

FREIDLIN–WENTZELL LDP IN PATH SPACE FOR MCKEAN–VLASOV EQUATIONS AND THE FUNCTIONAL ITERATED LOGARITHM LAW

BY GONÇALO DOS REIS^{*,1}, WILLIAM SALKELD^{*} AND JULIAN TUGAUT[†]

University of Edinburgh^{} and Université Jean Monnet[†]*

We show two Freidlin–Wentzell-type Large Deviations Principles (LDP) in path space topologies (uniform and Hölder) for the solution process of McKean–Vlasov Stochastic Differential Equations (MV-SDEs) using techniques which directly address the presence of the law in the coefficients and altogether avoiding decoupling arguments or limits of particle systems. We provide existence and uniqueness results along with several properties for a class of MV-SDEs having random coefficients and drifts of superlinear growth.

As an application of our results, we establish a functional Strassen-type result (law of iterated logarithm) for the solution process of a MV-SDE.

CONTENTS

1.	Introduction	1488
	Our contributions	1490
	Existence and uniqueness problem	1490
	Large deviations	1491
	Functional Strassen law	1491
	Our contribution in view of the “decoupling argument”	1492
2.	Preliminaries	1493
2.1.	Notation	1493
	Probability	1493
	Other spaces and norms	1493
2.2.	The Wasserstein metric	1494
3.	McKean–Vlasov equations with locally Lipschitz coefficients	1495
3.1.	Existence and uniqueness of solutions	1495
3.2.	Continuity in time behavior	1502
4.	Large deviations principle	1504
4.1.	Large deviations principle with the supremum norm	1504
4.1.1.	The main result	1506
	Methodology	1506

Received August 2017; revised February 2018.

¹Supported by the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações CMA/FCT/UNL).

MSC2010 subject classifications. Primary 60F10; secondary 60G07.

Key words and phrases. McKean–Vlasov equations, large deviations principle, path-space, Hölder topologies, superlinear growth, functional Strassen law.

4.1.2. Large Deviations principle for ODE Decoupled MV-SDE 1507
 4.1.3. Exponential equivalence 1512
 4.2. Large Deviations Principle in Hölder Topologies 1513
 4.2.1. The main result 1513
 4.2.2. LDP using a decoupling argument 1514
 4.2.3. Proofs 1515
 5. Functional iterated logarithm law 1525
 Functional iterated logarithm law for McKean–Vlasov SDEs 1526
 Appendix: A collection of auxiliary results 1533
 A.1. Classical large deviation principles 1533
 A.2. Large deviation principles in path space topologies 1534
 Acknowledgements 1538
 References 1538

1. Introduction. In this article, we study a general class of McKean–Vlasov Stochastic Differential Equations (MV-SDEs) having drifts of polynomial growth and examine Freidlin–Wentzell-type large deviations principle small noise asymptotics in related path spaces, namely with the supremum- and Hölder-topologies.

MV-SDEs are more involved than classical SDEs as their coefficients depend on the law of the solution. They are sometimes referred to as mean-field SDEs and were first studied in [29]. In a nutshell, these equations describe a limiting behaviour of individual particles having diffusive dynamics and which interact with each other in a “mean-field” sense. The analysis of stochastic particle systems and MV-SDEs interpreted as their limiting equations are of independent interest and appear widely in applications. Examples include molecular dynamics, fluid dynamics [34]; for the behaviour of large-scale interacting agents in economics or social networks or interacting neurons in biology, see [8, 9] and references therein. Recently, there has been a vigorous growth in the literature on MV-SDEs addressing existence and uniqueness [32], smoothness of associated PDEs [5, 12], numerical methods and many other aspects.

We begin by reminding the reader what a Large Deviations Principle (LDP) is. The main goal of the large deviations is to calculate the probability of a rare event. In the case of stochastic processes, the idea is to find a deterministic path around which the diffusion is concentrated with high probability. As a consequence, the stochastic motion can be interpreted as a small perturbation of the deterministic path.

As a simple example, we present the idea of the large deviations principle for a classical diffusion with a constant coefficient diffusion:

$$(1.1) \quad X_\varepsilon(t) = X(0) + \sqrt{\varepsilon}W(t) - \int_0^t \nabla V(X_\varepsilon(s)) ds,$$

where $X(0)$ is deterministic, W is a Brownian motion and V is the so-called confining potential. We also introduce the deterministic path

$$(1.2) \quad \varphi(t) = X(0) - \int_0^t \nabla V(\varphi(s)) ds.$$

Set ν_ε be the law of the diffusion $(X_\varepsilon(t))_{t \in [0,1]}$. Then we say that $(\nu_\varepsilon)_{\varepsilon > 0}$ satisfies a large deviations principle on $C([0, 1])$ equipped with the norm $\|\cdot\|_\infty$ with the rate function I if and only if for any Borel set Γ , we have

$$-\inf_{\phi \in \overset{\circ}{\Gamma}} I(\phi) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log(\nu_\varepsilon(\Gamma)) \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\nu_\varepsilon(\Gamma)) \leq -\inf_{\phi \in \overline{\Gamma}} I(\phi).$$

We will say that I is a good rate function if the level set $\{x : I(x) \leq \alpha\}$ is compact for any α .

In the case of (1.1), it is well known (see [22]) that $(\nu_\varepsilon)_{\varepsilon > 0}$ satisfies a LDP on $(C([0, 1]), \|\cdot\|_\infty)$ with the rate function I defined as

$$I(\phi) := \frac{1}{2} \int_0^1 \|\dot{\phi}(t) + \nabla V(\phi(t))\|^2 dt,$$

if ϕ is absolutely continuous and such that $\phi(0) = X(0)$; we set $I(\phi) := +\infty$ otherwise. It follows that a Borel set Γ of $C([0, 1])$ which contains the deterministic path $(\varphi(t))_{t \in [0,1]}$ in its interior is such that $\inf_{\phi \in \Gamma^0} I(\phi) = 0$.

We work with McKean–Vlasov SDEs satisfying $X_\varepsilon(0) = X_0$ and

$$(1.3) \quad dX_\varepsilon(t) = b_\varepsilon(t, X_\varepsilon(t), \mathcal{L}_t^{X_\varepsilon}) dt + \sqrt{\varepsilon} \sigma_\varepsilon(t, X_\varepsilon(t), \mathcal{L}_t^{X_\varepsilon}) dW(t),$$

where $\mathcal{L}_t^{X_\varepsilon}$ stands for $\text{Law}(X_\varepsilon(t))$. Since the law of the process is present in the coefficients, this equation is nonlinear—in the sense of McKean. Exact assumptions on σ_ε and on b_ε will be given subsequently. Let us discuss a particular case of this McKean–Vlasov diffusion (in dimension one):

$$(1.4) \quad X_\varepsilon(t) = X_\varepsilon(0) + \sqrt{\varepsilon} W(t) - \int_0^t V'(X_\varepsilon(s)) ds - \int_0^t F' * u_s^\varepsilon(X_\varepsilon(s)) ds,$$

where $u_s^\varepsilon = \mathcal{L}_s^{X_\varepsilon}$, $\sigma_\varepsilon(t, x, \mu) := 1$, $b_\varepsilon(t, x, \mu) := -V'(x) - F' * \mu(x)$ and “ $*$ ” is the usual convolution operator.

The motion of the process is generated by three concurrent forces. The first one is the derivative of a potential V —the confining potential. The second influence is a Brownian motion $(W(t))_{t \in \mathbb{R}_+}$. It allows the particle to move against the potential V . The third term—the so-called self-stabilizing term—represents the attraction between all the others trajectories. Indeed, we remark

$$F' * u_s^\varepsilon(X_\varepsilon(s)(\omega_0)) = \int_{\omega \in \Omega} F'(X_\varepsilon(s)(\omega_0) - X_\varepsilon(s)(\omega)) d\mathbb{P}(\omega),$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying measurable space; see [29, 30].

The particle X_ε which verifies (1.4) can be seen as one particle in a continuous mean-field system of an infinite number of particles. The mean-field system that we will consider is a random N -dimensional dynamical system for $i \in \{1, \dots, N\}$ where $X_\varepsilon^{i,N}(0) = X_0$ and

$$\begin{aligned} dX_\varepsilon^{i,N}(t) &= \sqrt{\varepsilon} dB^{i,N}(t) - V'(X_\varepsilon^{i,N}(t)) dt \\ &\quad - \frac{1}{N} \sum_{j=1}^N F'(X_\varepsilon^{i,N}(t) - X_\varepsilon^{j,N}(t)) dt, \end{aligned}$$

where the N Brownian motions $(B^i(t))_{t \in \mathbb{R}_+}$ are independent. Mean-field systems are the subject of a rich literature. The link between the self-stabilizing process and the mean-field system when N goes to $+\infty$ is called propagation of chaos. We say that there is *propagation of chaos* for the system of interacting particles when the law of k fixed particles $X^{i,N}$ tends to the distribution of k independent particles X solving (1.4) with same law when the size of the system N goes to infinity; see [37] under Lipschitz properties; [31] under Lipschitz assumptions but allowing for jumps; [4] if V is a constant; [10] for a uniform result in time in the nonuniformly convex case. For applications, see [11] about social interactions or [13] about the stochastic partial differential equations.

Another side to propagation of chaos are large deviations results which quantify the rate of convergence of the empirical measure in exponential scales. Many LDP results for McKean–Vlasov SDEs exist exploring Sanov-type large deviations for the N -particle empirical measures from the McKean–Vlasov limit. This is a huge field and a small selection of relevant references is given by [6, 14, 15, 17] (see references therein). As argued in [6], these results are a kind of Freidlin–Wentzell small noise asymptotics, but they are “small noise” at a different level (that of measure-valued processes or path-distribution-valued random variables) compared to the usual (process level) Freidlin–Wentzell results being discussed in this work.

Our contributions. We prove our results by dealing with the presence of the laws in the coefficients directly and avoiding arguments on empirical measures or approximation/convergence of measures. Moreover, our LDP result in the Hölder-topology sharpen existing ones in the classical SDE framework.

Existence and uniqueness problem. The existence problem for (1.4) has been investigated by two different methods. The first one consists in the application of a fixed-point theorem; see [4, 10, 27, 30] in the nonconvex case. The other consists in a propagation of chaos; see, for example, [31]. Moreover, it has been proved in [27], Theorem 2.13, that there is a unique strong solution. Further results on existence and uniqueness, but away from our setting, can be found in [7–9, 32]. We highlight [35] for a discussion on counterexamples on uniqueness of solutions.

We work with MV-SDE with dynamics (1.3) and the work closest to ours is that of [24]. There the author provides existence and uniqueness results for (1.3) under similar superlinear growth conditions but his methodology focuses on certain PDE arguments which force the coefficients to be deterministic, time-independent and impose a uniform ellipticity constraint on σ . Our methods are fully probabilistic in nature and lift these restrictions. We assume a random drift b with of spatial superlinear growth satisfying a (noncoercive) monotonicity condition and a random possibly degenerate Lipschitz diffusion coefficient (see Assumption 3.2 below).

Large deviations. The LDP results we present are with respect to the vanishing noise [when $\varepsilon \searrow 0$ in (1.3)] as in Freidlin–Wentzell theory. For instance, [27] investigates the large deviations principle for the McKean–Vlasov diffusion (1.4) in general dimensions, assuming superlinear growth of the drift but imposing coercivity in their monotonicity condition and a constant diffusion term. In particular, they show that the family of laws $(\nu_\varepsilon)_\varepsilon$ satisfies a large deviations principle on $C([0, 1])$ equipped with the uniform norm with the good rate function

$$I(\phi) := \frac{1}{2} \int_0^1 \|\dot{\phi}(t) + \nabla V(\phi(t)) + F'(\phi(t) - \varphi(t))\|_\infty^2 dt,$$

when ϕ is absolutely continuous such that $\phi(0) = X(0)$ and $I(\phi) := +\infty$ otherwise [φ as in (1.2)].

We show a similar result, in the uniform norm, for the family associated to (1.3). However, unlike [27], we assume a Lipschitz σ coefficient (not a constant one) and we do not impose any coercivity condition (strict negativity of the monotonicity constant). For this result, we combine aspects of their work jointly with [16].

Concerning the Hölder topologies LDP, we find inspiration in [3]. Studying standard SDEs, the authors find a way to transfer LDP results from a coarse topology to a finer one; in their case, from supremum norms to Hölder norms. Their method, explained later, relies on establishing the following inequality: $\forall R > 0, \forall \rho > 0, \exists \delta > 0$ and for ε small enough (see Theorem 4.9 below for the precise statement)

$$(1.5) \quad \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] \lesssim \exp\left(-\frac{R}{\varepsilon}\right),$$

for classical SDE’s where $\Phi^x(h)$ is the so-called Skeleton map (an ODE) associated with X_ε^x . This can be thought of as establishing that the probability of X having a high variation in the $\|\cdot\|_\alpha$ -norm given that the input signal (from the Brownian motion) is small in $\|\cdot\|_\infty$ -norm is exponentially small. For this, they assume boundedness and Lipschitz properties of the drift and diffusion coefficients of the SDE X_ε dependent only on the spatial variables. We provide results in the same vein but for the general class of MV-SDEs with drifts of polynomial growth (see Assumption 4.1). Their conditions are stronger than our conditions so our results extend existing results in classical SDE literature. To the best of our knowledge, LDPS in path space in Hölder topologies or general Besov–Orlicz spaces for MV-SDEs remain unexplored.

Our results on LDPS are of general interest and can be applied to the Monte-Carlo simulation of MV-SDEs. They can be used in the spirit of [25] as a way to find the optimal importance sampling measure; see [18].

Functional Strassen law. The final contribution of our work is a functional Strassen law (a type of law of iterated logarithm) for the solution of a MV-SDE.

Strassen's law for a Brownian motion W was originally stated in [36] and says for any t

$$\frac{W(nt)}{\sqrt{n \log \log(n)}} \rightarrow 0 \quad \text{in Prob. as } n \rightarrow \infty \text{ but a.s. convergence does not hold.}$$

We show that if one replaces the Brownian motion W by the solution of a MV-SDE then the result still holds; a by-product is that this statement allows one to characterize Lyapunov functions for such equations.

In broad strokes, the essence of our proof stems from arguments in [1] and is about showing that the set of rescaled paths is relatively compact in the path space topology which implies convergence in probability, but that the set of limit points of this set (connected to the Skeleton of MV-SDEs) is uncountable which implies the failure of almost sure convergence. In [2], the authors show Strassen's result in Hölder topologies for the Brownian motion.

The work closest to ours is [1]. A similar result is shown for standard SDEs with time-independent uniformly Lipschitz and bounded coefficients. These conditions are much stronger than the conditions we impose (roughly our Assumption 3.2 but with b, σ deterministic, time-independent and σ bounded), and hence our results extends the existing results of the classical SDE literature.

In methodological terms, we recast the notion of the re-scaling operators used in [1] to fit the MV-SDE setting (our Definitions 5.2 and 5.4) and most notably so that they act on the process and law in tandem and in the "right way". After this build-up, we prove our main result as described above.

Our contribution in view of the "decoupling argument". From a methodological point of view, many results of standard SDEs can be carried forward to the MV-SDE framework using the so-called "decoupling argument". The latter, is just that after establishing existence and (crucially) uniqueness for a MV-SDE, one can freeze the law (via an independent copy) and the dynamics that remains is that of a standard SDE with an added time dependency. As long as the new time dependency has the right properties, one can apply most of the known results of standard SDE and transfer them to the MV-SDEs setting. Concerning the LDPs, this topic is discussed in Section 4.2.2; concerning the functional Strassen results, see Remark 5.1.

