mathcal{V} \times \mathbf{H}^{-s}$ such that $\tilde{I}(V, \mathcal{J}) < \infty$. Fix $\delta > 0$ and let $\Theta \in \mathcal{P}^*(\mathcal{J})$ with $v_\Theta = V$ be δ-optimal for $\tilde{I}(V, \mathcal{J})$, namely

$$E_\Theta \left[\frac{1}{2} \int_{[0,T] \times \mathbb{R}^d} |y|^2 \, \rho(dt, dy) \right] \leq \tilde{I}(V, \mathcal{J}) + \delta. \qquad (3.17)$$

Disintegrate $\rho(dt, dy) = \rho_t(dy) \, dt$ and define

$$v(t) \doteq \int_{\mathbb{R}^d} y \, \rho_t(dy), \qquad \text{a.e. } t \in [0, T]. \qquad (3.18)$$

Also let $\eta_t^\Theta \doteq \Theta \circ (X(t), \sigma(V(t))v(t) + b(X(t), V(t)))^{-1} \in \mathcal{P}(\mathbb{R}^{2d})$. Then, since $v_\Theta = V$, η_t^Θ can be disintegrated as $\eta_t^\Theta(dx, dy) = \hat{\eta}_t^\Theta(x, dy) \, V(t, dx)$ for some $\hat{\eta}_t^\Theta : \mathbb{R}^d \to \mathcal{P}(\mathbb{R}^d)$. Define the function h on $[0, T] \times \mathbb{R}^d$ by

$$h(t, x) \doteq \int_{\mathbb{R}^d} y \, \hat{\eta}_t^\Theta(x, dy), \qquad (3.19)$$

and note that Condition 2.1 ensures that this is well-defined. Under Θ, $V(0) = \mu_0$ and

$$X(t) = X(0) + \int_0^t b(X(s), V(s)) \, ds + \int_{[0,t] \times \mathbb{R}^d} \sigma(V(s)) y \, \rho_s(dy) \, ds, \quad \text{a.s.,} \quad (3.20)$$

for each t, and so for $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$

$$0 = \varphi(T, X(T)) - \varphi(0, X(0))$$
$$= \int_0^T \left(\frac{\partial}{\partial t} \varphi(t, X(t)) + \nabla \varphi(t, X(t)) \cdot b(X(t), V(t)) + \nabla \varphi(t, X(t)) \cdot \sigma(V(t)) v(t) \right) dt,$$

where v is as in (3.18). Taking expectations with respect to Θ,

$$0 = \int_0^T \int_{\mathbb{R}^{2d}} \left[\frac{\partial}{\partial t} \varphi(t, x) + \nabla \varphi(t, x) \cdot y \right] \eta_t^\Theta(dx, dy) \, dt$$
$$= \int_0^T \left\langle V(t), \frac{\partial}{\partial t} \varphi(t, \cdot) + \nabla \varphi(t, \cdot) \int_{\mathbb{R}^d} y \, \hat{\eta}_t^\Theta(\cdot, dy) \right\rangle dt \qquad (3.21)$$
$$= \int_0^T \left\langle V(t), \frac{\partial}{\partial t} \varphi(t, \cdot) + \nabla \varphi(t, \cdot) \cdot h(t, \cdot) \right\rangle dt.$$

Similarly, since $\langle \mathcal{J}, \varphi \rangle = G_\varphi(\Theta)$, it is seen that for $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$

$$\langle \mathcal{J}, \varphi \rangle = \int_0^T \langle V(t), \varphi(t, \cdot) \cdot h(t, \cdot) \rangle \, dt.$$

Since $V = \nu_\Theta$, we now see from the above two identities that

$$
\begin{aligned}
\tilde{I}_0(V, \mathcal{J}) &\leq \frac{1}{2} \int_0^T \left\langle V(t), \int_{\mathbb{R}^d} \left| \sigma^{-1}(V(t))(y - b(\cdot, V(t))) \right|^2 \hat{\eta}_t^\Theta(\cdot, dy) \right\rangle dt \\
&= \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2d}} \left| \sigma^{-1}(V(t))(y - b(x, V(t))) \right|^2 \eta_t^\Theta(dx, dy) \, dt \\
&= \frac{1}{2} \int_0^T E_\Theta \left[|v(t)|^2 \right] dt \leq \frac{1}{2} E_\Theta \left[\int_{[0,T] \times \mathbb{R}^d} |y|^2 \rho(dt, dy) \right] \\
&\leq \tilde{I}(V, \mathcal{J}) + \delta,
\end{aligned}
$$

where the last inequality is from (3.17). Since $\delta > 0$ is arbitrary, the inequality $\tilde{I}_0(V, \mathcal{J}) \leq \tilde{I}(V, \mathcal{J})$ follows.

We now prove the reverse inequality, namely $\tilde{I}(V, \mathcal{J}) \leq \tilde{I}_0(V, \mathcal{J})$. Once more fix $\delta > 0$ and $(V, \mathcal{J}) \in \mathcal{V} \times \mathbf{H}^{-s}$ such that $\tilde{I}_0(V, \mathcal{J}) < \infty$, and let $\Theta \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z})$ be δ-optimal for $\tilde{I}_0(V, \mathcal{J})$, namely

$$
\frac{1}{2} \int_0^T \left\langle V(t), \int_{\mathbb{R}^d} \left| \sigma^{-1}(V(t))(y - b(\cdot, V(t))) \right|^2 \hat{\eta}_t^\Theta(\cdot, dy) \right\rangle dt \leq \tilde{I}_0(V, \mathcal{J}) + \delta, \quad (3.22)
$$

$V = \nu_\Theta$, and $(V, \mathcal{J})$ solves (2.18) with $h(t, x) = \int_{\mathbb{R}^d} y \, \hat{\eta}_t^\Theta(x, dy)$. In particular, for all $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$, (2.19) holds.

Now define an $\mathcal{R}_1$-valued random variable $\tilde{\rho}$ on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ as

$$
\tilde{\rho}(dt, dy) = \delta_{v(t)}(dy) \, dt,
$$

where v is defined in terms of the coordinate variable ρ as in (3.18). Defining $\tilde{\Theta} \in \mathcal{P}(\mathcal{Z})$ as $\tilde{\Theta} \doteq \Theta \circ (X, \tilde{\rho})^{-1}$, we have that $\nu_{\tilde{\Theta}} = \nu_\Theta = V$, and it can be seen from (3.20) that $\tilde{\Theta} \in \mathcal{S}(\mathcal{Z})$. Also, since (2.19) holds for any $\varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$,

$$
\begin{aligned}
\langle \mathcal{J}, \varphi \rangle &= \int_0^T \langle V(t), \varphi(t, \cdot) \cdot h(\cdot, t) \rangle \, dt \\
&= \int_0^T \left\langle V(t), \varphi(t, \cdot) \cdot \int_{\mathbb{R}^d} y \, \hat{\eta}_t^\Theta(\cdot, dy) \right\rangle dt \\
&= \int_0^T \int_{\mathbb{R}^{2d}} \varphi(t, x) y \, \eta_t^\Theta(dx, dy) \, dt \\
&= E_\Theta \left[\int_{[0,T] \times \mathbb{R}^d} \varphi(t, X(t))[\sigma(V(t))y + b(X(t), V(t))] \, \rho(dt, dy) \right] \\
&= E_{\tilde{\Theta}} \left[\int_{[0,T] \times \mathbb{R}^d} \varphi(t, X(t))[\sigma(V(t))y + b(X(t), V(t))] \, \rho(dt, dy) \right] \\
&= G_\varphi(\tilde{\Theta}),
\end{aligned}
$$

where the last line uses the fact that $\int y\, \rho_t(dy) = v(t) = \int y\, \delta_{v(t)}(dy) = \int y\, \tilde{\rho}_t(dy)$. Thus, $\tilde{\Theta} \in \mathcal{P}^*(\mathcal{J})$. Finally,

$$
\begin{aligned}
\tilde{I}(V,\mathcal{J}) &\leq E_{\tilde{\Theta}}\left[\frac{1}{2}\int_{[0,T]\times\mathbb{R}^d}|y|^2\,\rho(dt,dy)\right] = E_{\Theta}\left[\frac{1}{2}\int_0^T|v(t)|^2\,dt\right] \\
&= E_{\Theta}\left[\frac{1}{2}\int_0^T\left|\sigma^{-1}(V(t))[\sigma(V(t))v(t)+b(X(t),V(t))-b(X(t),V(t))]\right|^2 dt\right] \\
&= \frac{1}{2}\int_0^T\int_{\mathbb{R}^{2d}}\left|\sigma^{-1}(V(t))[y-b(x,V(t))]\right|^2\,\eta_t(dx,dy)\,dt \\
&= \frac{1}{2}\int_0^T\left\langle V(t), \int_{\mathbb{R}^d}\left|\sigma^{-1}(V(t))(y-b(\cdot,V(t))\right|^2\,\hat{\eta}_t^{\Theta}(\cdot,dy)\right\rangle dt \\
&\leq \tilde{I}_0(V,\mathcal{J})+\delta,
\end{aligned}
$$

where we used (3.22). Since $\delta > 0$ is arbitrary, the inequality $\tilde{I}(V,\mathcal{J}) \leq \tilde{I}_0(V,\mathcal{J})$ follows and completes the proof of the lemma.

4. Proofs of key lemmas

In this section we provide proofs of the results used in showing the Laplace upper and lower bounds. First we establish two estimates that will be used in subsequent subsections.

Lemma 4.1. *Suppose Conditions 2.1 and 2.3 are satisfied. Let $u^N = (u_1^N,\ldots,u_N^N) \in \mathcal{A}_N$ and let $\bar{X}^N$ be as defined in (3.3). Then, for each $N \in \mathbb{N}$,*

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^N E\left[\left\|\bar{X}_j^N\right\|_\infty^2\right] \\
&\leq c\left(1 + \frac{1}{N}\sum_{j=1}^N\left|x_j^N\right|^2 + E\left[\frac{1}{N}\sum_{j=1}^N\int_0^T\left|u_j^N(t)\right|^2 dt\right]\right),
\end{aligned}
\tag{4.1}
$$

and for any $\varepsilon > 0$ and any $\{\mathcal{F}(t)\}$-stopping time τ taking values in $[0, T-\varepsilon]$

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^N E\left[\left|\bar{X}_j^N(\tau+\varepsilon)-\bar{X}_j^N(\tau)\right|^2\right] \\
&\leq c\varepsilon\left(1 + \frac{1}{N}\sum_{j=1}^N\left|x_j^N\right|^2 + E\left[\frac{1}{N}\sum_{j=1}^N\int_0^T\left|u_j^N(t)\right|^2 dt\right]\right),
\end{aligned}
$$

where $c < \infty$ does not depend on N, u^N, or ε.

Proof. Condition 2.1 (see (2.5)) implies

$$\left| b\left(\bar{X}_j^N(t), \bar{V}^N(t) \right) \right|^2 \le 3L^2 \left(1 + \left| \bar{X}_j^N(t) \right|^2 + \frac{1}{N} \sum_{j=1}^N \left| \bar{X}_j^N(t) \right|^2 \right),$$

and so from (3.3) and since $|\sigma| \le L$ and $\varepsilon_N \le 1$, we have

$$\left| \bar{X}_j^N(t) \right|^2$$

$$\le 4 \left| x_j^N \right|^2 + 4 \left| \int_0^t b\left(\bar{X}_j^N(s), \bar{V}^N(s) \right) ds \right|^2$$

$$+ 4 \left| \varepsilon_N \int_0^t \sigma\left(\bar{X}_j^N(s), \bar{V}^N(s) \right) dW_j(s) \right|^2$$

$$+ 4 \left| \int_0^t \sigma\left(\bar{X}_j^N(s), \bar{V}^N(s) \right) u_j^N(s) ds \right|^2$$

$$\le 4 \left| x_j^N \right|^2 + 12L^2 T \left(1 + \int_0^t \sup_{0 \le r \le s} \left| \bar{X}_j^N(r) \right|^2 ds + \frac{1}{N} \sum_{j=1}^N \int_0^t \sup_{0 \le r \le s} \left| \bar{X}_j^N(r) \right|^2 ds \right)$$

$$+ 4 \sup_{0 \le r \le t} \left| \int_0^r \sigma\left(\bar{X}_j^N(s), \bar{V}^N(s) \right) dW_j(s) \right|^2 + 4L^2 T \int_0^T \left| u_j^N(s) \right|^2 ds.$$

Hence by the Burkholder-Davis-Gundy inequality, and using boundedness of σ once more,

$$\frac{1}{N} \sum_{j=1}^N E\left[\sup_{0 \le s \le t} \left| \bar{X}_j^N(s) \right|^2 \right]$$

$$\le \frac{4}{N} \sum_{j=1}^N \left| x_j^N \right|^2 + 24L^2 T \left(1 + \int_0^t \frac{1}{N} \sum_{j=1}^N E\left[\sup_{0 \le r \le s} \left| \bar{X}_j^N(r) \right|^2 \right] ds \right)$$

$$+ 16L^2 T + 4L^2 T E\left[\frac{1}{N} \sum_{j=1}^N \int_0^T \left| u_j^N(s) \right|^2 ds \right].$$

The first statement in the lemma then follows by Gronwall's inequality (see [12, Theorem A.5.1]) with $c = 24(L^2 T + 1)e^{24L^2 T^2}$.

Next, for any $t \in [0, T - \varepsilon]$, the linear growth of b, boundedness of σ, and the Cauchy-Schwarz inequality give

$$
\left| \bar{X}_j^N (t + \varepsilon) - \bar{X}_j^N (t) \right|^2
$$

$$
\leq 4 \left| \int_t^{t+\varepsilon} b \left(\bar{X}_j^N (s), \bar{V}^N (s) \right) ds \right|^2 + 4 \left| \varepsilon_N \int_t^{t+\varepsilon} \sigma \left(\bar{X}_j^N (s), \bar{V}^N (s) \right) dW_j (s) \right|^2
$$

$$
+ 4 \left| \int_t^{t+\varepsilon} \sigma \left(\bar{X}_j^N (s), \bar{V}^N (s) \right) u_j^N (s) \, ds \right|^2
$$

$$
\leq 12 T L^2 \varepsilon \left(1 + \sup_{0 \leq s \leq T} \left| \bar{X}_j^N (s) \right|^2 + \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq s \leq T} \left| \bar{X}_j^N (s) \right|^2 \right)
$$

$$
+ 4 \left| \int_t^{t+\varepsilon} \sigma \left(\bar{X}_j^N (s), \bar{V}^N (s) \right) dW_j (s) \right|^2 + 4 L^2 \varepsilon \int_0^T \left| u_j^N (s) \right|^2 \, ds.
$$

Since τ is a bounded stopping time, the optional sampling theorem gives

$$
E \left| \int_\tau^{\tau+\varepsilon} \sigma \left(\bar{X}_j^N (s), \bar{V}^N (s) \right) dW_j (s) \right|^2 \leq L^2 \varepsilon,
$$

and so

$$
\frac{1}{N} \sum_{j=1}^N E \left[\left| \bar{X}_j^N (\tau + \varepsilon) - \bar{X}_j^N (\tau) \right|^2 \right]
$$

$$
\leq 24 (T+1) L^2 \varepsilon \left(1 + E \left[\frac{1}{N} \sum_{j=1}^N \left\| \bar{X}_j^N \right\|_\infty^2 \right] + E \left[\frac{1}{N} \sum_{j=1}^N \int_0^T \left| u_j^N (s) \right|^2 \, ds \right] \right).
$$

The second estimate in the lemma now follows (with a possibly larger choice of c). $\qquad \square$

4.1. Proof of Lemma 3.3

The following general lemma will be useful in proving the tightness of $\{\bar{\mathcal{J}}^N\}$. The proof is standard (see, *e.g.*, [12, Exercise 3.11.18]) and is therefore omitted.

Lemma 4.2. *Let $\{Z_k, k \in \mathbb{N}\}$ be a sequence of random variables taking values in a separable Banach space with norm $\| \cdot \|$. Suppose that for each $\varepsilon > 0$ we can write $Z_k = Z_k^\varepsilon + R_k^\varepsilon$ for each $k \in \mathbb{N}$, where $\{Z_k^\varepsilon, k \in \mathbb{N}\}$ is tight and $\sup_{k \geq 1} E \left[\| R_k^\varepsilon \| \right] \leq \varepsilon$. Then $\{Z_k\}$ is tight.*

To prove tightness for the controlled stochastic currents, we will make use of a collection of test functions $\{g_M, M < \infty\}$ defined as follows.

Definition 4.3. Let $\{g_M, M < \infty\}$ be a collection of functions in $\mathcal{C}_c^\infty(\mathbb{R}^d, \mathbb{R})$ that satisfy $0 \leq g_M(x) \leq 1$ for all $M < \infty$ and $x \in \mathbb{R}^d$, and have the following properties:

  (i) For each M, $g_M(x) = 1$ on $|x| \leq M$;
  (ii) For each M, $g_M(x) = 0$ on $|x| \geq M + 1$;
 (iii) For every $k \in \mathbb{N}$, there is a constant $B(k) < \infty$ such that $|D^\alpha g_M(x)| \leq B(k)$ for all $x \in \mathbb{R}^d$, all $M < \infty$, and all $|\alpha| \leq k$.

Note that if $\{g_M, M < \infty\}$ is a collection as in Definition 4.3 then for every $k \in \mathbb{N}$ there is a constant $L(k) < \infty$ such that

$$|D^\alpha g_M(x) - D^\alpha g_M(y)| \leq L(k)|x - y| \tag{4.2}$$

for all $x, y \in \mathbb{R}^d$, all $M < \infty$, and all $|\alpha| \leq k$. We will need the following property of the collection $\{g_M, M < \infty\}$. A proof of the lemma is given in the Appendix.

Lemma 4.4. *For any $s > 0$ there is a constant $K = K(s) < \infty$ such that for any $f \in H^s(\mathbb{R}^d, \mathbb{R}^d)$*

$$\sup_{M < \infty} \|g_M f\|_s \leq K \|f\|_s.$$

The following is a simple extension of the well-known compact embedding result for Sobolev spaces on $\mathbb{R}^d$ known as Rellich's Theorem (see [15, Theorem 9.22]). Although the proof is standard, we provide details in the Appendix. For $\mathbf{s} \in \mathcal{O}_d$, $F \in \mathbf{H}^{-\mathbf{s}}$, and open $U_0 \subset U$, we say $F = 0$ on U_0 if for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ with support in U_0, $\langle F, \varphi \rangle = 0$. The support of F is the complement of the union of all open sets in U on which $F = 0$.

Lemma 4.5. *Let $\mathbf{s} = (s_1, s_2)$ and $\mathbf{s}' = (s_1', s_2')$ in $\mathcal{O}_d$ be such that $s_1' < s_1$ and $s_2' < s_2$. Suppose $A \subset \mathbf{H}^{-\mathbf{s}'}$ is such that for some compact $K \subset U \times \mathbb{R}^d$, every $F \in A$ has support contained in K. Suppose also that $\sup_{F \in A} \|F\|_{-\mathbf{s}'} < \infty$. Then A is relatively compact in $\mathbf{H}^{-\mathbf{s}}$.*

Finally, the lemma below establishes the required tightness for the controlled currents.

Lemma 4.6. *Suppose Conditions 2.1 and 2.3 are satisfied. Let $\{g_M, M < \infty\}$ be the collection of functions in $\mathcal{C}_c^\infty(\mathbb{R}^d, \mathbb{R})$ as in Definition 4.3. For each $N \in \mathbb{N}$, $M < \infty$, and $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, define*

$$\bar{J}^{N,M}(\varphi) \doteq \bar{J}^N(g_M \varphi), \qquad \bar{J}_c^{N,M}(\varphi) \doteq \bar{J}^N(\varphi) - \bar{J}^{N,M}(\varphi).$$

Then, the collections $\{\varphi \mapsto \bar{J}^{N,M}(\varphi)\}$ and $\{\varphi \mapsto \bar{J}_c^{N,M}(\varphi)\}$ have pathwise realizations $\bar{\mathcal{J}}^{N,M}, \bar{\mathcal{J}}_c^{N,M}$ in $\mathbf{H}^{-\mathbf{s}}$ for all $\mathbf{s} \in \mathcal{O}_d$. Furthermore, if

$$\sup_{N \geq 1} E\left[\frac{1}{N} \sum_{j=1}^N \int_0^T \left| u_j^N(t) \right|^2 dt \right] < \infty,$$

then for all $\mathbf{s} \in \mathcal{O}_d$

$$\sup_{M<\infty} \sup_{N\geq 1} E\left[\left\|\bar{\mathcal{J}}^{N,M}\right\|_{-\mathbf{s}}\right] < \infty$$

and

$$\lim_{M\to\infty} \sup_{N\geq 1} E\left[\left\|\bar{\mathcal{J}}_c^{N,M}\right\|_{-\mathbf{s}}\right] = 0.$$

In particular, $\{\bar{\mathcal{J}}^N, N \in \mathbb{N}\}$ *is tight in* $\mathbf{H}^{-\mathbf{s}}$ *for all* $\mathbf{s} \in \mathcal{O}_d$.

Proof. Fix $\mathbf{s} = (s_1, s_2) \in \mathcal{O}_d$, and for each N and $1 \leq j \leq N$, let $C_{j,\mathbf{s}}^N$ be the square-integrable random variable from Lemma 3.1, so that $|\bar{J}^N(\varphi)| \leq C_{\mathbf{s}}^N \|\varphi\|_{\mathbf{s}}$ a.s. for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, where $C_{\mathbf{s}}^N = \frac{1}{N}\sum_{j=1}^N C_{j,\mathbf{s}}^N$. As a consequence of Lemma 4.4, for some constant $K = K(s_2) < \infty$ we have, for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ and $M < \infty$,

$$\begin{aligned}
\|g_M\varphi\|_{\mathbf{s}}^2 &= \int_U \|g_M\varphi(u,\cdot)\|_{s_2}^2 \, du \\
&\quad + \int_U \int_U \frac{\|g_M(\varphi(u,\cdot) - \varphi(v,\cdot))\|_{s_2}^2}{|u-v|^{1+2s_1}} \, du \, dv \leq K^2 \|\varphi\|_{\mathbf{s}}^2.
\end{aligned} \tag{4.3}$$

Hence,

$$\left|\bar{J}^{N,M}(\varphi)\right| \leq C_{\mathbf{s}}^N \|g_M\varphi\|_{\mathbf{s}} \leq K C_{\mathbf{s}}^N \|\varphi\|_{\mathbf{s}} \qquad \text{a.s.,}$$

and

$$\left|\bar{J}_c^{N,M}(\varphi)\right| = \left|\bar{J}^N((1-g_M)\varphi)\right| \leq C_{\mathbf{s}}^N \|(1-g_M)\varphi\|_{\mathbf{s}} \leq (1+K) C_{\mathbf{s}}^N \|\varphi\|_{\mathbf{s}} \qquad \text{a.s.}$$

From [13, Lemma 5] it then follows that, for every $M < \infty$, there are $\mathbf{H}^{-\mathbf{s}}$-valued random variables $\bar{\mathcal{J}}^{N,M}$ and $\bar{\mathcal{J}}_c^{N,M}$ such that, for every $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ and $M < \infty$,

$$\begin{aligned}
\left\langle \bar{\mathcal{J}}^{N,M}(\omega), \varphi \right\rangle &= \left[J^{N,M}(\varphi)\right](\omega) && \text{and} \\
\left\langle \bar{\mathcal{J}}_c^{N,M}(\omega), \varphi \right\rangle &= \left[J_c^{N,M}(\varphi)\right](\omega) && \text{a.e. } \omega \in \Omega.
\end{aligned}$$

Then, from Lemma 3.1,

$$\sup_{M<\infty} \sup_{N\geq 1} E\left[\left\|\bar{\mathcal{J}}^{N,M}\right\|_{-\mathbf{s}}^2\right] \leq K^2 \sup_{N\geq 1} E\left[\left(C_{\mathbf{s}}^N\right)^2\right] < \infty. \tag{4.4}$$

Let $\bar{J}_j^N$ be as in (3.4) and define the stopping times $\tau_j^{N,M} = \inf\{t > 0 : |\bar{X}_j^N(t)| \geq M\}$. Then,

$$\bar{J}_c^{N,M}(\varphi) = \frac{1}{N}\sum_{j=1}^N \bar{J}_j^N((1-g_M)\varphi) = \frac{1}{N}\sum_{j=1}^N 1_{\left\{\tau_j^{N,M} < T\right\}} \bar{J}_j^N((1-g_M)\varphi),$$

and by Lemma 3.1,

$$\left| \bar{J}_j^N ((1 - g_M)\varphi) \right| \leq C_{j,\mathbf{s}}^N \|(1 - g_M)\varphi\|_{\mathbf{s}} \leq (1 + K)C_{j,\mathbf{s}}^N \|\varphi\|_{\mathbf{s}}.$$

Thus,

$$\left| \bar{J}_c^{N,M}(\varphi) \right| \leq \left(\frac{1 + K}{N} \sum_{j=1}^{N} 1_{\{\tau_j^{N,M} < T\}} C_{j,\mathbf{s}}^N \right) \|\varphi\|_{\mathbf{s}} \doteq \tilde{C}_{\mathbf{s}}^N \|\varphi\|_{\mathbf{s}}. \tag{4.5}$$

Also, by the Cauchy-Schwarz inequality,

$$E\left[(\tilde{C}_{\mathbf{s}}^N)^2 \right] \leq \frac{(1 + K)^2}{N} \left(\sum_{j=1}^{N} P\left(\tau_j^{N,M} < T\right) \right) \left(\frac{1}{N} \sum_{j=1}^{N} E\left[(C_{j,\mathbf{s}}^N)^2 \right] \right).$$

By Lemma 4.1, Condition 2.3, and the assumption that

$$\sup_{N \in \mathbb{N}} E\left[\frac{1}{N} \sum_{j=1}^{N} \int_0^T \left| u_j^N(t) \right|^2 dt \right] < \infty,$$

there is a constant $\tilde{K} < \infty$ such that

$$\sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^{N} P\left(\tau_j^{N,M} < T\right) \leq \sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^{N} P\left(\left\| \bar{X}_j^N \right\|_\infty \geq M \right) \leq \frac{\tilde{K}}{M^2}.$$

Thus,

$$E\left[\left\| \bar{\mathcal{J}}_c^{N,M} \right\|_{-\mathbf{s}}^2 \right] \leq E\left[(\tilde{C}_{\mathbf{s}}^N)^2 \right] \leq \frac{\tilde{K}(1 + K)^2}{M^2} \sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^{N} E\left[(C_{j,\mathbf{s}}^N)^2 \right],$$

and therefore, from Lemma 3.1,

$$\lim_{M \to \infty} \sup_{N \geq 1} E\left[\left\| \bar{\mathcal{J}}_c^{N,M} \right\|_{-\mathbf{s}}^2 \right] = 0. \tag{4.6}$$

Note that (4.4) and (4.6) are satisfied for every $\mathbf{s} \in \mathcal{O}_d$. Now for an arbitrary $\mathbf{s} \in \mathcal{O}_d$ choose $\mathbf{s}' = (s_1', s_2') \in \mathcal{O}_d$ such that $s_1' < s_1$ and $s_2' < s_2$. Then applying (4.4) for $\mathbf{s}'$ and observing that $\{\bar{\mathcal{J}}^{N,M}, N \in \mathbb{N}\}$ are compactly supported on $[0, T] \times \{|x| \leq M + 1\} \subset U \times \mathbb{R}^d$, we see from Lemma 4.5 and Markov's inequality that for each fixed M $\{\bar{\mathcal{J}}^{N,M}, N \in \mathbb{N}\}$ is a tight collection of $\mathbf{H}^{-\mathbf{s}}$-valued random variables. Finally, observing that $\bar{\mathcal{J}}^N = \bar{\mathcal{J}}^{N,M} + \bar{\mathcal{J}}_c^{N,M}$ for each M and applying (4.6) and Lemma 4.2, we obtain that $\{\bar{\mathcal{J}}^N, N \in \mathbb{N}\}$ is tight in $\mathbf{H}^{-\mathbf{s}}$. $\qquad\square$

The following general lemma will be useful in proving tightness of $\{\bar{\mu}^N\}$.