This work is organized as follows.

In Section 2 we introduce this work's notation and in Section 3 we prove the existence/uniqueness results as well as deriving properties of the associated dynamics. The LDP results appear in Section 4. In Section 5, we establish a functional iterated logarithm law (Strassen's law) for the solution of the MV-SDE. Some auxiliary results, including extensions of some other known ones, are provided in the Appendix.

2. Preliminaries.

2.1. *Notation.* We denote by $\mathbb{N} = \{1, 2, \dots\}$ the set of natural numbers; \mathbb{Z} and \mathbb{R} denote the set of integers and reals, respectively; $\mathbb{R}^+ = [0, \infty)$. By $a \lesssim b$, we denote the relation $a \leq Cb$ where $C > 0$ is a generic constant independent of the relevant parameters and may take different values at each occurrence. By $\lfloor x \rfloor$, we denote the largest integer less than or equal to x . Let A be a $d \times d'$ matrix, we denote the transpose of A by A^T .

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a differentiable function. Then we denote ∇f to be the gradient operator and $H[f]$ to be the Hessian operator. ∂_{x_i} is the 1st partial derivative with relation to the i th position.

Probability. Let $0 < T < \infty$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space carrying a d' -dimensional Brownian motion on the interval $[0, T]$. The filtration on this space satisfies the usual assumptions. We denote by \mathbb{E} and $\mathbb{E}[\cdot | \mathcal{F}_t]$ the usual expectation and conditional expectation operator, respectively. For a random variable X (RV in short), we denote its probability distribution (or law) by $\mathcal{L}^X = \mathbb{P} \circ X^{-1}$; the law of a process $(Y(t))_{t \in [0, T]}$ at time t is denoted by $\mathcal{L}_t^Y = \mathbb{P} \circ [Y(t)]^{-1}$.

Let $L^p(\mathcal{F}_t, \mathbb{R}^d)$, $t \in [0, T]$, is the space of \mathbb{R}^d -valued \mathcal{F}_t -measurable RVs X with norm $\|X\|_{L^p} = \mathbb{E}[|X|^p]^{1/p} < \infty$; L^∞ refers to the subset of essentially bounded RVs. $\mathcal{S}^p([0, T])$ is the space of \mathbb{R}^d -valued measurable \mathcal{F} -adapted processes Y satisfying $\|Y\|_{\mathcal{S}^p} = \mathbb{E}[\sup_{t \in [0, T]} |Y(t)|^p]^{1/p} < \infty$; \mathcal{S}^∞ refers to the subset of $\mathcal{S}^p(\mathbb{R}^d)$ of absolutely uniformly bounded processes.

Other spaces and norms. We set $C([0, T])$ as the space of continuous functions $f : [0, T] \rightarrow \mathbb{R}$ endowed with the uniform norm $\|\cdot\|_\infty$. For the space of continuous functions on the interval $[0, T]$ and $\alpha \in (0, 1)$, we define the uniform and the Hölder norm of a function ψ ,

$$\|\psi\|_\infty = \sup_{t \in [0, T]} |\psi(t)| \quad \text{and} \quad \|\psi\|_\alpha = \sup_{s, t \in [0, T]} \frac{|\psi(t) - \psi(s)|}{|t - s|^\alpha}.$$

With $\|\cdot\|_\alpha$, we define the space of α -Hölder continuous functions $f : [0, T] \rightarrow \mathbb{R}$ by $C^\alpha([0, T], \mathbb{R})$; a ball centered on the map ψ and with radius $r > 0$ in this topology $C^\alpha([0, T], \mathbb{R})$ is denoted as $B_\alpha(\psi, r)$; we use $B_\infty(\psi, r)$ to denote the same ball on the topology of the $\|\cdot\|_\infty$ -norm.

We define, for $t \in [0, T]$, the restricted norm $\|\cdot\|_{\alpha, t}$ and $\|\cdot\|_{\infty, t}$ based on $\|\cdot\|_\alpha$ and $\|\cdot\|_\infty$ such that $\|\cdot\|_{\alpha, T} = \|\cdot\|_\alpha$, $\|\cdot\|_{\infty, T} = \|\cdot\|_\infty$ and is defined as

$$\|f\|_{\infty, t} = \sup_{0 \leq s \leq t} |f(s)| \quad \text{and} \quad \|f\|_{\alpha, t} = \sup_{0 \leq r < s \leq t} \frac{|f(s) - f(r)|}{|s - r|^\alpha}.$$

This is similar to the Hölder/Supremum norm and they are also a monotone increasing function with respect to t . It is also clear that $\forall \psi \in C^\alpha([0, T])$ with

$\psi(0) = 0$ we have $\|\psi\|_{\infty,t} \leq \|\psi\|_{\infty} \leq \|\psi\|_{\alpha}$ and $\|\psi\|_{\infty,t} \leq \|\psi\|_{\alpha,t} \leq \|\psi\|_{\alpha}$ (see Lemma A.6).

Let $L^2([0, T])$ denote the space of square integrable functions $f : [0, T] \rightarrow \mathbb{R}$ satisfying $\|f\|_2 := (\int_0^T |f(r)|^2 dr)^{1/2} < \infty$. Let H be the usual Cameron–Martin–Hilbert space for Brownian motion; the space of all absolutely continuous paths on the interval $[0, T]$ which start at 0 and have a derivative almost everywhere which is $L^2([0, T])$ integrable,

$$H := \left\{ h : [0, T] \mapsto \mathbb{R}, h(0) = 0, h(\cdot) = \int_0^\cdot \dot{h}(s) ds; \dot{h} \in L^2([0, T]) \right\}.$$

It is easy to see that if $h \in H$ then $h(0) = 0$, $h \in C^{1/2}([0, T])$ and $\|h\|_{\infty} \leq \|h\|_{1/2} \leq \|\dot{h}\|_2$.

2.2. *The Wasserstein metric.* In this section, we introduce the Wasserstein metric and some results related to it, for in-depth treatments we refer the reader to [38] or [8], Chapter 5. Consider a measurable space (E, \mathcal{E}) and let $\mathcal{P}(E)$ be the class of probability measures in this space. Let $k \in \mathbb{N}$, let $\mathcal{P}_k(E)$ be the space of probability distributions on (E, \mathcal{E}) with finite k th moments. The Dirac delta measure concentrated at a point $x \in E$ is denoted by δ_x . We define a metric on the space of distributions.

DEFINITION 2.1 (Wasserstein metric). Let E be a complete, separable metric space with metric $d : E \times E \rightarrow \mathbb{R}^+$ and σ -algebra \mathcal{E} . Let $\mu, \nu \in \mathcal{P}_2(E)$. We define the Wasserstein distance to be

$$W^{(2)}(\mu, \nu) = \inf \left\{ \left(\int_{E^2} d(x, y)^2 \pi(dx, dy) \right)^{1/2}; \pi \in \mathcal{P}(E \times E) \right\},$$

where $\mu(A) = \int_{E^2} \chi_A(x) \pi(dx, dy)$ and $\nu(B) = \int_{E^2} \chi_B(y) \pi(dx, dy)$.

The Wasserstein metric is a metric and it induces a topology on $\mathcal{P}_2(E)$. This has been shown to be the topology of weak convergence of measure together with the convergence of all moments of order up to 2. It is important to define the Wasserstein distance for a generic complete separable metric space because later on we will be interchanging between measures on \mathbb{R}^d and $C([0, T]; \mathbb{R}^d)$. In order to distinguish between these two types of objects, we denote $m \in \mathcal{P}_2(C([0, T]; \mathbb{R}^d))$ and $m_t \in \mathcal{P}_2(\mathbb{R}^d)$ and we define for $A \subset \mathbb{R}^d$,

$$m_t(A) = \int_{C([0, T]; \mathbb{R}^d)} \mathbb{1}_{\{x(\cdot) \in C([0, T]; \mathbb{R}^d); x(t) \in A\}}(x) m(dx).$$

If one needed a metric on the entire space $\mathcal{P}(E)$ rather than the subset $\mathcal{P}_2(E)$, one could use the *modified Wasserstein distance*

$$W^{(0)}(\mu, \nu) = \inf \left\{ \int_{E^2} [1 \wedge d(x, y)] \pi(dx, dy); \pi \in \mathcal{P}(E \times E) \right\},$$

where π has marginals μ and ν as before. This metric induces that of weak convergence on $\mathcal{P}(E)$.

DEFINITION 2.2. Let $\mathcal{P}_2(E)$ be the set of all probability distributions (measures) on the separable vector space E with finite second orders. Endow this set with two operators called addition $+\mathcal{P}_2 : \mathcal{P}_2(E) \times \mathcal{P}_2(E) \rightarrow \mathcal{P}_2(E)$ and scalar multiplication $\times\mathcal{P}_2 : \mathbb{R}^d \times \mathcal{P}_2(E) \rightarrow \mathcal{P}_2(E)$ such that $\forall \mu, \nu \in \mathcal{P}_2(E), c \in \mathbb{R}^d$ and $A \subset E$ we have

$$(\mu +\mathcal{P}_2 \nu)[A] = \int_E \mu(y - A)\nu(dy) \quad \text{and} \quad (c \times\mathcal{P}_2 \mu)[A] = \mu\left[\frac{A}{c}\right].$$

These operators satisfy the vector axioms and so they form a vector space.

These vector operators are more intuitive if one thinks of the set of probability distributions as the set of all random variables on E . These measure operators represent addition and scalar multiplication of independent random variables with respect to the vector operators within the space E . This vector space, like all vector spaces, has a 0 element. This is the delta distribution centered at the 0 element of E , δ_0 . The convolution of the delta distribution with any other measure is that measure and it remains constant under stretches and compressions of the domain centered around 0.

The next result is a simple computation which we evaluate for the benefit of the reader.

LEMMA 2.3. Take δ_0 . Then for any $\mu \in \mathcal{P}_2(E)$ we have $W^{(2)}(\mu, \delta_0) = (\int_E y^2 \mu(dy))^{1/2}$.

PROOF. Consider a random variable with law δ_0 . We have $X : \Omega \rightarrow E$ with $\mathbb{P}[X \in A] = \delta_0(A)$ for any $A \subset E$. The σ -algebra generated by X is just $\{\Omega, \emptyset\}$. Let $\mu \in \mathcal{P}_2[E]$ be the law of a random variable $Y : \Omega \rightarrow E$ which generates a σ -algebra that X will be measurable with respect to. For any $B \in \sigma(Y)$, we have that $\mathbb{P}[\Omega \cap B] = \mathbb{P}[B] = 1\mathbb{P}[B] = \mathbb{P}[\Omega]\mathbb{P}[B]$ and $\mathbb{P}[B \cap \emptyset] = \mathbb{P}[\emptyset] = 0 = \mathbb{P}[B]\mathbb{P}[\emptyset]$. Hence $\sigma(X)$ and $\sigma(Y)$ are independent. Therefore, we have that the joint density function of X and Y is just $\mu(dy)\delta_0(dx)$ and the conclusion follows. \square

3. McKean–Vlasov equations with locally Lipschitz coefficients.

3.1. Existence and uniqueness of solutions. We start with a slight generalization of the existence and uniqueness result under Lipschitz conditions in [7], Theorem 1.7. Let W be a d' -dimensional Brownian motion and take the progressively measurable maps $b : [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d'}$.

We introduce, for $0 \leq t \leq T < \infty$ the dynamics of a process Y as

$$(3.1) \quad dY(t) = b(t, Y(t), \mathcal{L}_t^Y) dt + \sigma(t, Y(t), \mathcal{L}_t^Y) dW(t),$$

for $Y(0) \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$ and where \mathcal{L}_t^Y denotes the law of $Y(t)$.

THEOREM 3.1. *Suppose that b and σ are integrable in the sense that*

$$\mathbb{E} \left[\left(\int_0^T |b(t, \omega, 0, \delta_0)| dt \right)^2 \right] < \infty \quad \text{and} \quad \mathbb{E} \left[\int_0^T |\sigma(t, \omega, 0, \delta_0)|^2 dt \right] < \infty,$$

and Lipschitz in the sense that $\exists L > 0$ such that $\forall t \in [0, T], \forall \omega \in \Omega, \forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that

$$\begin{aligned} &|b(t, \omega, x, \mu) - b(t, \omega, x', \mu')| + |\sigma(t, \omega, x, \mu) - \sigma(t, \omega, x', \mu')| \\ &\leq L(|x - x'| + W^{(2)}(\mu, \mu')). \end{aligned}$$

Suppose further that $X(0) \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ is a square integrable random variable which is independent of the Brownian motion. Then there exists a unique solution for $Y \in \mathcal{S}^2([0, T]; \mathbb{R}^d)$ to the MV-SDE (3.1) and $\mathcal{L}_0^Y \in \mathcal{P}_2(\mathbb{R}^d)$ where \mathcal{L}_t^Y is the probability distribution of the random variable $Y(t)$.

PROOF. For $b(\cdot, 0, \delta_0)$ satisfying $\mathbb{E}[\int_0^T |b(t, \omega, 0, \delta_0)|^2 dt] < \infty$, the result is known, for example, [7], Theorem 1.7. A close inspection of that proof shows that this condition is not sharp. In particular, the result holds with the slightly weaker integrability condition found in the statement of the theorem we present here. The verification is straightforward and we do not carry it out. \square

We extend the previous result to the locally Lipschitz case; see [32] for other results. We work with general monotonicity assumptions without imposing coercivity restrictions. We also sharpen the integrability assumptions and leave it to the reader to verify that the proof in [7] can be sharpened.

ASSUMPTION 3.2. Let $p \geq 2$. The progressively measurable maps $b : [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d}$ satisfy that $\exists L > 0$ such that:

1. $Y(0) \in L^p(\mathcal{F}_0; \mathbb{R}^d; \mathbb{P})$ be independent of the Brownian motion.
2. Integrability: b and σ satisfy

$$\mathbb{E} \left[\left(\int_0^T |b(t, 0, \delta_0)| dt \right)^p \right], \mathbb{E} \left[\int_0^T |\sigma(t, 0, \delta_0)|^2 dt \right]^{\frac{p}{2}} < \infty,$$

3. σ is Lipschitz: $\forall t \in [0, T], \forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have

$$|\sigma(t, x, \mu) - \sigma(t, x', \mu')| \leq L(|x - x'| + W^{(2)}(\mu, \mu')),$$

4. b satisfies the monotone growth condition in x and is Lipschitz in μ : $\forall t \in [0, T], \forall x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that

$$\langle x - x', b(t, x, \mu) - b(t, x', \mu) \rangle_{\mathbb{R}^d} \leq L|x - x'|^2 \quad \text{and}$$

$$|b(t, x, \mu) - b(t, x, \mu')| \leq LW^{(2)}(\mu, \mu'),$$

5. b is Locally Lipschitz with polynomial growth in x : $\exists q \in \mathbb{N}$ such that $q > 1$ and $\forall t \in [0, T], \forall \mu \in \mathcal{P}_2(\mathbb{R}^d), \forall x, x' \in \mathbb{R}^d$ we have

$$|b(t, x, \mu) - b(t, x', \mu)| \leq L(1 + |x|^{q-1} + |x'|^{q-1})|x - x'|.$$

THEOREM 3.3. *Let $p \geq 2$. Recall the dynamics of Y given by (3.1), where the drift and diffusion coefficients b, σ and initial RV $Y(0)$ satisfy Assumption 3.2 with $p \geq 2$. Then there exists a unique solution $Y \in \mathcal{S}^p \cap \mathcal{S}^2$ to (3.1) and $\exists C > 0$ such that*

$$\mathbb{E}\left[\sup_{t \in [0, T]} |Y(t)|^p\right] \leq C\left(\mathbb{E}[|Y(0)|^p] + \mathbb{E}\left[\left(\int_0^T |b(t, 0, \delta_0)| dt\right)^p\right] + \mathbb{E}\left[\left(\int_0^T |\sigma(s, 0, \delta_0)|^2 ds\right)^{\frac{p}{2}}\right]\right)e^{CT}.$$

PROOF. Consider the operator

$$\Xi : \mathcal{P}_2(C([0, T], \mathbb{R}^d)) \rightarrow \mathcal{P}_2(C([0, T], \mathbb{R}^d)),$$

where $\Xi(\mu) = \mathcal{L}^{Y^\mu}$ denotes the law of the SDE’s solution Y^μ with dynamics

$$dY^\mu(t) = b(t, Y^\mu(t), \mu_t) dt + \sigma(t, Y^\mu(t), \mu_t) dW(t), \quad Y^\mu(0) = Y(0).$$

We start by showing that given some μ , a solution to the above SDE exists. Let $\mu \in \mathcal{P}_2(C([0, T], \mathbb{R}^d))$. Define

$$\hat{b}^\mu(t, x) = b(t, x, \mu_t), \quad \hat{\sigma}^\mu(t, x) = \sigma(t, x, \mu_t).$$

Then we have

$$\begin{aligned} &\mathbb{E}\left[\left(\int_0^T |\hat{b}^\mu(t, 0)| dt\right)^p\right] \\ &\leq \mathbb{E}\left[\left(\int_0^T |b(t, 0, \delta_0)| + L \cdot W^{(2)}(\mu_t, \delta_0) dt\right)^p\right] \\ &\leq 2^{p-1}\mathbb{E}\left[\left(\int_0^T |b(t, 0, \delta_0)| dt\right)^p\right] + 2^{p-1}L^p T^p \cdot W^{(2)}(\mu, \delta_0)^p < \infty, \end{aligned}$$

and similarly

$$\begin{aligned} &\mathbb{E}\left[\left(\int_0^T |\hat{\sigma}^\mu(t, 0)|^2 dt\right)^{\frac{p}{2}}\right] \\ &\leq 2^{p-1}\mathbb{E}\left[\left(\int_0^T |\sigma(t, 0, \delta_0)|^2 dt\right)^{\frac{p}{2}}\right] + 2^{p-1}L^p T^p \cdot W^{(2)}(\mu, \delta_0)^p < \infty. \end{aligned}$$

Also we have that $\hat{b}^\mu(t, x)$ is locally Lipschitz, satisfies a monotone growth condition and $\hat{\sigma}^\mu(t, x)$ has Lipschitz growth in its spacial variables. Therefore, by the methods in [28], Theorem 3.6, we have that a unique solution exists in $S^p([0, T])$. Since $p \geq 2$, we can conclude that $\mathcal{L}^{Y^\mu} \in \mathcal{P}_2(C([0, T], \mathbb{R}^d))$.

Using Itô’s formula, we have that

$$\begin{aligned}
 & W^{(2)}(\Xi(\mu), \Xi(\nu))^2 \\
 & \leq \mathbb{E} \left[\sup_{t \in [0, T]} |Y^\mu(t) - Y^\nu(t)|^2 \right] \\
 (3.2) \quad & \leq 2\mathbb{E} \left[\int_0^T |(Y^\mu(s) - Y^\nu(s), b(s, Y^\mu(s), \mu_s) - b(s, Y^\nu(s), \nu_s))| ds \right] \\
 (3.3) \quad & + 2\mathbb{E} \left[\sup_{t \in [0, T]} \int_0^t \langle Y^\mu(s) - Y^\nu(s), \right. \\
 & \left. [\sigma(s, Y^\mu(s), \mu_s) - \sigma(s, Y^\nu(s), \nu_s)] dW(s) \right] \\
 (3.4) \quad & + \mathbb{E} \left[\int_0^T |\sigma(s, Y^\mu(s), \mu_s) - \sigma(s, Y^\nu(s), \nu_s)|^2 ds \right].
 \end{aligned}$$

First, we apply the monotonicity and Lipschitz properties to get

$$\begin{aligned}
 (3.2) \quad & \leq 2L\mathbb{E} \left[\int_0^T |Y^\mu(s) - Y^\nu(s)|^2 ds \right] \\
 & + 2L\mathbb{E} \left[\int_0^T |Y^\mu(s) - Y^\nu(s)| W^{(2)}(\mu_s, \nu_s) ds \right] \\
 & \leq 2L \int_0^T \mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty, s}] ds + \frac{\mathbb{E}[\|Y^\mu - Y^\nu\|_\infty^2]}{3} \\
 & + 3L^2 \int_0^T W^{(2)}(\mu_s, \nu_s)^2 ds.
 \end{aligned}$$

Second, we use the Burkholder–Davis–Gundy inequality and Lipschitz properties,

$$\begin{aligned}
 (3.3) \quad & \leq 2\mathbb{E} \left[\left(\int_0^T |(Y^\mu(s) - Y^\nu(s))^T (\sigma(s, Y^\mu(s), \mu_s) - \sigma(s, Y^\nu(s), \nu_s))|^2 ds \right)^{\frac{1}{2}} \right] \\
 & \leq 2\mathbb{E} \left[\|Y^\mu - Y^\nu\|_\infty \left(\int_0^T |\sigma(s, Y^\mu(s), \mu_s) - \sigma(s, Y^\nu(s), \nu_s)|^2 ds \right)^{\frac{1}{2}} \right] \\
 & \leq \frac{\mathbb{E}[\|Y^\mu - Y^\nu\|_\infty^2]}{3} + 6L^2 \int_0^T \mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty, s}^2] ds \\
 & + 6L^2 \int_0^T W^{(2)}(\mu_s, \nu_s)^2 ds.
 \end{aligned}$$

Third, using the Lipschitz properties again we get

$$(3.3) \leq 2L^2 \int_0^T \mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty,s}^2] ds + 2L^2 \int_0^T W^{(2)}(\mu_s, \nu_s)^2 ds.$$

Combining (3.2), (3.3) and (3.4) gives that

$$\begin{aligned} \frac{\mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty}^2]}{3} &\leq (8L^2 + 2L) \int_0^T \mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty,s}^2] ds \\ &\quad + 11L^2 \int_0^T W^{(2)}(\mu_s, \nu_s)^2 ds. \end{aligned}$$

Applying Grönwall to this yields a control to the initial inequality

$$W^{(2)}(\Xi(\mu), \Xi(\nu))^2 \leq \mathbb{E}[\|Y^\mu - Y^\nu\|_{\infty}^2] \leq K \int_0^T W_s^{(2)}(\mu, \nu)^2 ds,$$

where $K = 11L^2 e^{(24L^2 + 6L)T}$. Applying Ξ inductively j times yields

$$\begin{aligned} W^{(2)}(\Xi^j(\mu), \Xi^j(\nu))^2 &\leq K^j \int_0^T \int_0^{t_1} \dots \int_0^{t_{j-1}} W_{t_j}^{(2)}(\mu, \nu)^2 dt_1 \dots dt_j \\ &\leq K^j \int_0^T \frac{(T - t_j)^{j-1}}{(j - 1)!} W_{t_j}^{(2)}(\mu, \nu)^2 dt_j \\ &\leq \frac{K^j T^j}{j!} W^{(2)}(\mu, \nu)^2. \end{aligned}$$

Choosing j large enough ensures that Ξ^j is a contraction operator. Therefore, Ξ has a unique fixed point. Hence we conclude that the Picard sequence of random processes $Y^0(t) = Y(0)$ and

$$dY^n(t) = b(t, Y^n(t), \mathcal{L}_t^{Y^{n-1}}) dt + \sigma(t, Y^n(t), \mathcal{L}_t^{Y^{n-1}}) dW(t),$$

converges in \mathcal{S}^2 and the limit solves the MV-SDE (3.1). From this we conclude that a unique solution exists.

Step 2: Moment calculations. Recall the dynamics of Y from (3.1). By Itô’s formula, we have

$$\begin{aligned} |Y(t)|^p &= |Y(0)|^p + p \int_0^t |Y(s)|^{p-2} \langle Y(s), b(s, Y(s), \mathcal{L}_s^Y) \rangle ds \\ &\quad + p \int_0^t |Y(s)|^{p-2} \langle Y(s), \sigma(s, Y(s), \mathcal{L}_s^Y) \rangle dW(s) \\ &\quad + \frac{p}{2} \int_0^t |Y(s)|^{p-2} |\sigma(s, Y(s), \mathcal{L}_s^Y)|^2 ds \\ &\quad + \frac{p(p-2)}{2} \int_0^t |Y(s)|^{p-4} \cdot |Y(s)^T \sigma(s, Y(s), \mathcal{L}_s^Y)|^2 ds. \end{aligned}$$

Therefore,

$$\begin{aligned}
 \mathbb{E}[\|Y\|_\infty^p] &= \mathbb{E}[|Y(0)|^p] \\
 (3.5) \quad &+ p\mathbb{E}\left[\int_0^T |Y(s)|^{p-2} \langle Y(s), b(s, Y(s), \mathcal{L}_s^Y) \rangle ds\right] \\
 (3.6) \quad &+ p\mathbb{E}\left[\sup_{0 \leq t \leq T} \int_0^t |Y(s)|^{p-2} \langle Y(s), \sigma(s, Y(s), \mathcal{L}_s^Y) \rangle dW(s)\right] \\
 (3.7) \quad &+ \frac{p}{2}\mathbb{E}\left[\int_0^T |Y(s)|^{p-2} |\sigma(s, Y(s), \mathcal{L}_s^Y)|^2 ds\right] \\
 (3.8) \quad &+ \frac{p(p-2)}{2}\mathbb{E}\left[\int_0^T |Y(s)|^{p-4} \cdot |Y(s)^T \sigma(s, Y(s), \mathcal{L}_s^Y)|^2 ds\right].
 \end{aligned}$$

By the triangle property, we have

$$\begin{aligned}
 (3.9) \quad (3.5) &\leq p\mathbb{E}\left[\int_0^T |Y(s)|^{p-2} \langle Y(s), b(s, Y(s), \mathcal{L}_s^Y) - b(s, 0, \mathcal{L}_s^Y) \rangle ds\right] \\
 (3.10) \quad &+ p\mathbb{E}\left[\int_0^T |Y(s)|^{p-2} \langle Y(s), b(s, 0, \mathcal{L}_s^Y) - b(s, 0, \delta_0) \rangle ds\right] \\
 (3.11) \quad &+ p\mathbb{E}\left[\int_0^T |Y(s)|^{p-2} \langle Y(s), b(s, 0, \delta_0) \rangle ds\right].
 \end{aligned}$$

Using the monotone property of b yields $(3.9) \leq pL \int_0^T \mathbb{E}[\|Y\|_{\infty,s}^p] ds$. Using the Lipschitz property of b in the distribution variable and Lemma 2.3 yields

$$\begin{aligned}
 (3.10) &\leq pL \int_0^T \mathbb{E}[\|Y\|_{\infty,s}^{p-1}] \mathbb{E}[\|Y\|_{\infty,s}^2]^{1/2} ds \\
 &\leq pL \int_0^T \left(\frac{(p-1)\mathbb{E}[\|Y\|_{\infty,s}^{p-1}]^{p/p-1}}{p} + \frac{\mathbb{E}[\|Y\|_{\infty,s}^2]^{p/2}}{p} \right) ds \\
 &\leq pL \int_0^T \mathbb{E}[\|Y\|_{\infty,s}^p] ds.
 \end{aligned}$$

Using the integrability properties of b yields

$$\begin{aligned}
 (3.11) &\leq \mathbb{E}\left[\|Y\|_\infty^{p-1} \int_0^T |b(s, 0, \delta_0)| ds\right] \\
 &\leq \frac{\mathbb{E}[\|Y\|_\infty^p]}{n} + n^{p-1}(p-1)^{p-1} \mathbb{E}\left[\left(\int_0^T |b(s, 0, \delta)| ds\right)^p\right],
 \end{aligned}$$

where $n \in \mathbb{N}$ which will be chosen later.

By the Burkholder–Davis–Gundy inequality, the Lipschitz properties and Lemma 2.3 we have

$$\begin{aligned}
 (3.6) &\leq pC_1\mathbb{E}\left[\left(\int_0^T|Y(s)|^{2p-4}|Y(s)^T\sigma(s,Y(s),\mathcal{L}_s^Y)|^2ds\right)^{\frac{1}{2}}\right] \\
 &\leq pC_1\mathbb{E}\left[\|Y\|_{\infty,s}^{\frac{p}{2}}\left(\int_0^T|Y(s)|^{p-2}\cdot|\sigma(s,Y(s),\mathcal{L}_s^Y)|^2ds\right)^{\frac{1}{2}}\right] \\
 &\leq\frac{\mathbb{E}[\|Y\|_{\infty,s}^p]}{n}+p^2C_1^2n\mathbb{E}\left[\int_0^T|Y(s)|^{p-2}\cdot|\sigma(s,Y(s),\mathcal{L}_s^Y)|^2ds\right] \\
 (3.12) &\leq\frac{\mathbb{E}[\|Y\|_{\infty,s}^p]}{n}+3p^2C_1^2nL^2\left(\int_0^T\mathbb{E}[\|Y\|_{\infty,s}^p]ds\right. \\
 &\quad\left.+\int_0^T\mathbb{E}[\|Y\|_{\infty,s}^{p-2}]\cdot\mathbb{E}[\|Y\|_{\infty,s}^2]ds\right)
 \end{aligned}$$

$$(3.13) \quad +3p^2C_1^2n\mathbb{E}\left[\int_0^T|Y(s)|^{p-2}|\sigma(s,0,\delta_0)|^2ds\right].$$

Terms (3.12) are dealt with in the same way as terms (3.9) and (3.10). For (3.13), we proceed as follows:

$$\begin{aligned}
 (3.13) &\leq\mathbb{E}\left[\|Y\|_{\infty}^{p-2}\left(3p^2C_1^2n\int_0^T|\sigma(s,0,\delta_0)|^2ds\right)\right] \\
 &\leq\frac{\mathbb{E}[\|Y\|_{\infty}^p]}{n}+2\cdot3^{\frac{p}{2}}\cdot n^{p-1}C_1^p(p-2)^{\frac{p-2}{2}}p^{\frac{p}{2}}\mathbb{E}\left[\left(\int_0^T|\sigma(s,0,\delta_0)|^2ds\right)^{\frac{p}{2}}\right].
 \end{aligned}$$

Third, we have

$$\begin{aligned}
 (3.7) + (3.8) &\leq\frac{p(p-1)}{2}\mathbb{E}\left[\int_0^T|Y(s)|^{p-2}|\sigma(s,Y(s),\mathcal{L}_s^Y)|^2ds\right] \\
 (3.14) &\leq\frac{3p(p-1)L^2}{2}\left(\int_0^T\mathbb{E}[\|Y\|_{\infty,s}^p]ds\right. \\
 &\quad\left.+\int_0^T\mathbb{E}[\|Y\|_{\infty,s}^{p-2}]\cdot\mathbb{E}[\|Y\|_{\infty,s}^2]ds\right) \\
 &\quad+\frac{3p(p-1)}{2}\mathbb{E}\left[\|Y\|_{\infty}^{p-2}\int_0^T|\sigma(s,0,\delta_0)|^2ds\right]
 \end{aligned}$$

and

$$\begin{aligned}
 (3.14) &\leq\frac{\mathbb{E}[\|Y\|_{\infty}^p]}{n}+\left(\frac{n(p-2)}{2}\right)^{\frac{p-2}{2}} \\
 &\quad\cdot(3(p-1))^{\frac{p}{2}}\mathbb{E}\left[\left(\int_0^T|\sigma(s,0,\delta_0)|^2ds\right)^{\frac{p}{2}}\right].
 \end{aligned}$$

Hence we choose $n = 5$ and this can all be rearranged to get

$$\begin{aligned} \frac{\mathbb{E}[\|Y\|_\infty^p]}{5} &\leq \mathbb{E}[|Y(0)|^p] + \tilde{C}_1 \mathbb{E}\left[\left(\int_0^T |\sigma(s, 0, \delta_0)|^2 ds\right)^{\frac{p}{2}}\right] \\ &\quad + \tilde{C}_2 \mathbb{E}\left[\left(\int_0^T |b(s, 0, \delta_0)| ds\right)^p\right] + \tilde{C}_3 \int_0^T \mathbb{E}[\|Y\|_{\infty,s}^p] ds, \end{aligned}$$

where the constants \tilde{C}_1 , \tilde{C}_2 and \tilde{C}_3 are dependent only on p and L . Applying Grönwall’s lemma provides us with the p moment upper bound

$$\begin{aligned} \mathbb{E}[\|Y\|_\infty^p] &\leq 5\left(\mathbb{E}[|Y(0)|^p] + \tilde{C}_1 \mathbb{E}\left[\left(\int_0^T |\sigma(s, 0, \delta_0)|^2 ds\right)^{\frac{p}{2}}\right]\right) \\ &\quad + \tilde{C}_2 \mathbb{E}\left[\left(\int_0^T |b(s, 0, \delta_0)| ds\right)^p\right] e^{\tilde{C}_3 T}. \quad \square \end{aligned}$$

3.2. *Continuity in time behavior.* We next give results describing time-continuity for the process and its law in the appropriate topologies.