Lemma 4.7. *Let (S, d_S) be a Polish space. If $\{\gamma_k, k \in \mathbb{N}\}$ is a tight sequence of $\mathcal{P}(S)$-valued random variables and for some $x_0 \in S$*

$$\sup_{k \in \mathbb{N}} E\left[\int_S d_S(x, x_0)^2 \, \gamma_k(dx)\right] < \infty, \tag{4.7}$$

then $\{\gamma_k\}$ is tight as a sequence of $\mathcal{P}_1(S)$-valued random variables.

Proof. Suppose that γ_k converges in distribution, along a subsequence, in $\mathcal{P}(S)$ to some γ, and denote the convergent subsequence once more as $\{\gamma_k\}$. From (4.7) it follows that each γ_k is in $\mathcal{P}_1(S)$ a.s. Furthermore, by lower semicontinuity of the map $\mu \mapsto \int_S d_S(x, x_0)^2 \, \mu(dx)$ on $\mathcal{P}(S)$ and Fatou's lemma, we see that

$$E\left[\int_S d_S(x, x_0)^2 \, \gamma(dx)\right] \leq E\left[\liminf_{k \to \infty} \int_S d_S(x, x_0)^2 \, \gamma_k(dx)\right]$$

$$\leq \sup_{k \geq 1} E\left[\int_S d_S(x, x_0)^2 \, \gamma_k(dx)\right] < \infty,$$

and so in particular $\gamma \in \mathcal{P}_1(S)$ a.s. By appealing to Skorohod's representation theorem we can assume that $\gamma_k \to \gamma$ a.s. in $\mathcal{P}(S)$. Recalling from Subsection 1.2 the metric d_{bl} on the space $\mathcal{P}(S)$, we have that $d_{\mathrm{bl}}(\gamma_k, \gamma) \to 0$ a.s.

It suffices now to show that γ_k converges in probability in $\mathcal{P}_1(S)$ to γ. Take $f \in \mathcal{L}(S)$ such that $f(x_0) = 0$. Fix $1 < M < \infty$ and define

$$f_M(x) \doteq \left(\frac{f(x)}{M} \vee (-1)\right) \wedge 1,$$

which is a function bounded by 1 in absolute value whose Lipschitz constant is also bounded by 1. Then,

$$\left|\int_S f(x) \, \gamma_k(dx) - \int_S f(x) \, \gamma(dx)\right|$$

$$\leq M\left|\int_S f_M(x) \, \gamma_k(dx) - \int_S f_M(x) \, \gamma(dx)\right| + \int_S |Mf_M(x) - f(x)| \, \gamma_k(dx)$$

$$+ \int_S |Mf_M(x) - f(x)| \, \gamma(dx)$$

$$\leq M d_{\mathrm{bl}}(\gamma_k, \gamma) + \int_S 2|f(x)| 1_{\{|f(x)| > M\}} \, \gamma_k(dx) + \int_S 2|f(x)| 1_{\{|f(x)| > M\}} \, \gamma(dx).$$

Since the Lipschitz constant of f is bounded by 1 and $f(x_0) = 0$, we have that $|f(x)| \leq d_S(x, x_0)$, and so

$$\int_S |f(x)| 1_{\{|f(x)| > M\}} \, \gamma_k(dx) \leq \frac{1}{M} \int_S d_S(x, x_0)^2 \, \gamma_k(dx),$$

and the equivalent inequality holds for γ. Now, since $\langle \mu, f \rangle - \langle \nu, f \rangle = \langle \mu, f - f(x_0) \rangle - \langle \nu, f - f(x_0) \rangle$ for any $\mu, \nu \in \mathcal{P}_1(S)$ and $f \in \mathcal{L}(S)$, the supremum in the definition of d_1 can be restricted to f such that $f(x_0) = 0$. Thus,

$$
\begin{aligned}
E\left[d_1(\gamma_k, \gamma)\right] &= E\left[\sup_{f \in \mathcal{L}(S), f(x_0)=0} \left| \int_S f(x)\,\gamma_k(dx) - \int_S f(x)\,\gamma(dx) \right| \right] \\
&\leq M E\left[d_{\mathrm{bl}}(\gamma_k, \gamma)\right] + \frac{2}{M} \sup_{l \in \mathbb{N}} E\left[\int_S d_S(x, x_0)^2\,\gamma_l(dx)\right] \\
&\quad + \frac{2}{M} E\left[\int_S d_S(x, x_0)^2\,\gamma(dx)\right].
\end{aligned}
$$

Sending first $k \to \infty$ and then $M \to \infty$, we have that $\lim_{k \to \infty} E\left[d_1(\gamma_k, \gamma)\right] = 0$, which completes the proof. $\qquad\square$

We can now complete the proof of Lemma 3.3.

Proof of Lemma 3.3(i). We begin by arguing that $\{\bar{\mu}^N\}$ is a tight sequence of $\mathcal{P}(\mathcal{X})$-valued random variables. For this it suffices to show (see [7, Theorem 2.11]) that $\{\gamma^N, N \in \mathbb{N}\}$ is a relatively compact set in $\mathcal{P}(\mathcal{X})$, where

$$
\gamma^N \doteq E\left[\bar{\mu}^N\right] = \frac{1}{N} \sum_{j=1}^N P\left(\bar{X}_j^N \in \cdot\right).
$$

Note that

$$
\int_{\mathcal{X}} \|\psi\|_\infty^2\,\gamma^N(d\psi) = \frac{1}{N} \sum_{j=1}^N E\left[\left\|\bar{X}_j^N\right\|_\infty^2\right],
$$

and so by Lemma 4.1 and the assumption on the controls in Lemma 3.3, we see that

$$
\sup_{N \geq 1} \int_{\mathcal{X}} \|\psi\|_\infty^2\,\gamma^N(d\psi) = \sup_{N \geq 1} \frac{1}{N} \sum_{j=1}^N E\left[\left\|\bar{X}_j^N\right\|_\infty^2\right] < \infty. \tag{4.8}
$$

Next, for $\varepsilon > 0$ let $\mathcal{T}_\varepsilon$ denote the collection of all $\{\sigma(X(s) : s \leq t)\}$-stopping times on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ taking values in $[0, T - \varepsilon]$ where $\{X(t)\}$ is the canonical coordinate process on $\mathcal{X}$. Then for each $N \in \mathbb{N}$ there are $\{\sigma(\bar{X}_j^N(s) : s \leq t)\}$-stopping times $\{\tau_j^N, 1 \leq j \leq N\}$ on $(\Omega, \mathcal{F})$ with values in $[0, T - \varepsilon]$ such that

$$
\int_{\mathcal{X}} |\psi(\tau + \varepsilon) - \psi(\tau)|^2\,\gamma^N(d\psi) = \frac{1}{N} \sum_{j=1}^N E\left[\left|\bar{X}_j^N\left(\tau_j^N + \varepsilon\right) - \bar{X}_j^N\left(\tau_j^N\right)\right|^2\right].
$$

Applying Lemma 4.1, we then have

$$\int_{\mathcal{X}} |\psi(\tau+\varepsilon) - \psi(\tau)|^2 \, \gamma^N(d\psi)$$

$$\leq c\varepsilon \left(1 + \sup_{N\geq 1} \frac{1}{N} \sum_{j=1}^N \left| x_j^N \right|^2 + \sup_{N\geq 1} E\left[\frac{1}{N} \sum_{J=1}^N \int_0^T \left| u_j^N(t) \right|^2 \, dt \right] \right),$$

and hence

$$\lim_{\varepsilon\to 0} \sup_{N\geq 1} \sup_{\tau\in\mathcal{T}_\varepsilon} \int_{\mathcal{X}} |\psi(\tau+\varepsilon) - \psi(\tau)|^2 \, \gamma^N(d\psi) = 0. \tag{4.9}$$

The relative compactness of $\{\gamma^N, N \in \mathbb{N}\}$ in $\mathcal{P}(\mathcal{X})$ is immediate from (4.8) and (4.9) (see [7, Theorem D.4]), which as noted previously shows that $\{\bar{\mu}^N\}$ is a tight sequence of $\mathcal{P}(\mathcal{X})$-valued random variables. The tightness of $\{\bar{\mu}^N\}$ as a sequence of $\mathcal{P}_1(\mathcal{X})$-valued random variables now follows from Lemma 4.7 and the uniform moment estimate in (4.8). Note also that since $\bar{\mu}^N = Q_{(1)}^N$ we have the tightness of the first marginals of $\{Q^N\}$ (as a sequence of $\mathcal{P}(\mathcal{X})$-valued random variables).

That the second marginals $\{Q_{(2)}^N\}$ is a tight sequence of $\mathcal{P}(\mathcal{R}_1)$-valued random variables follows by an argument similar to [8, Lemma 5.1], however we provide the details. Note that the function

$$h(r) = \int_{[0,T]\times\mathbb{R}^m} |y|^2 \, r(dt, dy)$$

has compact level sets on $\mathcal{R}_1$ (recall that $\mathcal{R}_1$ is equipped with the Wasserstein-1 metric). It then follows that

$$H(\theta) = \int_{\mathcal{R}_1} h(r) \, \theta(dr)$$

has relatively compact level sets on $\mathcal{P}(\mathcal{R}_1)$ (see [7, Lemma 2.10]). It now suffices to show that $\sup_{N\geq 1} E[H(Q_{(2)}^N)] < \infty$ (see [7, Lemmas 2.9]). However this is immediate as

$$\sup_{N\geq 1} E\left[H\left(Q_{(2)}^N \right) \right] = \sup_{N\geq 1} E\left[\frac{1}{N} \sum_{j=1}^N \int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho_j^N(dt, dy) \right]$$

$$= \sup_{N\geq 1} E\left[\frac{1}{N} \sum_{j=1}^N \int_0^T \left| u_j^N(t) \right|^2 \, dt \right] < \infty. \tag{4.10}$$

Thus we have shown that the second marginals of $\{Q^N\}$ are also tight, which in turn shows that $\{\bar{\mu}^N, Q^N\}$ is a tight sequence of $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z})$-valued random variables. Together with Lemma 4.6, this finishes the proof of Lemma 3.3(i). $\qquad\square$

Proof of Lemma 3.3(ii). Suppose now that $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N) \Rightarrow (\bar{\mu}, Q, \bar{\mathcal{J}})$ in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-s}$, where $(\bar{\mu}, Q, \bar{\mathcal{J}})$ is defined on some probability space. By appealing to Skorokhod's representation theorem, we can assume that $\{(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N)\}$ and $(\bar{\mu}, Q, \bar{\mathcal{J}})$ are defined on a common probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and that $(\bar{\mu}^N, Q^N, \bar{\mathcal{J}}^N) \to (\bar{\mu}, Q, \bar{\mathcal{J}})$ a.s. Let $\tilde{E}$ denote expectation on this space. The property $Q_{(1)} = \bar{\mu}$ is immediate from the identity $Q_{(1)}^N = \bar{\mu}^N$ for every $N \in \mathbb{N}$. We will complete the remainder of the proof in three steps: step 1 will establish that $Q \in \mathcal{P}_2(\mathcal{Z})$, step 2 that $Q \in \mathcal{S}(\mathcal{Z})$, and step 3 that $Q \in \mathcal{P}^*(\bar{\mathcal{J}})$, from which the result will follow.

Step 1. By Fatou's lemma,

$$
\begin{aligned}
&\tilde{E}\left[E_Q\left[\int_{\mathbb{R}^m \times [0,T]} |y|^2 \rho(dy\, dt)\right]\right] \\
&\leq \liminf_{N\to\infty} \tilde{E}\left[E_{Q^N}\left[\int_{\mathbb{R}^m \times [0,T]} |y|^2 \rho(dy\, dt)\right]\right] \\
&= \liminf_{N\to\infty} E\left[\frac{1}{N}\sum_{j=1}^N \int_0^T \left|u_j^N(t)\right|^2 dt\right] < \infty,
\end{aligned}
\tag{4.11}
$$

and hence $Q \in \mathcal{P}_2(\mathcal{Z})$ a.s.

Step 2. We now show that a.s. $Q \in \mathcal{S}(\mathcal{Z})$, namely it is a weak solution to (2.9). Define the generator $\mathcal{A}$ as follows. For each $f \in \mathcal{C}_c^2(\mathbb{R}^d, \mathbb{R})$, let

$$
\mathcal{A}f(\nu, x, y) = (b(x, \nu) + \sigma(x, \nu)y) \cdot \nabla f(x), \quad (\nu, x, y) \in \mathcal{P}_1\left(\mathbb{R}^d\right) \times \mathbb{R}^d \times \mathbb{R}^m.
$$

Now fix an $f \in \mathcal{C}_c^2(\mathbb{R}^d, \mathbb{R})$ and define, for each $V \in \mathcal{V}$, the $\mathbb{R}$-valued process $\{M^V(t), 0 \leq t \leq T\}$ on the measurable space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ by

$$
\begin{aligned}
M^V(t, (\xi, r)) &= f(\xi(t)) - f(\xi(0)) \\
&\quad - \int_{[0,t] \times \mathbb{R}^m} \mathcal{A}f(V(s), \xi(s), y)\, r(ds, dy), \quad (\xi, r) \in \mathcal{Z}.
\end{aligned}
\tag{4.12}
$$

Let $\bar{V} \doteq \nu_Q$. Since f is arbitrary, to establish that $Q \in \mathcal{S}(\mathcal{Z})$ a.s. it suffices to show that for each fixed $0 \leq t \leq T$ and a.e. $\omega \in \tilde{\Omega}$

$$
M^{\bar{V}(\omega)}(t, (\xi, r)) = 0, \qquad Q(\omega)\text{-a.e. } (\xi, r) \in \mathcal{Z}. \tag{4.13}
$$

We will supress ω from the notation for the remainder of the proof.

For each $1 \leq B < \infty$, let $\psi_B \in \mathcal{C}_c(\mathbb{R}^m, \mathbb{R}^m)$ be such that $\psi_B(y) = y$ on $\{|y| \leq B\}$ and $|\psi_B(y)| \leq |y| + 1$ everywhere. Note that since $B \geq 1$, this definition implies that

$$
|y - \psi_B(y)| \leq \frac{|y|(2|y| + 1)}{B} 1_{\{|y| > B\}} \leq \frac{3|y|^2}{B}. \tag{4.14}
$$

Also let $\eta_B \in \mathcal{C}_c(\mathbb{R}^d, \mathbb{R}^d)$ be such that $\eta_B(x) = x$ on $\{|x| \le B\}$ and $|\eta_B(x)| \le |x| + 1$ everywhere. As with ψ_B, we have that

$$|x - \eta_B(x)| \le \frac{3|x|^2}{B}. \tag{4.15}$$

Now define the "truncated generator" $\mathcal{A}_B$

$$\mathcal{A}_B f(v, x, y) = (\eta_B(b(x, v)) + \sigma(x, v)\psi_B(y)) \cdot \nabla f(x),$$
$$(v, x, y) \in \mathcal{P}_1\left(\mathbb{R}^d\right) \times \mathbb{R}^d \times \mathbb{R}^m,$$

and for each $V \in \mathcal{V}$ let $\{M_B^V(t)\}$ be the corresponding process defined as in (4.12) with $\mathcal{A}_B$ in place of $\mathcal{A}$. Let

$$K \doteq \sup_{x \in \mathbb{R}^d} \left(|f(x)| + |\nabla f(x)| + |D^2 f(x)|\right) < \infty,$$

and note that for all $V \in \mathcal{V}$, $0 \le s \le t$, and $(x, y) \in \mathbb{R}^d \times \mathbb{R}^m$

$$
\begin{aligned}
&|\mathcal{A} f(V(s), x, y) - \mathcal{A}_B f(V(s), x, y)| \\
&\le K \left(\frac{3\,|b(x, V(s))|^2}{B} + \frac{3L|y|^2}{B}\right) \\
&\le \frac{12K(L+1)^2}{B} \left(1 + |x|^2 + \int_{\mathbb{R}^d} |x'|^2\, V(s, dx') + |y|^2\right).
\end{aligned}
\tag{4.16}
$$

Now fix t, and define the maps Φ and Φ_B on $\mathcal{P}(\mathcal{Z}) \times \mathcal{V}$ by

$$\Phi(\Theta, V) = E_\Theta\left[\left|M^V(t)\right|\right], \qquad \Phi_B(\Theta, V) = E_\Theta\left[\left|M_B^V(t)\right|\right].$$

Note that $\bar{V}^N = \nu_{Q^N}$, were $\bar{V}^N$ is as in Subsection 3.1. We proceed by showing that:

(a) Φ_B is bounded and continuous on $\mathcal{P}(\mathcal{Z}) \times \mathcal{V}$;

(b) $\sup_{N \ge 1} \tilde{E}\left[\left|\Phi(Q^N, \bar{V}^N) - \Phi_B(Q^N, \bar{V}^N)\right|\right] \to 0$ and $\left|\Phi(Q, \bar{V}) - \Phi_B(Q, \bar{V})\right| \xrightarrow{\tilde{P}} 0$ as $B \to \infty$;

(c) $\Phi(Q^N, \bar{V}^N) \xrightarrow{\tilde{P}} 0$ as $N \to \infty$.

The convergence $(Q^N, \bar{V}^N) \to (Q, \bar{V})$ then yields that $\Phi(Q, \bar{V}) = 0$ a.s., from which the statement in (4.13) is immediate.

   We first show (a). Boundedness of Φ_B follows from the boundedness of η_B, ψ_B, σ, f, and ∇f. The continuity of Φ_B follows from the continuity of the map $(V, z) \mapsto M_B^V(t, z)$ on $\mathcal{V} \times \mathcal{Z}$.

For (b), note from (4.16) that

$$
\tilde{E}\left[\left|\Phi\left(Q^N, \bar{V}^N\right) - \Phi_B\left(Q^N, \bar{V}^N\right)\right|\right] \le \tilde{E}\left[E_{Q^N}\left[\left|M^{\bar{V}^N}(t) - M_B^{\bar{V}^N}(t)\right|\right]\right]
$$

$$
\le \frac{12K(L+1)^2}{B} \tilde{E}\left[E_{Q^N}\left[\int_0^T \left(1 + |X(s)|^2 + \int_{\mathbb{R}^d} |x|^2\, \bar{V}^N(s, dx)\right.\right.\right.
$$

$$
\left.\left.\left. + \int_{\mathbb{R}^m} |y|^2\, \rho_s(dy)\right) ds\right]\right] \tag{4.17}
$$

$$
\le \frac{12K(L+1)^2}{B} \sup_{N \ge 1} E\left[T + \frac{2T}{N}\sum_{j=1}^N \left\|\bar{X}_j^N\right\|_\infty^2 + \frac{1}{N}\sum_{j=1}^N \int_0^T \left|u_j^N(s)\right|^2 ds\right].
$$

From Lemma 4.1 and the assumption on the controls in Lemma 3.3, we see that the last term in the above chain of inequalities converges to 0 as $B \to \infty$. Similarly, since $Q \in \mathcal{P}_2(\mathcal{Z})$ a.s., the estimate

$$
\left|\Phi(Q, \bar{V}) - \Phi_B(Q, \bar{V})\right| \le E_Q\left[\left|M^{\bar{V}}(t) - M_B^{\bar{V}}(t)\right|\right]
$$

$$
\le \frac{12K(L+1)^2}{B}\left(\int_0^T\left(1 + 2\int_{\mathbb{R}^d} |x|^2\, \bar{V}(s, dx)\right)ds + E_Q\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2\, \rho(ds, dy)\right]\right)
$$

implies that

$$
\left|\Phi\left(Q, \bar{V}\right) - \Phi_B\left(Q, \bar{V}\right)\right| \to 0 \qquad \text{a.s.} \qquad \text{as } B \to \infty. \tag{4.18}
$$

This completes the proof of (b).

We now turn to (c). Note that

$$
\Phi\left(Q^N, \bar{V}^N\right) = E_{Q^N}\left[\left|M^{\bar{V}^N}(t)\right|\right] = \frac{1}{N}\sum_{j=1}^N \left|M^{\bar{V}^N}\left(t, (\bar{X}_j^N, \rho_j^N)\right)\right|
$$

$$
= \frac{1}{N}\sum_{j=1}^N \left|f\left(\bar{X}_j^N(t)\right) - f\left(x_j^N\right) - \int_0^t \mathcal{A}f\left(\bar{V}^N(s), \bar{X}_j^N(s), u_j^N(s)\right) ds\right|.
$$

By Itô's lemma, for each $1 \le j \le N$,

$$
f\left(\bar{X}_j^N(t)\right) - f\left(x_j^N\right) = \int_0^t \mathcal{A}f\left(\bar{V}^N, \bar{X}_j^N(s), u_j^N(s)\right) ds
$$

$$
+ \varepsilon_N \int_0^t \nabla f\left(\bar{X}_j^N(s)\right) \cdot \sigma\left(\bar{X}_j^N(s), \bar{V}^N(s)\right) dW_j(s)
$$

$$
+ \frac{\varepsilon_N^2}{2}\int_0^t \text{Tr}\left[D^2 f\left(\bar{X}_j^N(s)\right)\left(\sigma\sigma^\mathsf{T}\right)\left(\bar{X}_j^N(s), \bar{V}^N(s)\right)\right] ds.
$$

Hence,

$$\Phi\left(Q^N, \bar{V}^N\right) = \frac{1}{N} \sum_{j=1}^{N} \left| \varepsilon_N \int_0^t \nabla f\left(\bar{X}_j^N(s)\right) \cdot \sigma\left(\bar{X}_j^N(s), \bar{V}^N(s)\right) dW_j(s) \right.$$
$$\left. + \frac{\varepsilon_N^2}{2} \int_0^t \mathrm{Tr}\left[D^2 f\left(\bar{X}_j^N(s)\right)\left(\sigma\sigma^\mathsf{T}\right)\left(\bar{X}_j^N(s), \bar{V}^N(s)\right)\right] ds \right|.$$

From the boundedness of ∇f, $D^2 f$, and σ, it follows that

$$\tilde{E}\left[\Phi\left(Q^N, \bar{V}^N\right)\right] \le \left(KLT^{1/2}\right)\varepsilon_N + \frac{KL^2 T \varepsilon_N^2}{2} \to 0 \qquad \text{as } N \to \infty.$$

This completes (c), which as noted previously proves the statement in (4.13) which in turn shows that Q is a.s. a weak solution to (2.9).

Step 3. To complete the proof of Lemma 3.3, it only remains to establish that

$$G_\varphi(Q) = \langle \bar{\mathcal{J}}, \varphi \rangle \qquad \text{for all} \qquad \varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d), \qquad \tilde{P}\text{-a.s.} \quad (4.19)$$

By considering a countable, dense subset of $\mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$, it suffices to show that for each fixed $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ we have $G_\varphi(Q) = \bar{\mathcal{J}}(\varphi)$ a.s.

Fix φ, and let

$$K_\varphi \doteq \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \left(|\varphi(t,x)| + \sum_{k,l=1}^{d} \left| \frac{\partial \varphi_k}{\partial x_l}(t,x) \right| \right) < \infty.$$

Then, a.s.,

$$\langle \bar{\mathcal{J}}^N, \varphi \rangle = \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \circ d\bar{X}_j^N(t)$$
$$= \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \cdot d\bar{X}_j^N(t) + \frac{1}{2N} \sum_{j=1}^{N} \langle \varphi\left(\cdot, \bar{X}_j^N(\cdot)\right), \bar{X}_j^N(\cdot) \rangle_T$$
$$= \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \cdot d\bar{X}_j^N(t)$$
$$+ \frac{\varepsilon_N^2}{2N} \sum_{j=1}^{N} \int_0^T \sum_{k,l=1}^{d} \frac{\partial \varphi_k}{\partial x_l}\left(t, \bar{X}_j^N(t)\right) (\sigma\sigma^\mathsf{T})_{lk}\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) dt.$$

Define

$$G_\varphi^*\left(Q^N\right) \doteq \frac{1}{N} \sum_{j=1}^{N} \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \cdot d\bar{X}_j^N(t).$$

Since $|\sigma| \le L$,

$$\left| \langle \bar{\mathcal{J}}^N, \varphi \rangle - G_\varphi^*(Q^N) \right|$$

$$= \left| \frac{\varepsilon_N^2}{2N} \sum_{j=1}^N \int_0^T \sum_{k,l=1}^d \frac{\partial \varphi_k}{\partial x_l} \left(t, \bar{X}_j^N(t) \right) (\sigma\sigma^\mathsf{T})_{lk} \left(\bar{X}_j^N(t), \bar{V}^N(t) \right) dt \right| \le \frac{K_\varphi L^2 T \varepsilon_N^2}{2},$$

and hence $|\langle \bar{\mathcal{J}}^N, \varphi \rangle - G_\varphi^*(Q^N)| \to 0$ in L^1 as $N \to \infty$. Also, by the dominated convergence theorem,

$$\lim_{N\to\infty} E\left[\left| \langle \bar{\mathcal{J}}, \varphi \rangle - \langle \bar{\mathcal{J}}^N, \varphi \rangle \right| \wedge 1 \right] = 0.$$

Next, writing

$$\left| \langle \bar{\mathcal{J}}, \varphi \rangle - G_\varphi(Q) \right| \wedge 1 \le \left| \langle \bar{\mathcal{J}}, \varphi \rangle - \langle \bar{\mathcal{J}}^N, \varphi \rangle \right| \wedge 1 + \left| \langle \bar{\mathcal{J}}^N, \varphi \rangle - G_\varphi^*(Q^N) \right|$$
$$+ \left| G_\varphi^*(Q^N) - G_\varphi(Q) \right|,$$

we see that to prove (4.19) and thus to complete the proof it suffices to argue that the third term on the right-hand side converges to 0 in probability.