PROPOSITION 3.4. *Let Y be the solution of (3.1) satisfying Assumption 3.2 where $q \in \mathbb{N}$ is the order of the polynomial growth of b . Let $n \in \mathbb{N}$ and $n \geq 2$ and additionally assume that*

$$\mathbb{E}\left[\sup_{t \in [0, T]} |b(t, 0, \delta_0)|^{nq}\right], \mathbb{E}\left[\sup_{t \in [0, T]} |\sigma(t, 0, \delta_0)|^{\frac{nq}{2}}\right] < \infty.$$

Then for every $t, s \in [0, T]$

$$W^{(n)}(\mathcal{L}_t^Y, \mathcal{L}_s^Y) \leq \mathbb{E}[|Y(t) - Y(s)|^n]^{\frac{1}{n}} \lesssim |t - s|^{\frac{1}{2}}.$$

PROOF. The proposition’s conditions mean that by the arguments in the proof of Theorem 3.3 we have $\mathbb{E}[\sup_{t \in [0, T]} |Y(t)|^{nq}] < \infty$. Take $0 \leq s \leq t \leq T < \infty$ and a natural number $n \geq 2$. We have

$$|Y(t) - Y(s)|^n \leq \left| \int_s^t b(r, Y(r), \mathcal{L}_r^Y) dr + \int_s^t \sigma(r, Y(r), \mathcal{L}_r^Y) dW(r) \right|^n.$$

We use the growth condition of b and the Lipschitz property of σ and apply the Minkowski inequality to get

$$\begin{aligned} &\mathbb{E}[|Y(t) - Y(s)|^n]^{\frac{1}{n}} \\ &\leq \mathbb{E}\left[\left|\int_s^t b(r, Y(r), \mathcal{L}_r^Y) dr\right|^n\right]^{\frac{1}{n}} + \mathbb{E}\left[\left|\int_s^t \sigma(r, Y(r), \mathcal{L}_r^Y) dW(r)\right|^n\right]^{\frac{1}{n}} \\ &\leq \mathbb{E}\left[\left|\int_s^t |b(r, Y(r), \mathcal{L}_r^Y)| dr\right|^n\right]^{\frac{1}{n}} + \mathbb{E}\left[\left|\int_s^t |\sigma(r, Y(r), \mathcal{L}_r^Y)|^2 dr\right|^{\frac{n}{2}}\right]^{\frac{1}{n}} \end{aligned}$$

$$\begin{aligned} &\leq |t - s| \mathbb{E}[(\|b(\cdot, 0, \delta_0)\|_\infty + L\|Y\|_\infty^q + \mathbb{E}[\|Y\|^2]^\frac{1}{2})^n]^\frac{1}{n} \\ &\quad + |t - s|^\frac{1}{2} \mathbb{E}[(\|\sigma(\cdot, 0, \delta_0)\|_\infty + L\|Y\|_\infty + \mathbb{E}[\|Y\|_\infty^2]^\frac{1}{2})^n]^\frac{1}{n} \\ &\lesssim |t - s|^\frac{1}{2}. \end{aligned}$$

From the first part of the proposition, we have $\mathbb{E}[|Y(t) - Y(s)|^{2p}] \lesssim |t - s|^p$. The results now follow by applying Kolmogorov’s continuity criterion in a standard fashion. \square

COROLLARY 3.5. *Let Y be the solution of (3.1) under Assumption 3.2 and suppose additionally that $\forall n \in \mathbb{N}$ we have*

$$\mathbb{E}[\|b(\cdot, 0, \delta_0)\|_\infty^n] < \infty, \mathbb{E}[\|\sigma(\cdot, 0, \delta_0)\|_\infty^n] < \infty.$$

Then there is a modification of $Y(\cdot)$, $\tilde{Y}(\cdot)$, which is sample-continuous, almost surely equal to $Y(\cdot)$ and α -Hölder continuous for $\alpha < 1/2$.

PROOF. Under these stronger conditions, we have $\forall n \in \mathbb{N}$ that $\mathbb{E}[|Y(t) - Y(s)|^n] \lesssim |t - s|^{n/2}$. Therefore, we apply the Kolmogorov continuity criterion from [33], Theorem 2.2.3, and conclude. \square

The final result concerns C^1 -regularity (in time) of the expected value of maps of the MV-SDE’s solution.

PROPOSITION 3.6 (Regularity in time). *Let $\phi \in C^{1,2}([0, T] \times \mathbb{R}^d)$ and suppose that ϕ , its first derivative, $\nabla_x \phi(\cdot, \cdot)$, and Hessian, $H[\phi](\cdot, \cdot)$, have polynomial growth such that for some $r > 0$ and some $K > 0$,*

$$\max \left\{ \left| \frac{\partial \phi}{\partial t}(t, y) \right|, |\nabla_x \phi(t, y)|, |H[\phi](t, y)| \right\} \leq K(1 + |y|^r).$$

Suppose that Y is the solution to (3.1) under Assumptions 3.2 with $p := \max\{r + q, 2r + 2\}$ (q is the polynomial growth of b) and hence $Y \in \mathcal{S}^p$.

Then $t \mapsto \mathbb{E}[\phi(t, Y(t))] \in C^1$ and

$$\begin{aligned} \partial_t \mathbb{E}[\phi(t, Y(t))] &= \mathbb{E} \left[\frac{\partial \phi}{\partial t}(t, Y(t)) \right] + \mathbb{E}[\nabla \phi(t, Y(t))^T \cdot b(t, Y(t), \mathcal{L}_t^Y)] \\ &\quad + \mathbb{E}[\text{Tr}(\sigma(t, Y(t), \mathcal{L}_t^Y)^T \cdot H[\phi](t, Y(t)) \cdot \sigma(t, Y(t), \mathcal{L}_t^Y))]. \end{aligned}$$

PROOF. Use Itô’s formula on $\phi(t, Y(t))$, integrate over $[0, t]$ and take expectations. By the integrability/growth assumptions on b and σ , we have $Y \in \mathcal{S}^p$ and in particular $Y \in \mathcal{S}^{2r+2}$. Combining with the polynomial growth of $\nabla \phi(\cdot, \cdot)$ in its spatial variable, we easily conclude that the stochastic integral

$\int_0^\cdot \nabla \phi(s, Y(s)) \sigma(s, Y(s), \mathcal{L}_s^Y) dW(s)$ is a square-integrable martingale, and hence it vanishes under the expectation.

In the previous results, we have shown continuity in time of Y and \mathcal{L}^Y in the appropriate metrics. This, combined with the continuity of b and σ in their variables plus the integrability results, allows to apply Fubini and swap expectations and integrals. Lastly, using the continuity/integrability properties of the involved terms again (notice that here one requires $Y \in \mathcal{S}^{r+q}$), one can compute the time derivative of $t \mapsto \mathbb{E}[\phi(t, Y(t))]$ via the Leibniz differentiation rule for integrals. This yields the lemma’s formula. \square

4. Large deviations principle. In this section, we investigate the family of d -dimensional MV-SDEs indexed to a parameter $\varepsilon > 0$,

$$(4.1) \quad \begin{aligned} X_\varepsilon^x(t) = x + \int_0^t b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) ds \\ + \sqrt{\varepsilon} \int_0^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s). \end{aligned}$$

We derive two types of LDP for the above SDE. The first is an LDP for the supremum norm while the second is an LDP for the Hölder-norm. Throughout we make use of several known sources: [16], [27] and [3]. The main contribution of this section apart from the LDPs themselves, are the techniques needed to deal directly with the law of the solution process inside the coefficients avoiding measure arguments; time dependency of the coefficients is included. For technical convenience, we work on the time interval $[0, 1]$. The extension to the interval $[0, T]$ is straightforward.

ASSUMPTION 4.1. Let $\varepsilon > 0$. Let $b, b_\varepsilon : [0, 1] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\sigma, \sigma_\varepsilon : [0, 1] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d'}$ (deterministic maps) and $x \in \mathbb{R}^d$.

As $\varepsilon \searrow 0$, let the maps b_ε converge uniformly to b and σ_ε converge uniformly to σ . Let b and σ satisfy Assumption 3.2 with the additional restrictions that there exists $M > 0$ such that σ is bounded by M and that there exists $\beta \in (0, 1]$ such that for any $s, s' \in [0, 1]$, for any $y \in \mathbb{R}^d$ and for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we have

$$|\sigma(s, y, \mu) - \sigma(s', y, \mu)| \leq L|s - s'|^\beta, |b(s, y, \mu) - b(s', y, \mu)| \leq L|s - s'|^\beta.$$

REMARK 4.2. For this section, we only worry about the conditions on the coefficients b and σ . However, we will additionally assume that b_ε and σ_ε have adequate conditions to ensure the existence and uniqueness of a solution to the McKean–Vlasov SDE.

4.1. *Large deviations principle with the supremum norm.* To study (4.1) and establish an LDP in the supremum norm for (4.1), we will need to consider several

approximations for it. We start by considering the following ordinary differential equation:

$$(4.2) \quad \dot{\psi}(t) = b(t, \psi(t), \delta_{\psi(t)}), \quad \psi(0) = x.$$

Indeed, informally, when $\varepsilon \searrow 0$ in (4.1), the diffusion term vanishes and we have

$$X_0^x(t) = x + \int_0^t b(s, X_0^x(s), \mathcal{L}_s^{X_0^x}) ds.$$

Of course, since x is deterministic, we deduce that $\mathcal{L}_s^{X_0^x}$ is a Dirac measure centered on the path $X_0^x(\cdot)$. Thus, the ordinary differential equation (4.2) is, from a heuristically standpoint, a good approximation of the stochastic differential equation (4.1) as ε is small. Moreover, the law of $X_\varepsilon^x(t)$ can be approximated by $\delta_{\psi(t)}$. We thus define the following equation (which is closer to a standard SDE):

$$(4.3) \quad Y_\varepsilon^x(t) = x + \int_0^t b(s, Y_\varepsilon^x(s), \delta_{\psi(s)}) ds + \sqrt{\varepsilon} \int_0^t \sigma(s, Y_\varepsilon^x(s), \delta_{\psi(s)}) dW(s).$$

However, (4.3) has a diffusion coefficient which is not constant. As a consequence, we need to discretize

$$(4.4) \quad \begin{aligned} Y_{\varepsilon,m}^x(t) = x + \int_0^t b\left(\frac{\lfloor ms \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) ds \\ + \sqrt{\varepsilon} \int_0^t \sigma\left(\frac{\lfloor ms \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) dW(s), \end{aligned}$$

where $m \in \mathbb{N}$ and will go to infinity. Here, $\lfloor ms \rfloor$ stands for the floor of ms . Lastly, we state a simple result concerning the solvability of (4.2).

LEMMA 4.3. *Under Assumption 4.1, there exists a unique solution $\psi \in C([0, 1])$ to (4.2). Moreover, the map $t \rightarrow \psi(t)$ is Lipschitz continuous.*

PROOF. Existence of a local solution comes from the Peano existence theorem. Uniqueness follows from the monotonicity and Lipschitz/locally Lipschitz properties. In order to get a global solution, we consider the square of the solution, use the monotonicity condition and Grönwall to argue

$$\begin{aligned} \sup_{t \in [0, T]} |\psi(t)| &= |x| + 2 \sup_{t \in [0, T]} \int_0^t \langle \psi(s), b(s, \psi(s), \delta_{\psi(s)}) \rangle ds \\ &\leq |x| + \int_0^T (2L^2 \|\psi\|_{\infty, s} + |b(s, 0, \delta_0)|) ds \\ &\leq \left(|x| + \int_0^T |b(s, 0, \delta_0)| ds \right) e^{2LT}. \end{aligned}$$

From this we see

$$\begin{aligned}
 |\psi(t) - \psi(s)| &\leq \int_s^t (\|b(\cdot, 0, \delta_0)\|_\infty + L(1 + \|\psi\|_\infty^q) + L\|\psi\|_\infty) dr \\
 &\leq O(|t - s|),
 \end{aligned}$$

which yields the Lipschitz continuity. \square

4.1.1. *The main result.* We now state the main theorem concerning a LDP for (4.1) in the topology of the uniform norm and prove it in the remaining subsections.

THEOREM 4.4. *Under the hypotheses of the section, the diffusion X_ε^X satisfies a large deviations principle in $C([0, 1])$ equipped with the topology of the uniform norm with the good rate function*

$$I^X(f) := \inf \frac{1}{2} \int_0^t |\dot{g}(t)|^2 dt,$$

the infimum is taken over the set

$$\left\{ g \in H^{\otimes d'} : f(t) = x + \int_0^t b(s, f(s), \delta_{\psi(s)}) ds + \int_0^t \sigma(s, f(s), \delta_{\psi(s)}) \dot{g}(s) ds \right\},$$

and where ψ is the solution to (4.2).

The ODE appearing in the infimum is easily recognizable as the skeleton of SDE (4.3). Lastly, if σ is a square matrix and if $a := \sigma \sigma^T$ is uniformly strictly positive, the preceding formula for the rate function simplifies into

$$\begin{aligned}
 I^X(\varphi) &:= \frac{1}{2} \int_0^t (\dot{\varphi}(t) - b(t, \varphi(t), \delta_{\psi(t)}))^T \\
 &\quad \times a^{-1}(t, f(t), \delta_{\psi(t)}) (\dot{\varphi}(t) - b(t, \varphi(t), \delta_{\psi(t)})) dt.
 \end{aligned}$$

Methodology. From a methodological point of view, to show that the family $(X_\varepsilon^X)_{\varepsilon>0}$ satisfies a LDP in the supremum norm topology, we first show that the approximation $(Y_{\varepsilon,m}^X)_{\varepsilon>0, m \in \mathbb{N}}$ given in (4.4) satisfies a LDP with the good rate function I^X [defined below in (4.5)]. Then we prove that $(Y_{\varepsilon,m}^X)_{\varepsilon>0, m \in \mathbb{N}}$ is exponentially equivalent to $(Y_\varepsilon^X)_{\varepsilon>0}$ as m goes to infinity and ε goes to zero and since LDPs do not distinguish between exponentially equivalent families (see, e.g., [26], Theorem 2.21), we deduce that $(Y_\varepsilon^X)_{\varepsilon>0}$ satisfies a LDP with the good rate function I^X . Finally, we show that $(X_\varepsilon^X)_{\varepsilon>0}$ and $(Y_\varepsilon^X)_{\varepsilon>0}$ are exponentially equivalent as ε goes to zero. This implies, via the same argument, that $(X_\varepsilon^X)_{\varepsilon>0}$ satisfies a LDP with the good rate function I^X . We make use of standard results from [16], some of which are recalled in the [Appendix](#) below.

4.1.2. *Large deviations principle for Y_ε^x .* We follow [16] plus the techniques used in [27] for having a drift coefficient which is only *locally* Lipschitz, but adequately adapted to the current setting MV-SDE setting.

PROPOSITION 4.5. *Under the hypotheses of the section, the family of diffusions $(Y_\varepsilon^x)_{\varepsilon>0}$ satisfies a large deviations principle in $C([0, 1])$ equipped with the topology of the uniform norm with the good rate function*

$$(4.5) \quad I^x(f) := \inf_{\{g \in H^{\otimes d'} : f(t) = x + \int_0^t b(s, f(s), \delta_{\psi(s)}) ds + \int_0^t \sigma(s, f(s), \delta_{\psi(s)}) \dot{g}(s) ds\}} \frac{1}{2} \times \int_0^1 |\dot{g}(t)|^2 dt,$$

ψ is the unique solution to (4.2).

Before proving Proposition 4.5, we first show that the approximation $(Y_{\varepsilon,m})_{\varepsilon>0, m \in \mathbb{N}}$ satisfies a large deviation principle with the good rate function I_m^x defined as

$$I_m^x(f) := \inf \frac{1}{2} \int_0^1 |\dot{g}(t)|^2 dt,$$

where the infimum is taken over the set

$$\left\{ g \in H^{\otimes d'} : f(t) = x + \int_0^t b\left(\frac{\lfloor ms \rfloor}{m}, f\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) ds + \int_0^t \sigma\left(\frac{\lfloor ms \rfloor}{m}, f\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) \dot{g}(s) ds \right\}.$$

This is an easy exercise using the contraction principle (see [16]) so the proof is omitted. Let us just note that we need to introduce the map F^m defined via $h = F^m(g)$, where

$$h(t) = h\left(\frac{k}{m}\right) + b\left(\frac{k}{m}, h\left(\frac{k}{m}\right), \delta_{\psi\left(\frac{k}{m}\right)}\right)\left(t - \frac{k}{m}\right) + \sigma\left(\frac{k}{m}, h\left(\frac{k}{m}\right), \delta_{\psi\left(\frac{k}{m}\right)}\right)\left(g(t) - g\left(\frac{k}{m}\right)\right),$$

for $t \in [\frac{k}{m}, \frac{k+1}{m}]$, $0 \leq k \leq m - 1$, and $h(0) = x$.

Proposition 4.5 follows by showing exponential equivalence as ε goes to 0 and m goes to infinity of the the families $(Y_\varepsilon^x)_{\varepsilon>0}$ and $(Y_{\varepsilon,m}^x)_{\varepsilon>0, m \in \mathbb{N}}$.

LEMMA 4.6. *For any $\delta > 0$, we have*

$$\lim_{m \rightarrow +\infty} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[\|Y_\varepsilon^x - Y_{\varepsilon,m}^x\|_\infty > \delta]) = -\infty.$$

PROOF. Fix $\delta > 0$. Let $z_t := Y_{\varepsilon, m}^x(t) - Y_\varepsilon^X(t)$, and for any $\rho, R > 0$, define the stopping time

$$\tau := \min\{\inf\{t \geq 0 : |Y_{\varepsilon, m}^x(t)| \geq R + 1\}, \inf\{t \geq 0 : |Y_\varepsilon^X(t)| \geq R + 1\}\},$$

then

$$\tau_1 := \min\left\{1, \inf\left\{t \geq 0 : \left|Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right) - Y_\varepsilon^X(t)\right| \geq \rho\right\}, \tau\right\}.$$

The process $(z_t)_{t \in [0, 1]}$ is of the form (A.1), with $z_0 = 0$,

$$b_t := b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - b\left(t, Y_\varepsilon^X(t), \delta_{\psi(t)}\right),$$

$$\sigma_t := \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - \sigma\left(t, Y_\varepsilon^X(t), \delta_{\psi(t)}\right).$$

Thus, by the local Lipschitz continuity of b , by the global Lipschitz continuity of σ and the definition of τ_1 , it follows that Lemma A.1 is applicable here. Indeed, we have

$$\begin{aligned} |\sigma_t| &= \left| \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - \sigma\left(t, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) \right| \\ &\leq \left| \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi(t)}\right) \right| \\ &\quad + \left| \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi(t)}\right) - \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) \right| \\ &\quad + \left| \sigma\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) - \sigma\left(t, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) \right| \\ &\leq L \left| \psi\left(\frac{\lfloor mt \rfloor}{m}\right) - \psi(t) \right| + \frac{L}{m^\beta} + L|z_t| \\ &\leq M(\rho(m)^2 + |z_t|^2)^{\frac{1}{2}}, \end{aligned}$$

with M large enough and $\rho(m) := \max\{\sup_{t \in [0, 1]} |\psi\left(\frac{\lfloor mt \rfloor}{m}\right) - \psi(t)|, \frac{1}{m^\beta}\}$ which, using the continuity of ψ , goes to 0 as m goes to infinity.

We argue as follows for the drift coefficient $|b_t|$:

$$\begin{aligned} |b_t| &= \left| b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - b\left(t, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) \right| \\ &\leq \left| b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^x\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor mt \rfloor}{m}\right)}\right) - b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi(t)}\right) \right| \\ &\quad + \left| b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X\left(\frac{\lfloor mt \rfloor}{m}\right), \delta_{\psi(t)}\right) - b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) \right| \end{aligned}$$

$$\begin{aligned}
 & + \left| b\left(\frac{\lfloor mt \rfloor}{m}, Y_\varepsilon^X(t), \delta_{\psi(t)}\right) - b(t, Y_\varepsilon^X(t), \delta_{\psi(t)}) \right| \\
 & \leq L \left| \psi\left(\frac{\lfloor mt \rfloor}{m}\right) - \psi(t) \right| + \frac{L}{m^\beta} + L_{R+1}|z_t| \\
 & \leq B_R(\rho(m)^2 + |z_t|^2)^{\frac{1}{2}},
 \end{aligned}$$

with B_R large enough. This yields for any $\delta > 0$ and any $0 < \varepsilon \leq 1$,

$$\varepsilon \log\left(\mathbb{P}\left[\sup_{t \in [0, \tau_1]} |Y_{\varepsilon, m}^x(t) - Y_\varepsilon^X(t)| \geq \delta\right]\right) \leq K_R + \log\left(\frac{\rho(m)^2}{\rho(m)^2 + \delta^2}\right).$$

Hence, by considering first $\varepsilon \rightarrow 0$ and then $m \rightarrow +\infty$,

$$\lim_{m \rightarrow +\infty} \limsup_{\varepsilon \rightarrow 0^+} \varepsilon \log\left(\mathbb{P}\left[\sup_{t \in [0, \tau_1]} |Y_{\varepsilon, m}^x(t) - Y_\varepsilon^X(t)| \geq \delta\right]\right) = -\infty.$$

Now, since

$$(4.6) \quad \{\|Y_{\varepsilon, m}^x - Y_\varepsilon^X\|_\infty > \delta\} \subset \{\tau_1 < 1\} \cup \left\{ \sup_{t \in [0, \tau_1]} |Y_{\varepsilon, m}^x(t) - Y_\varepsilon^X(t)| \geq \delta \right\},$$

the lemma is proved as soon as we show that for all $\rho > 0$ and for any $R > 0$,

$$\lim_{m \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[\tau_1 < 1]) = -\infty.$$

To this end, observe first that for $t \leq \tau_1$

$$\begin{aligned}
 & \left| Y_\varepsilon^X\left(\frac{\lfloor mt \rfloor}{m}\right) - Y_\varepsilon^X(t) \right| \\
 & \leq C_R \left[\frac{1}{m} + \sqrt{\varepsilon} \max_{0 \leq k \leq m-1} \sup_{0 \leq s \leq \frac{1}{m}} \left| W\left(\frac{k}{m} + s\right) - W\left(\frac{k}{m}\right) \right| \right],
 \end{aligned}$$

where C_R is the maximum between the uniform bound of σ , the local bound (on the ball of center zero and radius $R + 1$) of b and the uniform bound of b and σ for the measure variable. Therefore, for all $m > 4C(R)/\rho$,

$$\begin{aligned}
 \mathbb{P}\left[\sup_{0 \leq t \leq \tau_1} \left| Y_\varepsilon^X\left(\frac{\lfloor mt \rfloor}{m}\right) - Y_\varepsilon^X(t) \right| \geq \frac{\rho}{2}\right] & \leq m \mathbb{P}\left[\sup_{0 \leq s \leq \frac{1}{m}} |W(s)| \geq \frac{\frac{\rho}{2} - \frac{C_R}{m}}{\sqrt{\varepsilon} C_R}\right] \\
 & \leq 4dm \exp\left(-m \frac{(\frac{\rho}{2} - \frac{C_R}{m})^2}{2d\varepsilon C_R^2}\right),
 \end{aligned}$$

where the second inequality is the bound of Lemma A.2.

By taking δ sufficiently small, if Y_ε^X exits the ball of center 0 and of radius $R + 1$, then, with high probability [quantified by the limit (4.6)], the process $Y_{\varepsilon, m}^x$ exits the ball of center 0 and of radius R . Consequently, to close the proof it is

sufficient to prove that the probability that $Y_{\varepsilon,m}^x$ leaves the ball of center 0 and radius R is very small as ε goes to zero.

Recall that $(Y_{\varepsilon,m}^x)_{\varepsilon>0}$ satisfies a large deviations principle with the good rate function I_m^x defined previously, and hence we can quantify the probability of exiting from aforementioned ball. We remark that the infimum of I_m^x on the set of paths exiting from the ball of center 0 and radius R goes to infinity as R goes to infinity provided that m is sufficiently large. This remark can be obtained as follows. We consider $f_0 := F^m(0)$ and f which is a path starting from x and exiting from the ball of center 0 and radius R . By g , we denote the function such that $f = F^m(g)$. We dominate $|f(t) - f_0(t)|$ as follows:

$$\begin{aligned} |f(t) - f_0(t)| &\lesssim \int_0^t (1 + |f(s)|^q + |f_0(s)|^q) |f(s) - f_0(s)| |\dot{g}(s)| ds \\ &\quad + \int_0^t |f(s) - f_0(s)| |\dot{g}(s)| ds \\ &\quad + \int_0^t \left| \sigma\left(\frac{\lfloor ms \rfloor}{m}, f_0\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)} \right) \right| |\dot{g}(s)| ds, \end{aligned}$$

by using the properties on b (locally Lipschitz with polynomial growth) and σ (uniformly Lipschitz). However, the last quantity in the integral can be bounded by $C|\dot{g}(s)|$ where C only depends on the function f_0 . In the same vein, we obtain

$$|f(t) - f_0(t)| \lesssim \int_0^t (1 + |f(s) - f_0(s)|^q) |f(s) - f_0(s)| |\dot{g}(s)| ds + \int_0^t |\dot{g}(s)| ds.$$

As $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$, we get

$$|f(t) - f_0(t)| \lesssim \int_0^t (1 + |f(s) - f_0(s)|^{2q}) |f(s) - f_0(s)|^2 ds + \int_0^t |\dot{g}(s)|^2 ds,$$

since we work on a finite time interval. However, if $\|f\|_\infty \geq R$, then we have that $\|f - f_0\|_\infty \geq \xi(R)$ with $\xi(+\infty) = +\infty$. A Grönwall argument suffices to prove that it implies $\int_0^t |\dot{g}(s)|^2 ds \geq \zeta(R)$ with $\zeta(+\infty) = +\infty$. \square

We now are able to prove Proposition 4.5.

PROOF OF PROPOSITION 4.5. Let F be defined on the space $H^{\otimes d'}$ such that $f = F(g)$ is the unique solution of the integral equation

$$f(t) = x + \int_0^t b(s, f(s), \delta_{\psi(s)}) ds + \int_0^t \sigma(s, f(s), \delta_{\psi(s)}) \dot{g}(s) ds.$$

The existence and the uniqueness of a continuous solution is a consequence of the assumptions on b and σ and is standard.² In view of Lemma 4.6, the proof of the

²Local existence of a solution to this ODE comes from Carathéodory’s existence theorem. Uniqueness comes from the monotonicity and Lipschitz/locally Lipschitz properties. Finally, global exist-

theorem is completed by combining Schilder’s theorem and Proposition A.3, as soon as we show that for every $\alpha < \infty$,

$$(4.7) \quad \lim_{m \rightarrow \infty} \sup_{g: \|g\|_{H^{\otimes d'}} \leq \alpha} \|F^m(g) - F(g)\|_\infty = 0.$$

To this end, fix $\alpha < \infty$ and $g \in H^{\otimes d'}$ such that $\|g\|_{H^{\otimes d'}} \leq \alpha$. Let $h = F^m(g)$, $f = F(g)$, and $e(t) := |f(t) - h(t)|^2$. Then for all $t \in [0, 1]$,

$$\begin{aligned} h(t) &= x + \int_0^t b\left(\frac{\lfloor ms \rfloor}{m}, h\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) ds \\ &\quad + \int_0^t \sigma\left(\frac{\lfloor ms \rfloor}{m}, h\left(\frac{\lfloor ms \rfloor}{m}\right), \delta_{\psi\left(\frac{\lfloor ms \rfloor}{m}\right)}\right) \dot{g}(s) ds. \end{aligned}$$

By the Cauchy–Schwarz inequality and the local Lipschitz property on b and the global Lipschitz property on σ ,

$$(4.8) \quad \sup_{0 \leq t \leq 1} \left| h(t) - h\left(\frac{\lfloor mt \rfloor}{m}\right) \right| \leq (\alpha + 1)L_\alpha \delta(m),$$

where $\delta(m)$ is independent of g for any m , and converges to zero as m goes to infinity. To prove the existence of the constant L_α , we remark that $\{g : \|g\|_{H^{\otimes d'}} \leq \alpha\}$ is a compact.