To this end, define the maps $\tilde{G}_\varphi$ and $\tilde{G}_\varphi^B$ on $\{\Theta \in \mathcal{P}_2(\mathcal{Z}) : \nu_\Theta \in \mathcal{V}\} \times \mathcal{V}$ by

$$\tilde{G}_\varphi(\Theta, V) \doteq E_\Theta \Bigg[\int_0^T \varphi(t, X(t)) \cdot b(X(t), V(t)) \, dt$$

$$+ \int_{[0,T]\times\mathbb{R}^m} \varphi(t, X(t)) \cdot \sigma(X(t), V(t)) y \, \rho(dt, dy) \Bigg],$$

$$\tilde{G}_\varphi^B(\Theta, V) \doteq E_\Theta \Bigg[\int_0^T \varphi(t, X(t)) \cdot \eta_B \left(b(X(t), V(t)) \right) dt \Bigg]$$

$$+ E_\Theta \Bigg[\int_{[0,T]\times\mathbb{R}^m} \varphi(t, X(t)) \cdot \sigma(X(t), V(t)) \psi_B(y) \, \rho(dt, dy) \Bigg]$$

for each $1 \le B < \infty$. Note by (2.10) that $\tilde{G}_\varphi(\Theta, \nu_\Theta) = G_\varphi(\Theta)$ whenever $\Theta \in \mathcal{S}(\mathcal{Z})$, and hence since $\bar{V} = \nu_Q$ and $Q \in \mathcal{S}(\mathcal{Z})$ a.s., we have that $\tilde{G}_\varphi(Q, \bar{V}) = G_\varphi(Q)$ a.s. Also, since

$$\tilde{G}_\varphi\left(Q^N, \bar{V}^N\right) = \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \cdot b\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) dt$$

$$+ \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \cdot \sigma\left(\bar{X}_j^N(t), \bar{V}^N(t)\right) u_j^N(t) \, dt$$

and $\varepsilon_N \to 0$, we see that $|\tilde{G}_\varphi(Q^N, \bar{V}^N) - G_\varphi^*(Q^N)| \xrightarrow{\tilde{P}} 0$ as $N \to \infty$. Thus it remains to argue that

$$\left| \tilde{G}_\varphi \left(Q^N, \bar{V}^N \right) - \tilde{G}_\varphi \left(Q, \bar{V} \right) \right| \xrightarrow{\tilde{P}} 0 \qquad \text{as } N \to \infty. \tag{4.20}$$

Now, since

$$\tilde{G}_\varphi^B \left(Q^N, \bar{V}^N \right) = \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi \left(t, \bar{X}_j^N(t) \right) \cdot \eta_B \left(b \left(\bar{X}_j^N(t), \bar{V}^N(t) \right) \right) dt$$

$$+ \frac{1}{N} \sum_{j=1}^N \int_0^T \varphi \left(t, \bar{X}_j^N(t) \right) \cdot \sigma \left(\bar{X}_j^N(t), \bar{V}^N(t) \right) \psi_B \left(u_j^N(t) \right) dt,$$

and the map

$$(\xi, r, V) \mapsto \int_0^T \varphi \left(t, \xi(t) \right) \cdot \eta_B (b(\xi(t), V(t))) \, dt$$

$$+ \int_{[0,T] \times \mathbb{R}^m} \varphi \left(t, \xi(t) \right) \cdot \sigma(\xi(t), V(t)) \psi_B(y) \, r(dt, dy)$$

is bounded and continuous on $\mathcal{Z} \times \mathcal{V}$, the a.s. convergence $(Q^N, \bar{V}^N) \to (Q, \bar{V})$ in $\mathcal{P}(\mathcal{Z}) \times \mathcal{V}$ implies that

$$\tilde{G}_\varphi^B \left(Q^N, \bar{V}^N \right) \to \tilde{G}_\varphi^B \left(Q, \bar{V} \right) \qquad \text{a.s.} \qquad \text{as } N \to \infty \tag{4.21}$$

for each B. Also, using (4.14) and (4.15), as in the proof of (4.17), we see

$$\left| \tilde{G}_\varphi^B \left(Q^N, \bar{V}^N \right) - \tilde{G}_\varphi \left(Q^N, \bar{V}^N \right) \right|$$

$$\leq \frac{18 K_\varphi L^2 T}{B} \left(1 + \frac{1}{N} \sum_{j=1}^N \left\| \bar{X}_j^N \right\|_\infty^2 \right) + \frac{3 K_\varphi L}{BN} \sum_{j=1}^N \int_0^T \left| u_j^N(t) \right|^2 dt,$$

which in view of Lemma 4.1 and the assumption on the controls in Lemma 3.3 shows that

$$\sup_{N \geq 1} \tilde{E} \left[\left| \tilde{G}_\varphi^B \left(Q^N, \bar{V}^N \right) - \tilde{G}_\varphi \left(Q^N, \bar{V}^N \right) \right| \right] \to 0 \qquad \text{as } B \to \infty. \tag{4.22}$$

Finally, along the same lines as in the proof of (4.18),

$$\left| \tilde{G}_\varphi^B \left(Q, \bar{V} \right) - \tilde{G}_\varphi \left(Q, \bar{V} \right) \right| \to 0 \qquad \text{a.s.} \qquad \text{as } B \to \infty.$$

Combining the above convergence with (4.21) and (4.22) shows (4.20), which as noted previously establishes that $Q \in \mathcal{P}^*(\bar{\mathcal{J}})$ a.s. and thus completes the proof of the lemma. □

4.2. Proof of Lemma 3.6

We first prove an estimate similar to that in Lemma 4.1 for the coordinate process $X(t)$ on the space $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}), \Theta)$ for each $\Theta \in \mathcal{P}_2(\mathcal{Z}) \cap \mathcal{S}(\mathcal{Z})$. By the definition of $\mathcal{S}(\mathcal{Z})$, the coordinate maps (X, ρ) satisfy

$$dX(t) = b\left(X(t), \nu_\Theta(t)\right) dt + \int_{\mathbb{R}^m} \sigma\left(X(t), \nu_\Theta(t)\right) y\, \rho_t(dy)\, dt \quad \Theta\text{-a.s.,} \quad (4.23)$$

with $X(0) \sim \mu_0$. By Condition 2.1,

$$
\begin{aligned}
|b\left(X(t), \nu_\Theta(t)\right)|^2 &\leq 3L^2 \left(1 + |X(t)|^2 + \int_{\mathbb{R}^d} |x|^2\, \nu_\Theta(t, dx)\right) \\
&= 3L^2 \left(1 + |X(t)|^2 + E_\Theta\left[|X(t)|^2\right]\right).
\end{aligned}
\quad (4.24)
$$

Applying the above bound in (4.23), taking expectation, using $|\sigma| \leq L$, and applying Gronwall's inequality, we have

$$
\begin{aligned}
&E_\Theta\left[\|X\|_\infty^2\right] \\
&\leq \tilde{c}\left(1 + \int_{\mathbb{R}^d} |x|^2 \mu_0(dx) + E_\Theta\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho(dt, dy)\right]\right) < \infty,
\end{aligned}
\quad (4.25)
$$

for some $\tilde{c} = \tilde{c}(L, T) < \infty$.

Now fix $\mathbf{s} \in \mathcal{O}_d$ and let $\{(\mu_k, \Theta_k, \mathcal{J}_k)\}$ be a sequence in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z}) \times \mathbf{H}^{-\mathbf{s}}$ that satisfies the hypotheses of the lemma. Note that, by (4.25),

$$
\begin{aligned}
&\sup_{k \geq 1} \int_{\mathcal{X}} \|\psi\|_\infty^2\, \mu_k(d\psi) \\
&= \sup_{k \geq 1} \int_{\mathcal{X}} \|\psi\|_\infty^2\, (\Theta_k)_{(1)}(d\psi) = \sup_{k \geq 1} E_{\Theta_k}\left[\|X\|_\infty^2\right] \\
&\leq \tilde{c}\left(1 + \int_{\mathbb{R}^d} |x|^2\, \mu_0(dx) + \sup_{k \geq 1} E_{\Theta_k}\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2\, \rho(dt, dy)\right]\right) < \infty.
\end{aligned}
\quad (4.26)
$$

If τ is a $\{\sigma(X(s), s \leq t)\}$-stopping time on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ taking values in $[0, T - \varepsilon]$, then for any $\varepsilon > 0$

$$
\begin{aligned}
&|X(\tau + \varepsilon) - X(\tau)|^2 \\
&\leq 2\left|\int_\tau^{\tau+\varepsilon} b(X(t), \nu_{\Theta_k}(t))\, dt\right|^2 + 2\left|\int_\tau^{\tau+\varepsilon} \int_{\mathbb{R}^m} \sigma(X(t), \nu_{\Theta_k}(t)) y\, \rho_t(dy)\, dt\right|^2 \\
&\leq 6L^2\varepsilon \int_0^T \left(1 + |X(t)|^2 + E_{\Theta_k}\left[|X(t)|^2\right]\right) dt + 2L^2\varepsilon \int_0^T \int_{\mathbb{R}^m} |y|^2\, \rho_t(dy)\, dt,
\end{aligned}
$$

Θ_k-a.s. for each k. Hence, using the bound in (4.25),

$$E_{\Theta_k} \left[|X(\tau + \varepsilon) - X(\tau)|^2 \right]$$

$$\leq 12 L^2 (1 + \tilde{c}) \varepsilon \left(1 + \int_{\mathbb{R}^d} |x|^2 \, \mu_0(dx) + \sup_{k \geq 1} E_{\Theta_k} \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right).$$

If $\mathcal{T}_\varepsilon$ denotes the collection of all such stopping times τ, it follows that

$$\sup_{k \geq 1} \sup_{\tau \in \mathcal{T}_\varepsilon} \int_{\mathcal{X}} |\psi(\tau + \varepsilon) - \psi(\tau)|^2 \, \mu_k(d\psi)$$

$$= \sup_{k \geq 1} \sup_{\tau \in \mathcal{T}_\varepsilon} \int_{\mathcal{X}} |\psi(\tau + \varepsilon) - \psi(\tau)|^2 \, (\Theta_k)_{(1)}(d\psi)$$

$$= \sup_{k \geq 1} \sup_{\tau \in \mathcal{T}_\varepsilon} E_{\Theta_k} \left[|X(\tau + \varepsilon) - X(\tau)|^2 \right]$$

$$\to 0$$

as $\varepsilon \to 0$. This and (4.26) prove relative compactness of $\{\mu_k\}$ (and hence of $\{(\Theta_k)_{(1)}\}$) in $\mathcal{P}(\mathcal{X})$. By Lemma 4.7 and (4.26), we in fact get relative compactness of $\{\mu_k\}$ in $\mathcal{P}_1(\mathcal{X})$.

For the second marginals $\{(\Theta_k)_{(2)}\}$, we recall from the proof of Lemma 3.3 that

$$H(\theta) = \int_{\mathcal{R}_1} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, r(dt, dy) \, \theta(dr)$$

has relatively compact level sets on $\mathcal{P}(\mathcal{R}_1)$. Hence, we have relative compactness of $\{(\Theta_k)_{(2)}\}$ in $\mathcal{P}(\mathcal{R}_1)$ on observing that

$$\sup_{k \geq 1} H\left((\Theta_k)_{(2)} \right) = \sup_{k \geq 1} E_{\Theta_k} \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \rho(dt, dy) \right] < \infty.$$

This establishes that $\{\Theta_k\}$ is relatively compact in $\mathcal{P}(\mathcal{Z})$.

For $\{\mathcal{J}_k\}$, we employ the following lemma, the proof of which is saved for the Appendix.

Lemma 4.8. *Suppose Conditions 2.1 and 2.3 are satisfied. Also suppose that, for some $\mathbf{s} \in \mathcal{O}_d$ and $(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$, $I(\mu, \mathcal{J}) < \infty$. Then, for each $\mathbf{s}' \in \mathcal{O}_d$, there is a constant $C_{\mathbf{s}'} < \infty$ such that for any $\Theta \in \mathcal{P}^*(\mathcal{J})$ with $\Theta_{(1)} = \mu$ and for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$*

$$|\langle \mathcal{J}, \varphi \rangle|^2 \leq E_\Theta \left[\left| \int_0^T \varphi(t, X(t)) \cdot dX(t) \right|^2 \right]$$

$$\leq C_{\mathbf{s}'} \left(1 + E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) \|\varphi\|_{\mathbf{s}'}^2,$$

where $C_{\mathbf{s}'}$ does not depend on $\mathcal{J}$, φ, or Θ. In particular, $\mathcal{J} \in \mathbf{H}^{-\mathbf{s}'}$ for all $\mathbf{s}' \in \mathcal{O}_d$.