Applying the Cauchy–Schwarz inequality again, it follows by the β -Hölder time continuity of b , σ , the Lipschitz and local Lipschitz continuity of b and the global Lipschitz continuity of σ that

$$\begin{aligned} |f(t) - h(t)| &\leq L_\alpha(\alpha + 1) \sqrt{\int_0^t \left| f(s) - h\left(\frac{\lfloor ms \rfloor}{m}\right) \right|^2 ds} \\ &\quad + L \sqrt{\int_0^t \left| \psi(s) - \psi\left(\frac{\lfloor ms \rfloor}{m}\right) \right|^2 ds} + \frac{L_\alpha(\alpha + 1) + L}{m^\beta}. \end{aligned}$$

Thus, due to (4.8) and the continuity of ψ ,

$$|f(t) - h(t)|^2 = e(t) \leq K_\alpha \int_0^t e(s) ds + K_\alpha \delta(m),$$

with $e(0) = K_\alpha \delta(m)$. Hence, by Grönwall’s lemma, $e(t) \leq K_\alpha \delta(m)^2 e^{K_\alpha t}$, and consequently

$$\|F(g) - F^m(g)\|_\infty \leq \sqrt{K_\alpha} \delta(m) e^{\frac{K_\alpha}{2}},$$

which establishes (4.7) and completes the proof. \square

tence comes from considering the square of the solution and using the monotonicity condition to obtain a linear growth upper bound condition which ensures the solution does not explode. The function f is 1/2-Hölder continuous.

4.1.3. $(Y_\varepsilon^X)_{\varepsilon>0}$ and $(X_\varepsilon^X)_{\varepsilon>0}$ are exponentially equivalent. In order to show that $(X_\varepsilon^X)_{\varepsilon>0}$ satisfies a large deviations principle for the uniform norm with the good rate function I^X , it now is sufficient to prove that the two families of processes $(X_\varepsilon^X)_{\varepsilon>0}$ and $(Y_\varepsilon^X)_{\varepsilon>0}$ are exponentially equivalent.

PROPOSITION 4.7. For any $\delta > 0$, we have

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log \left(\mathbb{P} \left[\sup_{0 \leq t \leq 1} |X_\varepsilon^X(t) - Y_\varepsilon^X(t)| \geq \delta \right] \right) = -\infty.$$

PROOF. The proof is similar to the one of Proposition 4.6 and is also inspired by the proof of [27], Theorem 3.4.

Without loss of generality, we may choose $R > 0$ such that x is in the ball of center 0 and radius $R + 1$. We also assume that $\psi(t)$ does not leave this ball up to time 2. By σ'_R , we denote the first time at which X_ε^X or Y_ε^X exits from the ball, then we put $\sigma_R := \min\{1, \sigma'_R\}$. We consider $z_t := X_\varepsilon^X(t) - Y_\varepsilon^X(t)$. This new process satisfies the following equation:

$$z_t = \int_0^t b_s ds + \sqrt{\varepsilon} \int_0^t \sigma_s dW(s),$$

with

$$b_t := b(t, X_\varepsilon^X(t), \mathcal{L}_t^{X_\varepsilon^X}) - b(t, Y_\varepsilon^X(t), \delta_{\psi(t)}),$$

$$\sigma_t := \sigma(t, X_\varepsilon^X(t), \mathcal{L}_t^{X_\varepsilon^X}) - \sigma(t, Y_\varepsilon^X(t), \delta_{\psi(t)}).$$

Both b_t and σ_t are progressively measurable processes. We now assume that $t \leq \sigma_R$. Then, b and σ are Lipschitz in the spatial variable

$$\begin{aligned} |b_t| &= |b(t, X_\varepsilon^X(t), \mathcal{L}_t^{X_\varepsilon^X}) - b(t, Y_\varepsilon^X(t), \delta_{\psi(t)})| \\ &\leq |b(t, X_\varepsilon^X(t), \mathcal{L}_t^{X_\varepsilon^X}) - b(t, X_\varepsilon^X(t), \delta_{\psi(t)})| \\ &\quad + |b(t, X_\varepsilon^X(t), \delta_{\psi(t)}) - b(t, Y_\varepsilon^X(t), \delta_{\psi(t)})| \\ &\leq L\sqrt{\mathbb{E}[|X_\varepsilon^X(t) - \psi(t)|^2]} + L_R|X_\varepsilon^X(t) - Y_\varepsilon^X(t)| \\ &\leq B_R\sqrt{\rho(\varepsilon)^2 + |z_t^2|}, \end{aligned}$$

where B_R depends only on R and $\rho(\varepsilon) := \sup_{t \in [0, T]} \mathbb{E}\{|X_\varepsilon^X(t) - \psi(t)|^2\}$ goes to 0 as ε goes to 0. Indeed, we can proceed as in [27], Lemma 3.1, to show that $\rho(\varepsilon)$ is small as ε goes to 0.

In the same vein, where M is a constant which does not depend on R , we have

$$|\sigma_t| \leq M\sqrt{\rho(\varepsilon)^2 + |z_t^2|}.$$

Thus, Lemma A.1 is applicable and for any $\delta, \rho > 0$ and for any ε small enough, we have

$$\varepsilon \log\left(\mathbb{P}\left[\sup_{0 \leq t \leq \sigma_R} |z_t| \geq \delta\right]\right) \leq B_R + M^2\left(1 + \frac{d}{2}\right) + \frac{1}{2} \log\left(\frac{\rho^2}{\rho^2 + \delta^2}\right).$$

As $\rho(\varepsilon)$ converges to 0 as $\varepsilon \searrow 0$, we deduce that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log\left(\mathbb{P}\left[\sup_{0 \leq t \leq \sigma_R} |z_t| \geq \delta\right]\right) = -\infty.$$

Now, since

$$\{\|X_\varepsilon^X - Y_\varepsilon^X\|_\infty \geq \delta\} \subset \{\sigma_R < 1\} \cup \left\{ \sup_{0 \leq t \leq \sigma_R} |X_\varepsilon^X(t) - Y_\varepsilon^X(t)| \geq \delta \right\},$$

we can conclude as soon as we show that

$$\lim_{R \rightarrow +\infty} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[\sigma_R < 1]) = -\infty.$$

By τ_R , we denote the first time that Y_ε exits from the ball of center 0 and radius R . If, $Y_\varepsilon^X(\sigma_R)$ is not in the ball of center 0 and radius $R + 1$, then we have immediately $\tau_R < 1$. Conversely, if $X_\varepsilon^X(\sigma_R)$ is not in the ball of center 0 and radius $R + 1$, by taking $\delta < \frac{1}{2}$, we know that with high probability $Y_\varepsilon^X(\sigma_R)$ is not in the ball of center 0 and radius R , which means again $\tau_R < 1$.

However, Y_ε^X satisfies a large deviations principle for the uniform norm with the good rate function I^x . As a consequence,

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[\sigma_R < 1]) &= \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[\tau_R < 1]) \\ &\leq - \inf_{\{g \in H^{\otimes d'} : f(t) = F(g), \|f\|_\infty \geq R\}} \frac{1}{2} \int_0^1 |\dot{g}|^2 dt. \end{aligned}$$

It is not difficult to see that the latter expression approaches $-\infty$ as R goes to ∞ by using the same arguments as those used from the end of the proof of Lemma 4.6. □

Theorem 4.4 now follows easily.

PROOF OF THEOREM 4.4. Following the methodology described in Section 4.1.1 and the above results in combination with [26], Theorem 2.21, our main Theorem 4.4 follows. □

4.2. Large deviations principle in the Hölder topology for X_ε .

4.2.1. *The main result.* Recall the Stochastic process (4.1). We introduce the so-called Skeleton operator Φ for the MV-SDE (4.1) on the Cameron–Martin

space H . In other words, $\Phi : H^{\otimes d'} \rightarrow C([0, 1])$,

$$(4.9) \quad \begin{aligned} \Phi^x(h)(t) = x + \int_0^t & b(s, \Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) ds \\ & + \int_0^t \sigma(s, \Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) \dot{h}(s) ds. \end{aligned}$$

The operator Φ for each $h \in H$ outputs the unique solution to the above ODE. For existence and uniqueness of a solution, see footnote 2.

Following the same method as in Lemma 4.3 and using the Hölder inequality, one can see that

$$\begin{aligned} & |\Phi^x(h)(t) - \Phi^x(h)(s)| \\ & \leq O(|t - s|) + M|t - s|^{\frac{1}{2}} \sqrt{\int_0^T |\dot{h}(r)|^2 dr} \leq O(|t - s|^{\frac{1}{2}}), \end{aligned}$$

so $\Phi(h) \in C^{\frac{1}{2}}([0, T])$. We are now able to state the two main results of this section.

THEOREM 4.8. *Let $\alpha \in (0, 1/2)$. Let A be a Borel set of the space of \mathbb{R}^d -valued continuous paths over $[0, 1]$ in the Hölder topology of $C^\alpha([0, 1])$. Let $\Delta(A) := \inf\{\|\dot{h}\|_2^2/2; h \in H^{\otimes d'}, \Phi^x(h)(\cdot) \in A\}$. Then*

$$-\Delta(\overset{\circ}{A}) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}[X_\varepsilon^x \in A] \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}[X_\varepsilon^x \in A] \leq -\Delta(\bar{A}),$$

where $\overset{\circ}{A}$ and \bar{A} are the interior and closure of the set A with respect to the topology generated by the Hölder norm.

In order to prove the Theorem 4.8, we first prove another LDP-type result [compare with (1.5)].

PROPOSITION 4.9. *Let $h \in H^{\otimes d'}$. Take $\forall R, \rho > 0, \exists \delta, \nu > 0$ such that $\forall 0 < \varepsilon < \nu$,*

$$\mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] \lesssim \exp\left(-\frac{R}{\varepsilon}\right).$$

Intuitively, Proposition 4.9 (proof given below) quantifies the probability of a highly varying process (in $\|\cdot\|_\alpha$ -norm) when the equation’s input signal is small (in $\|\cdot\|_\infty$) [see (1.5)].

4.2.2. LDP using a decoupling argument. In this section, we discuss another method for proving the LDP results for X the solution process to (4.1), called a decoupling argument. The main idea is to freeze the law of X in the original MV-SDE

and understand the outcome as a standard SDE, with solution \hat{X} , where the coefficients $\hat{b}_\varepsilon(t, x) := b_\varepsilon(t, x, \mathcal{L}_t^{X_\varepsilon^x})$ and $\hat{\sigma}_\varepsilon(t, x) := \sigma_\varepsilon(t, x, \mathcal{L}_t^{X_\varepsilon^x})$ are just functions of time and space that have no measure dependency. Observe that by Proposition 3.4, the time regularity will not be affected by the measure dependency since we assume β -Hölder continuity in time for $\beta < \frac{1}{2}$; see Assumption 4.1.

One can prove that $\mathcal{L}^{X_\varepsilon^x} \rightarrow \delta_{\psi, x}$ in distribution as $\varepsilon \searrow 0$ [where ψ solves (4.2)]. The LDP in the uniform topology for the MV-SDE would now follow from a similar LDP under our core conditions for the SDE of \hat{X} . To the best of our knowledge, we were unable to find LDP results in Hölder topologies for SDEs with coefficients which allow for time dependency or monotone growth in the spacial variables. Such LDPs do exist, for example, [3], present the right LDP but under assumptions of uniform Lipschitzness and uniformly boundedness of b and σ plus no time dependency.

Hence the methods and results we present contribute to the MV-SDE literature, but also they are of general interest for the literature on classical SDEs.

4.2.3. *Proofs.* PROOF OF PROPOSITION 4.9. Let $t \in [0, 1]$. Fix $R, \rho > 0$. In order to progress with a Local Lipschitz condition, we first need to consider the function $\Phi^x(h)(\cdot)$ [recall (4.9)] for $h \in H$. This is a continuous solution of an ODE on the compact interval $[0, 1]$. Therefore, it is bounded and we can say that $\exists N > 0$ such that $\|\Phi(h)\|_\infty < N$.

We condition on the event that the process $X_\varepsilon^x(\cdot)$ remains in the ball of radius N and we see

$$\begin{aligned} &\mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] \\ &\leq \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta, \|X_\varepsilon^x\|_\infty < N] \\ &\quad + \mathbb{P}[\|X_\varepsilon^x\|_\infty \geq N]. \end{aligned}$$

We use that we have the LDP result for X_ε^x in a supremum norm and choose N large enough so that

$$\mathbb{P}[\|X_\varepsilon^x\|_\infty \geq N] < \exp\left(-\frac{R}{\varepsilon}\right).$$

We also introduce a step function approximation to discretize the process X_ε^x in (4.1) as, for $l \in \mathbb{N}$,

$$X_\varepsilon^{x,l}(t) = X_\varepsilon^x\left(\frac{j}{l}\right) \quad \text{on the interval } t \in \left(\frac{j}{l}, \frac{j+1}{l}\right], \text{ with } X_\varepsilon^{x,l}(0) = x.$$

Step 1. Analysis of the diffusion term for $h = 0$. Consider

$$\begin{aligned} &\mathbb{P}\left[\left\|\sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s)\right\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X_\varepsilon^x\|_\infty < N\right] \\ (4.10) \quad &\leq \mathbb{P}\left[\left\|\sqrt{\varepsilon} \int_0^\cdot \left[\sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - \sigma_\varepsilon\left(\frac{\lfloor sl \rfloor}{l}, X_\varepsilon^{x,l}, \mathcal{L}_{\frac{\lfloor sl \rfloor}{l}}^{X_\varepsilon^x}\right)\right] dW(s)\right\|_\alpha \right. \end{aligned}$$

$$\begin{aligned}
 &\geq \frac{\rho}{2}, \\
 (4.11) \quad &\frac{1}{l^\beta} + \|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty + \mathbb{E}[\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty^2]^{1/2} \leq \gamma \Big] \\
 (4.12) \quad &+ \mathbb{P}\left[\frac{1}{l^\beta} + \|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty + \mathbb{E}[\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty^2]^{1/2} > \gamma, \right. \\
 &\left. \|X_\varepsilon^x\|_\infty < N \right] \\
 (4.13) \quad &+ \mathbb{P}\left[\left\|\sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon\left(\frac{\lfloor sl \rfloor}{l}, X_\varepsilon^{x,l}(s), \mathcal{L}_{\frac{\lfloor sl \rfloor}{l}}^{X_\varepsilon^x}\right) dW(s)\right\|_\alpha \geq \frac{\rho}{2}, \right. \\
 &\left. \|\sqrt{\varepsilon} W\|_\infty \leq \delta \right].
 \end{aligned}$$

We analyze each term in the RHS separately. First, consider the term (4.11). We denote

$$\eta_\varepsilon = \sup_{s,y,\mu} \{|b(s, y, \mu) - b_\varepsilon(s, y, \mu)|, |\sigma(s, y, \mu) - \sigma_\varepsilon(s, y, \mu)|\}.$$

By uniform convergence of b_ε to b and σ_ε to σ , we have that $\lim_{\varepsilon \rightarrow 0} \eta_\varepsilon = 0$. We choose ε small enough so that $\eta_\varepsilon \leq \frac{L\gamma}{4}$. Then

$$\begin{aligned}
 (4.11) \leq &\mathbb{P}\left[\left\|\int_0^\cdot \frac{2[\sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - \sigma_\varepsilon(\frac{\lfloor sl \rfloor}{l}, X_\varepsilon^{x,l}(s), \mathcal{L}_{\frac{\lfloor sl \rfloor}{l}}^{X_\varepsilon^x})]}{L\gamma} dW(s)\right\|_\alpha \right. \\
 &\geq \frac{\rho}{\sqrt{\varepsilon}L\gamma}, \frac{2\|\sigma_\varepsilon(\cdot, X_\varepsilon^x(\cdot), \mathcal{L}_\cdot^{X_\varepsilon^x}) - \sigma_\varepsilon(\frac{\lfloor \cdot \rfloor}{l}, X_\varepsilon^{x,l}(\cdot), \mathcal{L}_\cdot^{X_\varepsilon^{x,l}})\|_\infty}{L\gamma} \leq 1 \Big] \\
 &\leq C' \exp\left(\frac{-\rho^2}{C'L^2\gamma^2\varepsilon}\right),
 \end{aligned}$$

using Lemma A.5. Thus choose γ such that $\frac{\rho^2}{C'L^2R} \geq \gamma^2$.

Second, consider the term (4.12). We take ε small enough so that $\eta_\varepsilon < 1$. Applying Itô's formula to $|X_\varepsilon^x(t)|^2$ gives

$$\begin{aligned}
 |X_\varepsilon^x(t)|^2 &= |x|^2 + 2\sqrt{\varepsilon} \int_0^t \langle X_\varepsilon^x(s), \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \rangle \\
 &\quad + 2 \int_0^t \langle X_\varepsilon^x(s), b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) \rangle ds \\
 &\quad + \varepsilon \int_0^t \text{Tr}(\sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x})^T) ds.
 \end{aligned}$$

Following the estimation methods used to prove Theorem 3.3, we have

$$\mathbb{E}[\|X_\varepsilon^x\|_\infty^2] \leq K(|x|^2 + \mathbb{E}[\|b(\cdot, 0, \delta_0)\|_\infty^2]) \exp(K + \mathbb{E}[\|b(\cdot, 0, \delta_0)\|_\infty^2]) < \infty.$$

In the same way, we can additionally prove $\mathbb{E}[\|X_\varepsilon^x\|_\infty^{2q}] < \infty$. Let $j = \lfloor tl \rfloor$. We can rewrite our SDE, for $t \in [j/l, (j + 1)/l]$, as

$$X_\varepsilon^x(t) - X_\varepsilon^x(j/l) = \sqrt{\varepsilon} \int_{j/l}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) + \int_{j/l}^t b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) ds.$$

We evaluate the strong error term in the same way as above to see that

$$\begin{aligned} \sup_{t \in [0,1]} |X_\varepsilon^x(t) - X_\varepsilon^{x,l}(t)|^2 &\leq 2 \sup_{t \in [j/l, (j+1)/l]} \left| \int_{j/l}^t b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) ds \right|^2 \\ &\quad + 2 \sup_{t \in [j/l, (j+1)/l]} \left| \sqrt{\varepsilon} \int_{j/l}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right|^2. \end{aligned}$$

Taking expectations yields

$$\begin{aligned} &\mathbb{E}[\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty^2] \\ &\leq 2\mathbb{E}\left[\sup_{t \in [j/l, (j+1)/l]} \left| \int_{j/l}^t b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) ds \right|^2 \right] \\ &\quad + 2\mathbb{E}\left[\sup_{t \in [j/l, (j+1)/l]} \left| \sqrt{\varepsilon} \int_{j/l}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right|^2 \right] \\ &\leq \frac{1}{l^2} (4\eta_\varepsilon + 32L^2\mathbb{E}[\|X_\varepsilon^x\|_\infty^{2q}] + 8L^2 4\mathbb{E}[\|X_\varepsilon^x\|_\infty^2] + 4\|b(\cdot, 0, \delta_0)\|_\infty^2) \\ &\quad + \frac{8\varepsilon M^2}{l} \\ &\lesssim \frac{1}{l}, \end{aligned}$$

and we write $\mathbb{E}[\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty^2]^{1/2} \leq K_1/\sqrt{l}$. In the same way we also have that

$$\begin{aligned} &\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty \\ &= \sup_{j=0, \dots, l-1} \sup_{t \in [j/l, (j+1)/l]} \left| \sqrt{\varepsilon} \int_{j/l}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(r) \right| \\ &\quad + \sup_{j=0, \dots, l-1} \sup_{t \in [j/l, (j+1)/l]} \left| \int_{j/l}^t b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dr \right| \\ &\leq \sup_{j,t} \left| \sqrt{\varepsilon} \int_{j/l}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right| \end{aligned}$$

$$\begin{aligned}
 & + \sup_{j,t} \int_{\frac{j}{l}}^t \left[\eta_\varepsilon + L(1 + \|X\|_\infty^q) + L\mathbb{E}[\|X\|_\infty^2]^{1/2} + \sup_{r \in [0,1]} |b(r, 0, \delta_0)| \right] ds \\
 & \leq \sup_{j,t} \left| \sqrt{\varepsilon} \int_{\frac{j}{l}}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right| + K_2 \frac{1 + \|X_\varepsilon^x\|_\infty^q}{l}.
 \end{aligned}$$

Hence we have for term (4.12) that

$$\begin{aligned}
 & \mathbb{P} \left[\frac{1}{l^\beta} + \|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty + \mathbb{E}[\|X_\varepsilon^x - X_\varepsilon^{x,l}\|_\infty^2]^{1/2} > \gamma, \|X_\varepsilon^x\|_\infty < N \right] \\
 & \leq \mathbb{P} \left[\sup_{j,t} \left| \sqrt{\varepsilon} \int_{\frac{j}{l}}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right| \right. \\
 & \quad \left. + K_2 \frac{1 + \|X_\varepsilon^x\|_\infty^q}{l} + \frac{K_1}{\sqrt{l}} + \frac{1}{l^\beta} > \gamma, \|X_\varepsilon^x\|_\infty < N \right] \\
 & \leq \mathbb{P} \left[\sup_{j,t} \left| \sqrt{\varepsilon} \int_{\frac{j}{l}}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right| > \gamma - \frac{K_3}{l^{\frac{1}{2} \wedge \beta}} \right],
 \end{aligned}$$

where $K_3 = K_1 + K_2(1 + N^q) + 1$.

Therefore, using Chernoff’s inequality,

$$\begin{aligned}
 (4.12) & \leq \mathbb{P} \left[\sup_{j,t} \exp \left(\lambda \sup_{j,t} \left| \sqrt{\varepsilon} \int_{\frac{j}{l}}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right| \right) \right] \\
 & > \exp \left(\frac{\lambda}{l^{\frac{1}{2} \wedge \beta}} (\gamma l^{\frac{1}{2} \wedge \beta} - K_3) \right) \\
 & \leq \frac{\sup_{j,t} \mathbb{E}[\exp(\lambda |\sqrt{\varepsilon} \int_{\frac{j}{l}}^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s)|) \mathbb{1}_{\|X_\varepsilon^x\|_\infty < N}]}{\exp(-\frac{\lambda}{l^{\frac{1}{2} \wedge \beta}} (\gamma l^{\frac{1}{2} \wedge \beta} - K_3))} \\
 & \lesssim \exp \left(\lambda^2 \varepsilon \frac{M}{2l} - \frac{\lambda}{l^{\frac{1}{2} \wedge \beta}} (\gamma l^{\frac{1}{2} \wedge \beta} - K_3) \right) \\
 & \lesssim \exp \left(\frac{-(\gamma l^{\frac{1}{2} \wedge \beta} - K_3)^2}{2\varepsilon M} \frac{l}{l^{\frac{1}{2} \wedge \beta}} \right),
 \end{aligned}$$

by optimizing over the arbitrary choice of λ . We can now choose the constant l such that $\frac{(\gamma l^{\frac{1}{2} \wedge \beta} - K_3)^2}{2M} l^{1 - (1 \wedge 2\beta)} > R$.

Finally, to evaluate equation (4.13), we first consider $\sigma_\varepsilon(X_\varepsilon^{x,l}(\cdot), \mathcal{L}_{\lfloor \frac{\cdot}{l} \rfloor}^{X_\varepsilon^x})$. This process is constant over the interval $(\frac{j}{l}, \frac{j+1}{l}]$. Then taking the Hölder norm we get

$$\begin{aligned} & \left\| \int_0^\cdot \sigma_\varepsilon\left(\frac{\lfloor ls \rfloor}{l}, X_\varepsilon^{x,l}(s), \mathcal{L}_{\lfloor \frac{ls}{l} \rfloor}^{X_\varepsilon^x}\right) dW(s) \right\|_\alpha \\ &= \left\| \sum_{j=0}^{\lfloor l-1 \rfloor} \sigma_\varepsilon\left(\frac{j}{l}, X_\varepsilon^{x,l}\left(\frac{j}{l}\right), \mathcal{L}_{\frac{j}{l}}^{X_\varepsilon^x}\right) \left[W\left(\frac{j+1}{l} \wedge \cdot\right) - W\left(\frac{j}{l} \wedge \cdot\right) \right] \right\|_\alpha \\ &\leq 2lM \|W\|_\alpha, \end{aligned}$$

using $\|\sigma_\varepsilon\|_\infty \leq M$ and the triangle inequality. Then

$$\begin{aligned} & \mathbb{P}\left[\left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^{x,l}(s), \mathcal{L}_s^{X_\varepsilon^{x,l}}) dW(s) \right\|_\alpha \geq \frac{\rho}{2}, \|\sqrt{\varepsilon}W\|_\infty \leq \delta\right] \\ &\leq \mathbb{P}\left[\|W\|_\alpha \geq \frac{\rho}{4\sqrt{\varepsilon}lM}, \|W\|_\infty \leq \frac{\delta}{\sqrt{\varepsilon}}\right] \\ &\leq C \max\left(1, \left(\frac{\rho}{4Ml\delta}\right)^{1/\alpha}\right) \exp\left(\frac{-1}{\varepsilon} \frac{1}{C} \left(\frac{\rho}{4Ml\delta}\right)^{1/\alpha} \delta^2\right), \end{aligned}$$

where we applied Lemma A.4 and chose δ such that $\frac{\rho}{R^\alpha 4MlC^\alpha} \geq \delta^{1-2\alpha}$.

Injecting these three results in (4.10) gives us that

$$\begin{aligned} & \mathbb{P}\left[\left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_{\lfloor \frac{ls}{l} \rfloor}^{X_\varepsilon^x}) dW(s) \right\|_\alpha \geq \rho, \right. \\ (4.14) \quad & \left. \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X_\varepsilon^x\|_\infty < N\right] \\ & \lesssim \exp\left(-\frac{R}{\varepsilon}\right). \end{aligned}$$

Step 2. The Hölder norm of the whole process when $h = 0$. We have

$$(4.15) \quad \begin{aligned} & \|X_\varepsilon^x - \Phi^x(0)\|_{\alpha,t} \\ & \leq \left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_{\alpha,t} \end{aligned}$$

$$(4.16) \quad + \left\| \int_0^\cdot [b_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - b(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x})] ds \right\|_{\alpha,t}$$

$$(4.17) \quad + \left\| \int_0^\cdot [b(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - b(s, \Phi^x(0)(s), \delta_{\Phi^x(0)(s)})] ds \right\|_{\alpha,t}.$$

Equation (4.15) is the term in (4.14) that we desire. Equation (4.16) is bounded above by η_ε . We only consider the cases when $\|X_\varepsilon^x\|_\infty, \|\Phi^x(0)\|_\infty < N$ since we know that $\Phi^x(0)(t)$ remains in this ball and we conditioned on $X_\varepsilon^x(t)$ remaining

in the same ball. This means that by the locally Lipschitz condition, we can say that $b(t, x, \mu)$ is Lipschitz in the spacial variable with constant L_N . Therefore, for (4.17) we have

$$\begin{aligned} & \left\| \int_0^\cdot b(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - b(s, \Phi^x(0)(s), \delta_{\Phi^x(0)(s)}) ds \right\|_{\alpha,t} \\ & \leq \sup_{p,q \in [0,t]} \frac{\int_p^q |b(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) - b(s, \Phi^x(0)(s), \delta_{\Phi^x(0)(s)})|}{|q - p|^\alpha} \\ & \leq \sup_{p,q \in [0,t]} \frac{L_N}{|q - p|^\alpha} \int_p^q |X_\varepsilon^x(s) - \Phi^x(0)(s)| ds \\ & \quad + \frac{L}{|q - p|^\alpha} \int_p^q \mathbb{E}[|X_\varepsilon^x(s) - \Phi^x(0)(s)|^2]^{1/2} ds \end{aligned} \tag{4.18}$$

$$\begin{aligned} & \leq L_N \|X_\varepsilon^x(\cdot) - \Phi^x(0)(\cdot)\|_{\infty,t} + L_N \int_0^t \|X_\varepsilon^x(\cdot) - \Phi^x(0)(\cdot)\|_{\alpha,s} ds \\ & \tag{4.19} + L \mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_\infty^2]^{1/2}, \end{aligned}$$

using Lemma 2.3.

Next, we want to show that the strong error $\mathbb{E}[\|X_\varepsilon^x - \Phi(0)\|_\infty^2]$ can be controlled by ε . Using that

$$\begin{aligned} d(X_\varepsilon^x - \Phi^x(0))(t) &= \sigma_\varepsilon(t, X_\varepsilon^x(t), \mathcal{L}_t^{X_\varepsilon^x}) dW(t) \\ & \quad + (b_\varepsilon(t, X_\varepsilon^x(t), \mathcal{L}_t^{X_\varepsilon^x}) - b(t, X_\varepsilon^x(t), \mathcal{L}_t^{X_\varepsilon^x})) dt \\ & \quad + (b(t, X_\varepsilon^x(t), \mathcal{L}_t^{X_\varepsilon^x}) - b(t, \Phi^x(0)(t), \delta_{\Phi^x(0)(t)})) dt, \end{aligned}$$

and Itô's formula for $f(x) = |x|^2$ with $X_\varepsilon^x(0) - \Phi^x(0)(0) = 0$ gives that

$$\begin{aligned} & \|X_\varepsilon^x - \Phi^x(0)\|_{\infty,t}^2 \\ & \leq 2\sqrt{\varepsilon} \sup_{0 \leq s \leq t} \left| \int_0^s \langle X_\varepsilon^x(r) - \Phi^x(0)(r), \sigma_\varepsilon(r, X_\varepsilon^x(r), \mathcal{L}_r^{X_\varepsilon^x}) dW(r) \rangle \right| \\ & \quad + \varepsilon M^2 t d + \int_0^t 2\eta_\varepsilon |X_\varepsilon^x(r) - \Phi^x(0)(r)| dr \\ & \quad + 2 \int_0^t |\langle X_\varepsilon^x(r) - \Phi^x(0)(r), b(r, X_\varepsilon^x(r), \mathcal{L}_r^{X_\varepsilon^x}) \\ & \quad - b(r, \Phi^x(0)(r), \delta_{\Phi^x(0)(r)}) \rangle| dr. \end{aligned}$$

Squaring and taking expectations give

$$\begin{aligned} &\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,t}^4] \\ &\leq 64pM^2\varepsilon \int_0^t (\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,r}^4] + 1) dr + 4\varepsilon^2M^4t^2 \\ &\quad + 16t\eta_\varepsilon \int_0^t (\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,r}^4] + 1) dr \\ &\quad + 16t \int_0^t 4L^2\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,r}^4] dr. \end{aligned}$$

Refining, we then obtain $\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_\infty^2]^{1/2} \leq K(\eta_\varepsilon^{1/4} \vee \varepsilon^{1/4})e^K$. We have shown that this expectation is of order $\varepsilon^{1/4}$. Now we consider $\|X_\varepsilon^x - \Phi^x(0)\|_\infty$. Since the supremum norm can be made to appear inside the integrals, we have

$$\begin{aligned} &\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,t} \\ &\leq \left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_{\infty,t} + \eta_\varepsilon t \\ &\quad + \int_0^t L_N \|X_\varepsilon^x - \Phi(0)\|_{\infty,r} dr + Lt\mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_\infty^2]^{1/2}, \end{aligned}$$

and by using Grönwall, we get

$$\begin{aligned} &\|X_\varepsilon^x - \Phi^x(0)\|_{\infty,t} \\ &\leq \left(\left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_{\infty,t} \right. \\ &\quad \left. + (\eta_\varepsilon + \mathbb{E}[\|X_\varepsilon^x - \Phi^x(0)\|_\infty^2]^{1/2}) \right) e^{L_N t} \\ &\leq \left(\left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_{\alpha,t} \right. \\ &\quad \left. + (\varepsilon^{1/4} \vee \eta_\varepsilon^{1/4})K' \right) e^{K't}. \end{aligned} \tag{4.20}$$