Recall the collection of test functions $\{g_M, M < \infty\}$ from Definition 4.3, which by Lemma 4.4 (see (4.3)) satisfy

$$\|g_M \varphi\|_{\mathbf{s}} \leq K \|\varphi\|_{\mathbf{s}} \tag{4.27}$$

for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ and $\mathbf{s} \in \mathcal{O}_d$, with $K < \infty$ depending only on $\mathbf{s}$. For each $k \geq 1$ and $M < \infty$, define $\mathcal{J}_k^M, \mathcal{J}_k^{M,c} \in \mathbf{H}^{-\mathbf{s}}$ by

$$\left\langle \mathcal{J}_k^M, \varphi \right\rangle \doteq \left\langle \mathcal{J}_k, g_M \varphi \right\rangle, \quad \left\langle \mathcal{J}_k^{M,c}, \varphi \right\rangle \doteq \left\langle \mathcal{J}_k, \varphi \right\rangle - \left\langle \mathcal{J}_k^M, \varphi \right\rangle, \quad \varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d).$$

Fix some $\mathbf{s}' \in \mathcal{O}_d$ such that $s_1' < s_1$ and $s_2' < s_2$. Since $\Theta_k \in \mathcal{P}^*(\mathcal{J}_k)$ for each k and (3.15) holds, $I(\mu_k, \mathcal{J}_k) < \infty$ for each k, so, by Lemma 4.8, $\mathcal{J}_k \in \mathbf{H}^{-\mathbf{s}'}$ for each k. Then for each k and M, in view of (4.27), $\mathcal{J}_k^M$ and $\mathcal{J}_k^{M,c}$ are in $\mathbf{H}^{-\mathbf{s}'}$ as well, and furthermore

$$\left| \left\langle \mathcal{J}_k^M, \varphi \right\rangle \right|^2 \leq C_{\mathbf{s}'} \left(1 + E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) \|g_M \varphi\|_{\mathbf{s}'}^2$$

$$\leq C_{\mathbf{s}'} K^2 \left(1 + E_{\Theta_k} \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) \|\varphi\|_{\mathbf{s}'}^2,$$

and hence

$$\sup_{M < \infty, k \geq 1} \left\| \mathcal{J}_k^M \right\|_{-\mathbf{s}'}^2$$
$$\leq C_{\mathbf{s}'} K^2 \left(1 + \sup_{k \geq 1} E_{\Theta_k} \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) < \infty. \tag{4.28}$$

Noting that, for each M, $\{\mathcal{J}_k^M\}$ are all supported on $[0, T] \times \{|x| \leq M + 1\} \subset U \times \mathbb{R}^d$, by Lemma 4.5 $\{\mathcal{J}_k^M, k \geq 1\}$ is relatively compact in $\mathbf{H}^{-\mathbf{s}}$. Now define the collection of stopping times $\{\tau^M, M < \infty\}$ on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$ by $\tau^M \doteq \inf\{t > 0 : |X(t)| \geq M\}$. Note that

$$\left\langle \mathcal{J}_k^{M,c}, \varphi \right\rangle = E_{\Theta_k} \left[\int_0^T (1 - g_M(X(t))) \varphi(t, X(t)) \cdot dX(t) \right]$$

$$= E_{\Theta_k} \left[1_{\{\tau^M < T\}} \int_0^T (1 - g_M(X(t))) \varphi(t, X(t)) \cdot dX(t) \right],$$

and so by Lemma 4.8 and (4.27)

$$\left|\left\langle \mathcal{J}_k^{M,c}, \varphi \right\rangle\right|^2$$

$$\leq \Theta_k \left(\tau^M < T\right) E_{\Theta_k} \left[\left|\int_0^T (1 - g_M(X(t)))\varphi(t, X(t)) \cdot dX(t)\right|^2\right]$$

$$\leq \Theta_k \left(\tau^M < T\right) C_s \left(1 + E_{\Theta_k}\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho(dt, dy)\right]\right) \|(1 - g_M)\varphi\|_s^2$$

$$\leq 2\Theta_k \left(\tau^M < T\right) C_s \left(1 + K^2\right) \left(1 + E_{\Theta_k}\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho(dt, dy)\right]\right) \|\varphi\|_s^2,$$

and hence

$$\sup_{k \geq 1} \left\|\mathcal{J}_k^{M,c}\right\|_{-s}^2$$

$$\leq 2 \sup_{k \geq 1} \Theta_k \left(\tau^M < T\right) C_s \left(1 + K^2\right)$$

$$\times \left(1 + \sup_{k \geq 1} E_{\Theta_k}\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho(dt, dy)\right]\right) \tag{4.29}$$

$$\leq \frac{2}{M^2} \sup_{k \geq 1} E_{\Theta_k}\left[\|X\|_\infty^2\right] C_s \left(1 + K^2\right)$$

$$\times \left(1 + \sup_{k \geq 1} E_{\Theta_k}\left[\int_{[0,T]\times\mathbb{R}^m} |y|^2 \rho(dt, dy)\right]\right)$$

$$\to 0$$

as $M \to \infty$, by (4.25). Then by Lemma 4.2 (applied to the constant random variables $\mathcal{J}_k = \mathcal{J}_k^M + \mathcal{J}_k^{M,c}$ on $(\mathcal{Z}, \mathcal{B}(\mathcal{Z}))$), we obtain from (4.28) and (4.29) that $\{\mathcal{J}_k\}$ is relatively compact in $\mathbf{H}^{-s}$. Lemma 3.6 now follows on combining the above with the relative compactness of $\{(\mu_k, \Theta_k)\}$ in $\mathcal{P}_1(\mathcal{X}) \times \mathcal{P}(\mathcal{Z})$ shown previously. □

4.3. Proof of Lemma 3.5

Recall that we assume that Conditions 2.1, 2.3 and 2.4 hold. In particular, $\sigma(x, \mu) = \sigma(\mu)$. Let $\Theta_1, \Theta_2 \in \mathcal{S}(\mathcal{Z}) \cap \mathcal{P}_2(\mathcal{Z})$ be such that $\Theta_1 \circ \theta^{-1} = \Theta_2 \circ \theta^{-1}$, and let $\Lambda = \Theta_1 \circ \theta^{-1}$. Then for $j \in \{1, 2\}$, we can disintegrate Θ_j as

$$\Theta_j(dx, dr) = \widetilde{\Theta}_j(x_0, r, dx) \, \Lambda(dx_0, dr)$$

for some measurable map $\widetilde{\Theta}_j : \mathbb{R}^d \times \mathcal{R}_1 \to \mathcal{P}(\mathcal{X})$. Define the probability measure Ξ on the space $\mathbb{R}^d \times \mathcal{R}_1 \times \mathcal{X} \times \mathcal{X}$ as

$$\Xi(dx_0, dr, dx_1, dx_2) = \widetilde{\Theta}_1(x_0, r, dx_1) \, \widetilde{\Theta}_2(x_0, r, dx_2) \, \Lambda(dx_0, dr),$$

and let (ξ_0, ρ, X_1, X_2) denote the coordinate maps on this space. Then, $X_1(0) = X_2(0) = \xi_0$, and to prove the lemma it suffices to show that $X_1 = X_2$ Ξ-a.s.

Letting $u(t) = \int_{\mathbb{R}^m} y \, \rho_t(dy)$ and $V_j(t) = \Xi \circ (X_j(t))^{-1}$, we have that $E_\Xi[\int_0^T |u(t)|^2 \, dt] < \infty$ and

$$X_j(t) = \xi_0 + \int_0^t b\left(X_j(s), V_j(s)\right) ds + \int_0^t \sigma\left(V_j(s)\right) u(s) \, ds, \qquad j \in \{1, 2\}.$$

By the Lipschitz property of the coefficients and the fact that

$$d_1\left(V_1(s), V_2(s)\right)^2 \le \left(E_\Xi\left[\|X_1(s) - X_2(s)\|\right]\right)^2 \le E_\Xi\left[\sup_{0 \le r \le s} |X_1(r) - X_2(r)|^2\right],$$

it follows from Condition 2.1 that for every $0 \le t \le T$,

$$
\begin{aligned}
&|X_1(t) - X_2(t)|^2 \\
&\le 2T \int_0^t |b\left(X_1(s), V_1(s)\right) - b\left(X_2(s), V_2(s)\right)|^2 \, ds \\
&\quad + 2\left(\int_0^T |u(s)|^2 \, ds\right) \int_0^T |\sigma\left(V_1(s)\right) - \sigma\left(V_2(s)\right)|^2 \, ds \\
&\le 2L^2 T \int_0^t \left(|X_1(s) - X_2(s)| + d_1\left(V_1(s), V_2(s)\right)\right)^2 \, ds \\
&\quad + 2L^2 \left(\int_0^t |u(s)|^2 \, ds\right) \int_0^T d_1\left(V_1(s), V_2(s)\right)^2 \, ds \\
&\le 4L^2 T \int_0^t \sup_{0 \le r \le s} |X_1(r) - X_2(r)|^2 \, ds \\
&\quad + 2L^2 \left(2T + \int_0^T |u(t)|^2 \, dt\right) \int_0^t E_\Xi\left[\sup_{0 \le r \le s} |X_1(r) - X_2(r)|^2\right] ds.
\end{aligned}
$$

Then, taking expectation with respect to Ξ, for all $0 \le t \le T$

$$
\begin{aligned}
&E_\Xi\left[\sup_{0 \le s \le t} |X_1(s) - X_2(s)|^2\right] \\
&\le 2L^2 \left(4T + E_\Xi\left[\int_0^T |u(s)|^2 \, ds\right]\right) \int_0^t E_\Xi\left[\sup_{0 \le r \le s} |X_1(r) - X_2(r)|^2\right] ds.
\end{aligned}
$$

Gronwall's inequality now shows that $E_\Xi[\|X_1 - X_2\|_\infty^2] = 0$, which completes the proof.

Appendix A.

In this section we provide proofs of some Sobolev space results that are used in our work. It will be convenient to introduce an alternate norm on $\mathbf{H^s}$ equivalent to (2.8), and which is similar to norms used in [3] and [29]. Let $\{e_1, \ldots, e_d\}$ denote the canonical basis in $\mathbb{R}^d$, recall that $U = (a, b) \supset [0, T]$, let $\mathcal{I} \doteq \mathbb{Z} \times \mathbb{R}^d \times \{1, \ldots, d\}$, and define the functions $e_{n,\xi}^k : U \times \mathbb{R}^d \to \mathbb{R}^d$ for $(n, \xi, k) \in \mathcal{I}$ by

$$e_{n,\xi}^k(t, x) = \frac{1}{b-a} e^{2\pi i n t/(b-a)} e^{2\pi i \xi \cdot x} e_k.$$

Consider the Fourier coefficients of $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$ given by

$$\hat{\varphi}(n, \xi) = (\hat{\varphi}_1(n, \xi), \ldots, \hat{\varphi}_d(n, \xi)),$$
$$\hat{\varphi}_k(n, \xi) = \int_U \int_{\mathbb{R}^d} e_{-n,-\xi}^k(t, x) \cdot \varphi(t, x) \, dx \, dt. \tag{A.1}$$

Then an equivalent norm on $\mathbf{H^s}$, $\mathbf{s} = (s_1, s_2) \in \mathbb{R}_+^2$, is given by

$$\|\varphi\|_{*,\mathbf{s}}^2 = \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} |\hat{\varphi}(n, \xi)|^2 \left(1 + n^2\right)^{s_1} \left(1 + |\xi|^2\right)^{s_2} \, d\xi. \tag{A.2}$$

A.1. Proof of Lemma 3.1

From the equivalence of the norms, it suffices to prove the statement in the lemma with $\| \cdot \|_{\mathbf{s}}$ replaced with $\| \cdot \|_{*,\mathbf{s}}$. In what follows, we will abuse notation and denote $\| \cdot \|_{*,\mathbf{s}}$ once more as $\| \cdot \|_{\mathbf{s}}$. Recall that for $N \in \mathbb{N}$, $1 \le j \le N$, and $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$,

$$\bar{J}_j^N(\varphi) = \int_0^T \varphi\left(t, \bar{X}_j^N(t)\right) \circ d\bar{X}_j^N(t).$$

Any such φ can be written in terms of its Fourier coefficients as

$$\varphi(t, x) = \sum_{k=1}^d \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \hat{\varphi}_k(n, \xi) e_{n,\xi}^k(t, x) \, d\xi.$$

As in [13, Lemma 8] it follows that

$$\bar{J}_j^N(\varphi) = \sum_{k=1}^d \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \hat{\varphi}_k(n, \xi) Z_{j,k}^N(n, \xi) \, d\xi,$$

where

$$Z_{j,k}^N(n, \xi) \doteq \int_0^T e_{n,\xi}^k\left(t, \bar{X}_j^N(t)\right) \circ d\bar{X}_j^N(t).$$

Note that

$$
\begin{aligned}
Z_{j,k}^N(n,\xi) &= \int_0^T e_{n,\xi}^k\left(t,\bar{X}_j^N(t)\right)\cdot d\bar{X}_j^N(t) + \frac{1}{2}\left\langle e_{n,\xi}^k\left(\cdot,\bar{X}_j^N(\cdot)\right),\bar{X}_j^N(\cdot)\right\rangle_T \\
&= \int_0^T e_{n,\xi}^k\left(t,\bar{X}_j^N(t)\right)\cdot b\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)\,dt \\
&\quad + \int_0^T e_{n,\xi}^k\left(t,\bar{X}_j^N(t)\right)\cdot \sigma\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)u_j^N(t)\,dt \\
&\quad + \varepsilon_N\int_0^T e_{n,\xi}^k\left(t,\bar{X}_j^N(t)\right)\cdot \sigma\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)\,dW_j(t) \\
&\quad + \pi i\,\varepsilon_N^2\xi_k\int_0^T \left(e_{n,\xi}^k\right)_k\left(t,\bar{X}_j^N(t)\right)\left(\sigma\sigma^{\mathsf{T}}\right)_{kk}\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)\,dt,
\end{aligned}
$$

since the kth component $(e_{n,\xi}^k)_k$ is the only nonzero component of $e_{n,\xi}^k$. By the Cauchy-Schwarz inequality, for all $\varphi\in\mathcal{C}_c^\infty(U\times\mathbb{R}^d,\mathbb{R}^d)$,

$$
\left|\bar{J}_j^N(\varphi)\right|^2 \le \|\varphi\|_{\mathbf{s}}^2\sum_{k=1}^d\sum_{n\in\mathbb{Z}}\int_{\mathbb{R}^d}\frac{\left|Z_{j,k}^N(n,\xi)\right|^2}{(1+n^2)^{s_1}(1+|\xi|^2)^{s_2}}\,d\xi = \|\varphi\|_{\mathbf{s}}^2\left(C_{j,\mathbf{s}}^N\right)^2, \quad \text{(A.3)}
$$

where

$$
C_{j,\mathbf{s}}^N \doteq \left(\sum_{k=1}^d\sum_{n\in\mathbb{Z}}\int_{\mathbb{R}^d}\frac{\left|Z_{j,k}^N(n,\xi)\right|^2}{(1+n^2)^{s_1}(1+|\xi|^2)^{s_2}}\,d\xi\right)^{1/2}.
$$

Since $|e_{n,\xi}^k|\le T^{-1}$ and $|\sigma|\le L$, the Burkholder-Davis-Gundy inequality gives

$$
\begin{aligned}
E\left[\left|Z_{j,k}^N(n,\xi)\right|^2\right] &\le 4E\left[\int_0^T\left|b\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)\right|^2\,dt\right] \\
&\quad + \frac{4L^2}{T}E\left[\int_0^T\left|u_j^N(t)\right|^2\,dt\right] + \frac{4\varepsilon_N^2L^2}{T} \\
&\quad + \frac{4\pi^2\varepsilon_N^4L^4\xi_k^2}{T}.
\end{aligned} \qquad \text{(A.4)}
$$

By the linear growth property of b from Condition 2.1,

$$
\left|b\left(\bar{X}_j^N(t),\bar{V}^N(t)\right)\right|^2 \le 3L^2\left(1+\left|\bar{X}_j^N(t)\right|^2 + \frac{1}{N}\sum_{l=1}^N\left|\bar{X}_l^N(t)\right|^2\right),
$$

and from Lemma 4.1 $E\left[\|\bar{X}_j^N\|_\infty^2\right] < \infty$ for each $N \in \mathbb{N}$ and $1 \leq j \leq N$. Using the last two estimates and (A.4), we see that

$$\sup_{(n,\xi,k)\in\mathcal{I}} E\left[\left|Z_{j,k}^N(n,\xi)\right|^2\right] < \infty.$$

Thus, for each $N \in \mathbb{N}$ and $1 \leq j \leq N$, $E[|C_{j,\mathbf{s}}^N|^2] < \infty$ for any $\mathbf{s} \in \mathcal{O}_d$. Following [13], we now have from (A.3) the existence of a pathwise realization $\bar{\mathcal{J}}^N$ of $\{\varphi \mapsto \bar{J}^N(\varphi)\}$ in $\mathbf{H}^{-\mathbf{s}}$ for every $N \in \mathbb{N}$ and any $\mathbf{s} \in \mathcal{O}_d$. This proves the first part of the lemma.

For the second part, note that by Lemma 4.1

$$E\left[\left|b\left(\bar{X}_j^N(t), \bar{V}^N(t)\right)\right|^2\right]$$

$$\leq 4L^2(c+1)\left(1 + \left|x_j^N\right|^2 + E\left[\int_0^T \left|u_j^N(t)\right|^2 dt\right] + \frac{1}{N}\sum_{l=1}^N \left|x_l^N\right|^2\right.$$

$$\left. + E\left[\frac{1}{N}\sum_{l=1}^N \int_0^T \left|u_l^N(t)\right|^2 dt\right]\right).$$

Thus, for some constant $K < \infty$ depending only on d, T, and L,

$$\frac{1}{N}\sum_{j=1}^N \sum_{k=1}^d E\left[\left|Z_{j,k}^N(n,\xi)\right|^2\right]$$

$$\leq K\left(1 + |\xi|^2 + \frac{1}{N}\sum_{j=1}^N \left|x_j^N\right|^2 + E\left[\frac{1}{N}\sum_{j=1}^N \int_0^T \left|u_j^N(t)\right|^2 dt\right]\right).$$

Letting $C_{\mathbf{s}}^N = \frac{1}{N}\sum_{j=1}^N C_{j,\mathbf{s}}^N$, we have from (A.3) that, for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$,

$$\left|\bar{J}^N(\varphi)\right| \leq \frac{1}{N}\sum_{j=1}^N \left|\bar{J}_j^N(\varphi)\right| \leq C_{\mathbf{s}}^N \|\varphi\|_{\mathbf{s}}.$$

Finally,

$$E\left[(C_{\mathbf{s}}^N)^2\right] \leq \frac{1}{N}\sum_{j=1}^N E\left[(C_{j,\mathbf{s}}^N)^2\right] \leq \sum_{n\in\mathbb{Z}}\int_{\mathbb{R}^d} \frac{K}{(1+n^2)^{s_1}(1+|\xi|^2)^{s_2}}$$

$$\times\left(1 + |\xi|^2 + \sup_{N\geq 1}\frac{1}{N}\sum_{j=1}^N \left|x_j^N\right|^2 + \sup_{N\geq 1} E\left[\frac{1}{N}\sum_{j=1}^N \int_0^T \left|u_j^N(t)\right|^2 dt\right]\right) d\xi,$$

which is finite by Condition 2.3 and (3.5) since $\mathbf{s} \in \mathcal{O}_d$.

A.2. Proof of Lemma 4.8

As in the proof of Lemma 3.1, it suffices to prove the statement in the lemma with $\|\cdot\|_{\mathbf{s}}$ replaced by $\|\cdot\|_{*,\mathbf{s}}$, and once again, abusing notation, we will denote $\|\cdot\|_{*,\mathbf{s}}$ as $\|\cdot\|_{\mathbf{s}}$. Suppose that $\mathbf{s} \in \mathcal{O}_d$ and $(\mu, \mathcal{J}) \in \mathcal{P}_1(\mathcal{X}) \times \mathbf{H}^{-\mathbf{s}}$ are such that $I(\mu, \mathcal{J}) < \infty$. Then there is some $\Theta \in \mathcal{P}^*(\mathcal{J})$ such that $\Theta_{(1)} = \mu$ and

$$\langle \mathcal{J}, \varphi \rangle = G_\varphi(\Theta) = E_\Theta \left[\int_0^T \varphi(t, X(t)) \cdot dX(t) \right]$$

for all $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$. Furthermore, the estimate (4.25) holds for this Θ. By an argument as in the proof of Lemma 3.1,

$$\int_0^T \varphi(t, X(t)) \cdot dX(t) = \sum_{k=1}^d \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \hat{\varphi}_k(n, \xi) Z_k(n, \xi) \, d\xi \qquad \Theta\text{-a.s.,}$$

where $\hat{\varphi}_k$ is defined in (A.1) and

$$\begin{aligned}
Z_k(n, \xi) &\doteq \int_0^T e_{n,\xi}^k(t, X(t)) \cdot dX(t) \\
&= \int_0^T e_{n,\xi}^k(t, X(t)) \cdot b(X(t), \nu_\Theta(t)) \, dt \\
&\quad + \int_{[0,T] \times \mathbb{R}^m} e_{n,\xi}^k(t, X(t)) \cdot \sigma(X(t), \nu_\Theta(t)) y \, \rho(dt, dy)
\end{aligned}$$

Θ-a.s. Since $|e_{n,\xi}^k| \leq T^{-1}$, using (4.24) we have

$$\begin{aligned}
|Z_k(n, \xi)|^2 &\leq \frac{6L^2}{T} \int_0^T \left(1 + |X(t)|^2 + E_\Theta \left[|X(t)|^2 \right] \right) dt \\
&\quad + \frac{2L^2}{T} \int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy),
\end{aligned}$$

and then the bound in (4.25) gives

$$\begin{aligned}
&\sup_{(n,\xi,k) \in \mathcal{I}} E_\Theta \left[|Z_k(n, \xi)|^2 \right] \\
&\leq c' \left(1 + \int_{\mathbb{R}^d} |x|^2 \, \mu_0(dx) + E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right)
\end{aligned}$$

for some $c' < \infty$. Thus, by the Cauchy-Schwarz inequality, for any $\mathbf{s}' = (s_1', s_2') \in \mathcal{O}_d$ and $\varphi \in \mathcal{C}_c^\infty(U \times \mathbb{R}^d, \mathbb{R}^d)$,

$$
\begin{aligned}
|\langle \mathcal{J}, \varphi \rangle|^2 &\le E_\Theta \left[\left| \int_0^T \varphi(t, X(t)) \cdot dX(t) \right|^2 \right] \\
&\le E_\Theta \left[\sum_{k=1}^d \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \frac{|Z_k(n, \xi)|^2}{(1 + n^2)^{s_1'} (1 + |\xi|^2)^{s_2'}} \, d\xi \right] \|\varphi\|_{\mathbf{s}'}^2 \\
&\le c' \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \frac{d\xi}{(1 + n^2)^{s_1'} (1 + |\xi|^2)^{s_2'}} \\
&\quad \times \left(1 + \int_{\mathbb{R}^d} |x|^2 \, \mu_0(dx) + E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) \|\varphi\|_{\mathbf{s}'}^2 \\
&\le C_{\mathbf{s}'}^2 \left(1 + E_\Theta \left[\int_{[0,T] \times \mathbb{R}^m} |y|^2 \, \rho(dt, dy) \right] \right) \|\varphi\|_{\mathbf{s}'}^2
\end{aligned}
$$

where

$$
C_{\mathbf{s}'}^2 \doteq c' \left(1 + \int_{\mathbb{R}^d} |x|^2 \, \mu_0(dx) \right) \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \frac{d\xi}{(1 + n^2)^{s_1'} (1 + |\xi|^2)^{s_2'}} < \infty,
$$

since $\mathbf{s}' = (s_1', s_2') \in \mathcal{O}_d$. The result follows.

A.3. Proof of Lemma 4.4

We will only consider the case where s is not an integer, the proof for the case when s is an integer is a simpler version of the proof given below. An equivalent norm to $\| \cdot \|_s$ in (2.7) can be given as follows (see [28, page 527]): write $s = k + r$ where $k \in \mathbb{N}$ and $r \in (0, 1)$. Then, for $h \in H^s(\mathbb{R}^d, \mathbb{R}^d)$, define

$$
\|h\|_{*,s}^2 \doteq \|h\|_k^2 + \sum_{|\alpha|=k} \|D^\alpha h\|_r^2,
$$

where $\| \cdot \|_k$ is the usual integer Sobolev norm

$$
\|h\|_k^2 = \sum_{0 \le |\alpha| \le k} \|D^\alpha h\|_{L^2}^2,
$$

and $\| \cdot \|_r$ is the fractional Gagliardo-type Sobolev norm

$$
\|h\|_r^2 = \|h\|_{L^2}^2 + [h]_r^2 = \int_{\mathbb{R}^d} |h(x)|^2 \, dx + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|h(x) - h(y)|^2}{|x - y|^{d+2r}} \, dx \, dy. \quad \text{(A.5)}
$$

The norm $\| \cdot \|_{*,s}$ is equivalent to the norm $\| \cdot \|_s$ in (2.7) and thus it suffices to prove Lemma 4.4 by $\| \cdot \|_s$ replaced by $\| \cdot \|_{*,s}$. Henceforth, abusing notation, we will denote this new norm once more as $\| \cdot \|_s$. Now let f and g_M be as in the statement of the lemma. With $B(k)$ as in Definition 4.3(iii), the Leibniz product formula gives, for a multi-index α with $|\alpha| \leq k$,

$$|D^\alpha(g_M(x)f(x))| = \left| \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} g_M(x) D^\beta f(x) \right| \leq B(k) \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \left| D^\beta f(x) \right|$$

and hence for all $M < \infty$

$$\|g_M f\|_k^2 = \sum_{0 \leq |\alpha| \leq k} \int_{\mathbb{R}^d} |D^\alpha g_M(x) f(x)|^2 \, dx$$

$$\leq c_1 \sum_{0 \leq |\beta| \leq k} \int_{\mathbb{R}^d} \left| D^\beta f(x) \right|^2 \, dx = c_1 \|f\|_k^2$$

(A.6)

for some $c_1 = c_1(k) < \infty$. For the r term we follow the proof of [28, Lemma 5.3]. If $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^d, \mathbb{R})$ is such that $0 \leq \psi \leq B_\psi < \infty$ and $h \in H^r(\mathbb{R}^d, \mathbb{R}^d)$ for some $0 < r < 1$, then $\|\psi h\|_{L^2}^2 \leq B_\psi^2 \|h\|_{L^2}^2$. If L_ψ denotes the Lipschitz constant of ψ, then

$$[\psi h]_r^2 = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\psi(x)h(x) - \psi(y)h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$\leq 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\psi(x)h(x) - \psi(x)h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$+ 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\psi(x)h(y) - \psi(y)h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$\leq 2B_\psi^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|h(x) - h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$+ 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|\psi(x) - \psi(y)|^2 |h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$\leq 2B_\psi^2 [h]_r^2 + 2L_\psi^2 \int_{\mathbb{R}^d} \int_{\{|x-y| \leq 1\}} \frac{|h(y)|^2}{|x-y|^{d+2(r-1)}} \, dx \, dy$$

$$+ 8B_\psi^2 \int_{\mathbb{R}^d} \int_{\{|x-y|>1\}} \frac{|h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy$$

$$\leq 2B_\psi^2 [h]_r^2 + 2 \left(L_\psi^2 + 4B_\psi^2 \right) c_2 \|h\|_{L^2}^2,$$

for $c_2 = c_2(r) < \infty$. In the last line, we used the fact that, for some $c_3, c_4 < \infty$ depending on r,

$$\int_{\mathbb{R}^d} \int_{\{|x-y|\leq 1\}} \frac{|h(y)|^2}{|x-y|^{d+2(r-1)}} \, dx \, dy$$

$$\leq \int_{\mathbb{R}^d} \left(\int_{\{|z|\leq 1\}} \frac{1}{|z|^{d+2(r-1)}} \, dz \right) |h(y)|^2 \, dy \leq c_3 \|h\|_{L^2}^2$$

since $d + 2(r-1) < d$, and

$$\int_{\mathbb{R}^d} \int_{\{|x-y|>1\}} \frac{|h(y)|^2}{|x-y|^{d+2r}} \, dx \, dy \leq \int_{\mathbb{R}^d} \left(\int_{\{|z|>1\}} \frac{1}{|z|^{d+2r}} \, dz \right) |h(y)|^2 \, dy \leq c_4 \|h\|_{L^2}^2$$

since $d + 2r > d$. Thus we have that

$$\|\psi h\|_r^2 \leq 8 \left(B_\psi^2 + L_\psi^2 \right) (c_2 + 1) \|h\|_r^2.$$

Then, with $B(k)$ as in Definition 4.3 and $L(k)$ as in (4.2), we obtain that for $|\alpha| = k$

$$\|D^\alpha g_M f\|_r^2 = \left\| \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} g_M D^\beta f \right\|_r^2$$

$$\leq 2^{\alpha!} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta}^2 \left\| D^{\alpha-\beta} g_M D^\beta f \right\|_r^2$$

$$\leq 2^{\alpha!} 8 \left(B(k)^2 + L(k)^2 \right) (c_2 + 1) \sum_{\beta \leq \alpha} \binom{\alpha}{\beta}^2 \left\| D^\beta f \right\|_r^2.$$

Next, for $|\beta| < k$ and some constant $c_5 = c_5(r) < \infty$, we have that

$$\left\| D^\beta f \right\|_r^2 \leq c_5 \left\| D^\beta f \right\|_1^2 = c_5 \left\| D^\beta f \right\|_{L^2}^2 + c_5 \sum_{|\alpha|=1} \left\| D^\alpha D^\beta f \right\|_{L^2}^2$$

$$\leq c_5 \left\| D^\beta f \right\|_{L^2}^2 + c_5 \sum_{|\alpha|=|\beta|+1} \| D^\alpha f \|_{L^2}^2,$$

and hence for some $c_6 = c_6(k, r) < \infty$ and all $M < \infty$

$$\sum_{|\alpha|=k} \|D^\alpha g_M f\|_r^2 \leq c_6 \sum_{|\alpha|=k} \|D^\alpha f\|_r^2 + c_6 \|f\|_k^2. \tag{A.7}$$

Finally, from (A.6) and (A.7), for all $M < \infty$,

$$\|g_M f\|_s^2 = \|g_M f\|_k^2 + \sum_{|\alpha|=k} \|D^\alpha g_M f\|_r^2$$

$$\leq (c_1 + c_6)\|f\|_k^2 + c_6 \sum_{|\alpha|=k} \|D^\alpha f\|_r^2 \leq K\|f\|_s^2,$$

where $K = c_1 + c_6$.

A.4. Proof of Lemma 4.5

Let $\mathbf{s}$, $\mathbf{s}'$, A and K be as in the statement of the lemma. In particular $A \subset \mathbf{H}^{-\mathbf{s}'}$ is such that

$$B \doteq \sup_{F \in A} \|F\|_{-\mathbf{s}'} < \infty, \tag{A.8}$$

and every $F \in A$ has support contained in K. Recall the functions $e_{n,\xi}^k$ for $(n, \xi, k) \in \mathcal{I}$ introduced above (A.1). Let $\{F^l\}_{l \in \mathbb{N}}$ be a sequence in A, and for $l \in \mathbb{N}$ and $(n, \xi) \in \mathbb{Z} \times \mathbb{R}^d$ let

$$\hat{F}^l(n, \xi) \doteq \left(\hat{F}_1^l(n, \xi), \dots, \hat{F}_d^l(n, \xi) \right),$$
$$\hat{F}_k^l(n, \xi) \doteq \left\langle F^l, e_{-n,-\xi}^k \right\rangle, \quad 1 \leq k \leq d. \tag{A.9}$$

Since F^l has compact support, the evaluation on the right-hand side of the second equality above is indeed meaningful (see, *e.g.*, [15, Theorem 9.8]), and for each $l \in \mathbb{N}$ and $n \in \mathbb{Z}$, $\xi \mapsto \hat{F}^l(n, \xi)$ is in $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$. Also, using (A.8) and the compact support property, one can verify (see [15, Theorem 9.22]) that for each $n \in \mathbb{Z}$

$$\sup_{l \geq 1} \sup_{\xi \in \mathbb{R}^d} \left| \hat{F}^l(n, \xi) \right| < \infty \quad \text{and} \quad \sup_{l \geq 1} \sup_{\xi \in \mathbb{R}^d} \left| D_\xi \hat{F}^l(n, \xi) \right| < \infty.$$

Thus, for each $n \in \mathbb{Z}$, $\{\hat{F}^l(n, \cdot), l \in \mathbb{N}\}$ is relatively compact in $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^d)$. By a standard diagonalization procedure, we can pick a subsequence $\{l_j\}$ such that $\{\hat{F}^{l_j}(n, \cdot), j \in \mathbb{N}\}$ converges in $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^d)$ for every n to a limit. We will now show that F^{l_j} is Cauchy in $\mathbf{H}^{-\mathbf{s}}$, which will complete the proof.

By an argument similar to [15, Proposition 9.16], there are constants $c_1(\mathbf{t}, K)$, $c_2(\mathbf{t}, K) < \infty$ for $\mathbf{t} = \mathbf{s}, \mathbf{s}'$ such that, for any $F \in \mathbf{H}^{-\mathbf{s}'} \subset \mathbf{H}^{-\mathbf{s}}$ supported on the compact set K and both $\mathbf{t} = (t_1, t_2) = \mathbf{s}, \mathbf{s}'$,

$$c_1(\mathbf{t}, K) \|F\|_{-\mathbf{t}}^2 \leq \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \left| \hat{F}(n, \xi) \right|^2 \left(1 + n^2 \right)^{-t_1} \left(1 + |\xi|^2 \right)^{-t_2} d\xi$$

$$\leq c_2(\mathbf{t}, K) \|F\|_{-\mathbf{t}}^2, \tag{A.10}$$

where $\hat{F}(n, \xi)$ is defined as in (A.9). In particular, for $j, m \in \mathbb{N}$,

$$c_1(\mathbf{s}, K) \left\| F^{lj} - F^{lm} \right\|_{-\mathbf{s}}^2$$
$$\leq \sum_{n \in \mathbb{Z}} \int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi.$$

Fix $M \in \mathbb{N}$. Then, using $(1 + |\xi|^2)^{-s_2} \leq (1 + |\xi|^2)^{-s_2'}$, we have

$$c_1(\mathbf{s}, K) \left\| F^{lj} - F^{lm} \right\|_{-\mathbf{s}}^2$$
$$\leq \sum_{-M \leq n \leq M} \int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$+ \sum_{|n| > M} \int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$\leq \sum_{-M \leq n \leq M} \int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$+ c_2(\mathbf{s}', K) \left\| F^{lj} - F^{lm} \right\|_{\mathbf{s}'}^2 \frac{1}{(1 + (M + 1)^2)^{s_1 - s_1'}}$$
$$\leq \sum_{-M \leq n \leq M} \int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$+ \frac{4B^2 c_2(\mathbf{s}', K)}{(1 + (M + 1)^2)^{s_1 - s_1'}}.$$

Next, for each $|n| \leq M$ and $R < \infty$, there is a $C(R) < \infty$ such that

$$\int_{\mathbb{R}^d} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$= \int_{\{|\xi| \leq R\}} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$+ \int_{\{|\xi| > R\}} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 \left(1 + n^2\right)^{-s_1} \left(1 + |\xi|^2\right)^{-s_2} d\xi$$
$$\leq C(R) \sup_{|\xi| \leq R} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 + \frac{c_2(\mathbf{s}', K)}{(1 + R^2)^{s_2 - s_2'}} \left\| F^{lj} - F^{lm} \right\|_{\mathbf{s}'}^2$$
$$\leq C(R) \sup_{|\xi| \leq R} \left| \hat{F}^{lj}(n, \xi) - \hat{F}^{lm}(n, \xi) \right|^2 + \frac{4B^2 c_2(\mathbf{s}', K)}{(1 + R^2)^{s_2 - s_2'}}.$$

Combining the above estimates and sending $j, m \to \infty$, since $\{\hat{F}^{l_j}(n, \cdot)\}$ converges for every n, we get

$$\limsup_{j, m \to \infty} \left\| F^{l_j} - F^{lm} \right\|_{\mathbf{s}}^2 \leq \frac{4B^2(2M+1)c_2(\mathbf{s}', K)}{c_1(\mathbf{s}, K)(1+R^2)^{s_2 - s_2'}} + \frac{4B^2 c_2(\mathbf{s}', K)}{c_1(\mathbf{s}, K)(1+(M+1)^2)^{s_1 - s_1'}}.$$

The result now follows on first sending $R \to \infty$ and then $M \to \infty$.

References

[1] R. A. ADAMS and J. F. FOURNIER, "Sobolev Spaces", Pure and Applied Mathematics, Academic Press, Amsterdam, 2003.

[2] M. BOUÉ and P. DUPUIS, *A variational representation for certain functionals of Brownian motion*, Ann. Probab. **26** (1998), 1641–1659.

[3] L. BERTINI, P. BUTTÀ and A. PISANTE, *Stochastic Allen-Cahn approximation of the mean curvature flow: large deviations upper bound*, Arch. Ration. Mech. Anal. **224** (2017), 659–707.

[4] W. BRAUN and K. HEPP, *The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles*, Comm. Math. Phys. **56** (1977), 101–113.

[5] A. BUDHIRAJA and M. CONROY, *Empirical measure and small noise asymptotic under large deviation scaling for interacting diffusions*, J. Theoret. Probab. **35** (2022), 295–349.

[6] A. BUDHIRAJA and P. DUPUIS, *A variational representation for positive functionals of infinite dimensional Brownian motion*, Probab. Math. Statist. **20** (2000), 39–61.

[7] A. BUDHIRAJA and P. DUPUIS, "Analysis and Approximation of Rare Events: Representations and Weak Convergence Methods", Probability Theory and Stochastic Modelling, Vol. 94, Springer, New York, 2019.

[8] A. BUDHIRAJA, P. DUPUIS and M. FISCHER, *Large deviation properties of weakly interacting processes via weak convergence methods*, Ann. Probab. **40** (2012), 74–102.

[9] D. A. DAWSON, *Critical dynamics and fluctuations for a mean field model of cooperative behavior*, J. Stat. Phys. **31** (1983), 29–85.

[10] D. A. DAWSON and J. GÄRTNER, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics **20** (1987), 247–308.

[11] P. DUPUIS and R. S. ELLIS, "A Weak Convergence Approach to the Theory of Large Deviations", Wiley Series in Probability and Statistics, John Wiley & Sons, New York, 1997.

[12] S. N. ETHIER and T. G. KURTZ, "Markov Processes: Characterization and Convergence", Wiley Series in Probability and Statistics, John Wiley & Sons, New York, 1986.

[13] F. FLANDOLI, M. GUBINELLI, M. GIAQUINTA and V. M. TORTORELLI, *Stochastic currents*, Stochastic Process. Appl. **115** (2005), 1583–1601.

[14] F. FLANDOLI and C. A. TUDOR, *Brownian and fractional Brownian stochastic currents via Malliavin calculus*, J. Funct. Anal. **258** (2010), 279–306.

[15] G. B. FOLLAND, "Real Analysis: Modern Techniques and Their Applications", 2nd ed., Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, John Wiley & Sons, New York, 1999.

[16] M. FORNASIER, S. LISINI, C. ORRIERI and G. SAVARÉ, *Mean-field optimal control as Gamma-limit of finite agent controls*, European J. Appl. Math. **30** (2019), 1153–1186.

[17] M. GIAQUINTA, G. MODICA and J. SOUČEK, "Cartesian Currents in the Calculus of Variations I", A Series of Modern Surveys in Mathematics, Vol. 37, Springer, Berlin, 1998.

[18] M. GUBINELLI, *Controlling rough paths*, J. Funct. Anal. **216** (2004), 86–140.

[19] S. HERRMANN, P. IMKELLER and D. PEITHMANN, *Large deviations and a Kramers' type law for self-stabilizing diffusions*, Ann. Appl. Probab. **18** (2008), 1379–1423.

[20] S. HERRMANN and J. TUGAUT, *Mean-field limit versus small-noise limit for some inter-acting particle systems*, Commun. Stoch. Anal. **10** (2016), 39–55.

[21] A. KLENKE, "Probability Theory: A Comprehensive Course", 2nd ed., Universitext, Springer, London, 2014.

[22] T. G. KURTZ and J. XIONG, *Particle representations for a class of nonlinear SPDEs*, Stochastic Process. Appl. **83** (1999), 103–126.

[23] T. J. LYONS, *Differential equations driven by rough signals*, Rev. Mat. Iberoam. **14** (1998), 215–310.

[24] M. MARIANI, *A Γ-convergence approach to large deviations*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **18** (2018), 951–976.

[25] H. P. MCKEAN, *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Natl. Acad. Sci. USA **56** (1966), 1907–1911.

[26] S. MÉLÉARD, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, In: "Probabilistic Models for Nonlinear Partial Differential Equations" (Montecatini Terme, 1995), D. Talay and L. Tubaro (eds.), Lecture Notes in Mathematics, Vol. 1627, Springer, Berlin, 1996, 42–95.

[27] K. OELSCHLÄGER, *A martingale approach to the law of large numbers for weakly inter-acting stochastic processes*, Ann. Probab. **12** (1984), 458–479.

[28] E. DI NEZZA, G. PALATUCCI and E. VALDINOCI, *Hitchhiker's guide to fractional Sobolev spaces*, Bull. Sci. Math. **136** (2012), 521–573.

[29] C. ORRIERI, *Large deviations for interacting particle systems: joint mean field and small-noise limit*, Electron. J. Probab. **25** (2020), 1–44.

[30] G. D. REIS, W. SALKELD and J. TUGAUT, *Freidlin-Wentzell LDP in path space for Mckean-Vlasov equations and the functional iterated logarithm law*, Ann. Appl. Probab. **29** (2019), 1487–1540.

[31] T. SHIGA and H. TANAKA, *Central limit theorem for a system of Markovian particles with mean field interactions*, Probab. Theory Related Fields **69** (1985), 439–459.

[32] C. VILLANI, "Optimal Transport, Old and New", Grundlehren der mathematischen WissenschaftenSpringer, Vol. 338, Springer, Berlin, 2009.

Department of Statistics and Operations Research
University of North Carolina
Chapel Hill, NC 27599, USA
budhiraj@email.unc.edu
mconroy@live.unc.edu