Combining equation (4.18), equation (4.19) and equation (4.20) give

$$\begin{aligned} \|X_\varepsilon^x - \Phi^x(0)\|_\alpha &\leq \left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_\alpha (1 + L_N e^{K'}) e^{L_N t} \\ &\quad + (\eta_\varepsilon^{1/4} \vee \varepsilon^{1/4})(1 + L_N K' e^{K'} + KLe^K) e^{L_N t} \\ &\leq \left[\left\| \sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s) \right\|_\alpha + (\eta_\varepsilon^{1/4} \vee \varepsilon^{1/4}) \right] K_4. \end{aligned}$$

Thus for any choice of ρ we see that

$$\begin{aligned} &\mathbb{P}[\|X_\varepsilon^x - \Phi^x(0)\|_\alpha \geq \rho, \|X_\varepsilon^x\|_\infty < N] \\ &\leq \mathbb{P}\left[\left(\left\|\sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s)\right\|_\alpha + (\eta_\varepsilon^{1/4} \vee \varepsilon^{1/4})\right) K_4 \geq \rho, \right. \\ &\quad \left. \|X_\varepsilon^x\|_\infty < N\right] \end{aligned}$$

and by choosing ε small enough such that $(\eta_\varepsilon^{1/4} \vee \varepsilon^{1/4}) < \frac{\rho}{2K_4}$ we get

$$\begin{aligned} &\mathbb{P}[\|X_\varepsilon^x - \Phi^x(0)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X_\varepsilon^x\|_\infty < N] \\ &\leq \mathbb{P}\left[\left\|\sqrt{\varepsilon} \int_0^\cdot \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathcal{L}_s^{X_\varepsilon^x}) dW(s)\right\|_\alpha \geq \frac{\rho}{2K_4}, \right. \\ &\quad \left. \|\sqrt{\varepsilon}W\|_\infty \leq \delta, \|X_\varepsilon^x\|_\infty < N\right] \\ &\lesssim \exp\left(-\frac{R}{\varepsilon}\right), \end{aligned}$$

since in equation (4.14) the choice of ρ is arbitrary.

Step 3. The case when $h \neq 0$. For the final step, we use the same method as in [3] to extend the results to Wiener processes with drift. Using a Girsanov transformation, we have that there is a measure $\tilde{\mathbb{P}}$ absolutely continuous to the standard probability measure \mathbb{P} .

Note that the law of the stochastic process is not changed by perturbing the path of the Brownian motion by some element of the Cameron–Martin space. When solving a McKean–Vlasov equation (unlike classical SDEs), one has to fix the law of the probability space in order to define $\mathcal{L}^X = \mathbb{P} \circ X^{-1}$. Hence the law is not changed when one considers a different driving noise for the SDE. This is most obvious in expression (4.9) where the delta distribution follows the path of the skeleton with input $h = 0$.

We rewrite the SDE and skeleton process

$$\begin{aligned} X_\varepsilon^x(t) &= x + \int_0^t b_\varepsilon(s, X_\varepsilon^x(s), \mathbb{P} \circ [X_\varepsilon^x(s)]^{-1}) ds \\ &\quad + \int_0^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathbb{P} \circ [X_\varepsilon^x(s)]^{-1}) \dot{h}(s) ds \\ &\quad + \sqrt{\varepsilon} \int_0^t \sigma_\varepsilon(s, X_\varepsilon^x(s), \mathbb{P} \circ [X_\varepsilon^x(s)]^{-1}) d\tilde{W}(s), \\ \Phi^x(h)(t) &= x + \int_0^t b(s, \Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) ds \\ &\quad + \int_0^t \sigma(s, \Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) \dot{h}(s) ds, \end{aligned}$$

where $\tilde{W} = W - h/\sqrt{\varepsilon}$, $\tilde{\mathbb{P}}$ is the measure where \tilde{W} is a Brownian motion and $\mathbb{P} \circ [X_\varepsilon^x(t)]^{-1} = \mathcal{L}_t^{X_\varepsilon^x}$. The drift term $b_\varepsilon + \sigma_\varepsilon \dot{h}$ satisfies the properties from before and matches the skeleton process $\Phi^x(h)$.

Also note that

$$W^{(2)}(\mathbb{P} \circ [X_\varepsilon^x(t)]^{-1}, \delta_{\Phi^x(0)(t)}) = \mathbb{E}^{\mathbb{P}}[|X_\varepsilon^x(t) - \Phi(0)(t)|^2]^{\frac{1}{2}},$$

which we have already showed to go to 0 as $\varepsilon \rightarrow 0$. Thus we argue in the same way as in Step 2 and conclude

$$\begin{aligned} &\mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] \\ &\quad \lesssim \tilde{\mathbb{P}}[\|\tilde{X}_\varepsilon^x - \Phi^x(0)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}\tilde{W}\|_\infty \leq \delta] \\ &\quad \lesssim \exp\left(-\frac{R}{\varepsilon}\right). \end{aligned} \quad \square$$

We are now in position to prove our second main result, Theorem 4.8.

PROOF OF THEOREM 4.8. *Proving the upper bound.* First, consider the case where $0 \notin A$ and A is closed in the Hölder topology. Then there exists a r such that $\Delta(A) > r > 0$. Let us consider the ball in the Cameron–Martin space H

$$\left\{ h \in H^{\otimes d'} : h(t) = \int_0^t \dot{h}(s) ds, \frac{\|\dot{h}\|_2^2}{2} \leq r \right\}.$$

Recall that if $h \in H^{\otimes d'}$ then $h \in C^{\frac{1}{2}}([0, 1]; \mathbb{R}^m)$ and is bounded and, moreover, that $\|h\|_\infty \leq \|h\|_{\frac{1}{2}} \leq \|\dot{h}\|_2$. Therefore, we can apply the Arzelà–Ascoli theorem [19] to get that this set is compact. Hence we can find a finite open cover of this set and we can restrict the radius of the open balls. We write

$$\left\{ h \in H^{\otimes d'} : h(t) = \int_0^t \dot{h}(s) ds, \frac{\|\dot{h}\|_2^2}{2} \leq r \right\} \subset \bigcup_{i=1}^N B_\infty(h_i, \delta_{h_i}) = U.$$

These balls are in the uniform topology and the elements h_i are all have $\|\dot{h}\|_2^2/2 < r$. By this property, $\Phi(h_i) \notin A$. If it were, $\|\dot{h}\|_2^2/2 > \Delta(A)$. The set A is closed in the Hölder topology so A^c is open in the Hölder topology. Therefore, there exists a ρ_{h_i} such that in the Hölder topology $B_\alpha(\Phi(h_i), \rho_{h_i})$ is in A^c and, therefore, does not intersect with A . Hence when $X_\varepsilon^x \in A$, we can say that $\|X_\varepsilon^x - \Phi^x(h_i)\|_\alpha \geq \rho_{h_i}$. Finally, we can estimate

$$\begin{aligned} &\mathbb{P}[X_\varepsilon^x \in A] \\ &\quad = \mathbb{P}[X_\varepsilon^x \in A, \sqrt{\varepsilon}W \notin U] + \mathbb{P}[X_\varepsilon^x \in A, \sqrt{\varepsilon}W \in U] \\ &\quad \leq \mathbb{P}[\sqrt{\varepsilon}W \notin U] + \mathbb{P}[X_\varepsilon^x \in A, \sqrt{\varepsilon}W \in U] \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{P}[\sqrt{\varepsilon}W \notin U] + \sum_{i=1}^N \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h_i)\|_\alpha \geq \rho_{h_i}, \|\sqrt{\varepsilon}W - h_i\|_\infty \leq \delta_{h_i}] \\ &\leq \mathbb{P}[\sqrt{\varepsilon}W \notin U] + N \exp\left(-\frac{2r}{\varepsilon}\right), \end{aligned}$$

where for the last line we apply Proposition 4.9, with δ_{h_i} and ε chosen sufficiently small for the given ρ_{h_i} . The δ_{h_i} are dependent on our choice of open cover for the compact set, so we can make them as small as required. We already have a large deviation principle for a Wiener process on the uniform norm by [26]. Hence we have for ε sufficiently small that

$$\mathbb{P}[\sqrt{\varepsilon}W \notin U] \leq \exp\left(-\frac{\Delta(U^c)}{\varepsilon}\right).$$

If $h \notin U$, then we have that $\frac{\|\dot{h}\|_2^2}{2} > r$ and consequently $\mathbb{P}[\sqrt{\varepsilon}W \notin U] \leq \exp(-2r/\varepsilon)$. Combining all of this together, we get

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[X_\varepsilon^x \in A]) \leq -r,$$

where r was chosen arbitrarily so that $r < \Delta(A)$ where A is closed. We optimize for our choice of r and get

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[X_\varepsilon^x \in A]) \leq -\Delta(A),$$

which is the upper inequality for the theorem.

Proving the lower bound. Now consider A to be an open set in the Hölder topology and let $h \in H^{\otimes d'}$ such that $\Phi^x(h) \in A$. There exists a $\rho > 0$ such that the Hölder ball $B_\alpha(\Phi^x(h), \rho) \subset A$. Also we have that

$$\begin{aligned} &\mathbb{P}[\|\sqrt{\varepsilon}W - h\|_\infty < \delta] \\ &\leq \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] + \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha < \rho]. \end{aligned}$$

Hence

$$\begin{aligned} &\mathbb{P}[X_\varepsilon^x \in A] \\ &\geq \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha < \rho] \\ &\geq \mathbb{P}[\|\sqrt{\varepsilon}W - h\|_\infty < \delta] - \mathbb{P}[\|X_\varepsilon^x - \Phi^x(h)\|_\alpha \geq \rho, \|\sqrt{\varepsilon}W - h\|_\infty \leq \delta] \\ &\geq \mathbb{P}[\|\sqrt{\varepsilon}W - h\|_\infty < \delta] - \exp\left(-\frac{R}{\varepsilon}\right). \end{aligned}$$

Applying the LDP for the Brownian motion (see Lemma A.2) and using that $\frac{\|\dot{h}\|_2^2}{2} \geq \Delta(B_\infty(h, \delta))$, we see that

$$\mathbb{P}[\|\sqrt{\varepsilon}W - h\|_\infty < \delta] \geq \exp\left(-\frac{\|\dot{h}\|_2^2}{2\varepsilon}\right)$$

and hence

$$\mathbb{P}[X_\varepsilon^x \in A] \geq \exp\left(-\frac{\|\dot{h}\|_2^2}{2\varepsilon}\right) - \exp\left(-\frac{R}{\varepsilon}\right),$$

where we can choose R to take any value. Choosing $R = \|\dot{h}\|_2^2$ and rearranging, we get

$$\mathbb{P}[X_\varepsilon^x \in A] \geq \exp\left(-\frac{\|\dot{h}\|_2^2}{2\varepsilon}\right) \left(1 - \exp\left(-\frac{\|\dot{h}\|_2^2}{2\varepsilon}\right)\right).$$

Hence

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[X_\varepsilon^x \in A]) \geq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log\left(1 - \exp\left(-\frac{\|\dot{h}\|_2^2}{2\varepsilon}\right)\right) - \frac{\|\dot{h}\|_2^2}{2}.$$

The limit goes to 0 for any choice of $h \in H^{\otimes d}$. Finally, as h was arbitrarily chosen in A , we take the infimum over h and get

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log(\mathbb{P}[X_\varepsilon^x \in A]) \geq -\Delta(A).$$

This completes the proof of the theorem. \square

5. Functional iterated logarithm law. Strassen’s law, or the law of iterated logarithm describes the magnitude of the fluctuations of a Brownian motion. It was first proved in [36]. Observe that for a Brownian motion $W(t)$, we have that $X_n^{(1)}(t) = W(nt)/n \rightarrow 0$ as $n \rightarrow \infty$ both in probability and almost surely. However, $X_n^{(2)}(t) = W(nt)/\sqrt{n}$ is also a Brownian motion for any choice of n . Therefore, something is happening between n and \sqrt{n} which is turning a stochastic process into a deterministic constant in the limit as $n \rightarrow \infty$. Strassen’s law says that

$$X_n^{(3)}(t) = \frac{W(nt)}{\sqrt{n \log \log(n)}},$$

converges to 0 in probability but does not converge almost surely. In particular,

$$\limsup_{n \rightarrow \infty} X_n^{(3)}(1) = \sqrt{2}, \quad \text{almost surely.}$$

In this section, we are interested in studying whether stochastic processes have a similar type of property. We will consider the solution of the SDE run over a large time interval of order n and rescaled to order $\sqrt{n \log(\log(n))}$. Similar to the proof of Strassen’s law, we will show that the set of rescaled paths is relatively compact in the Hölder topology but that the set of limit points of this set is uncountable which implies the failure of almost sure convergence.

In [1], Baldi proves a law of iterated logarithm for classical SDEs for the uniform topology. This was then extended in [20] and later [21] to other coarser

pathspace topologies. Standard LDP results easily give us convergence in probability. We calculate the set of possible limit points of the scaled diffusions which for a classical SDE are

$$\{\Phi^x(h) : d\Phi^x(h)(t) = b(\Phi^x(h)(t)) dt + \sigma(\Phi^x(h)(t))\dot{h}(t) dt, \\ \Phi^x(h)(0) = x \text{ and } \|\dot{h}\|_2 \leq \sqrt{2}\}.$$

We show below, that similarly for a McKean–Vlasov SDE these are

$$\{\Phi^x(h) : d\Phi^x(h)(t) = b(\Phi^x(h)(t), \delta_{\Phi^x(0)(t)}) dt + \sigma(\Phi^x(h)(t), \delta_{\Phi^x(0)(t)})\dot{h}(t) dt, \\ \Phi^x(h)(0) = x \text{ and } \|\dot{h}\|_2 \leq \sqrt{2}\}.$$

We will follow the methods of [1], [20] and [21] to extend the LDP results to prove an iterated logarithm law for the class of McKean–Vlasov SDEs in Theorem 3.3. It seems possible to use microscopic rescaling of the Brownian motion such as in [23] to provide an alternative proof of our result, however, we do not pursue this point.

REMARK 5.1 (Decoupling argument). In this section, we are unable to use a decoupling argument as highlighted in Section 4.2.2.

To the best of our knowledge, there are no results proving a Strassen-type law for SDEs with coefficients which can vary in time and we were unable to establish any such results while working on this paper. The conditions that we require on the measure dependency are similar to those of the spacial dependency and do not naturally translate into conditions for time dependency. Therefore, proving that they are satisfied is much easier in the MV-SDE setting when they are written as properties on the measure dependency than for some general time dependent coefficient.

Functional iterated logarithm law for McKean–Vlasov SDEs. First, we need to define in what sense we will be rescaling our MV-SDE.

DEFINITION 5.2. Let $\alpha \in \mathbb{R}^+$. A family of continuous bijections $\Gamma_\alpha : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is said to be a system of contractions centered at x if:

1. $\Gamma_\alpha(x) = x$ for every $\alpha \in \mathbb{R}^+$.
2. If $\alpha \geq \beta$, then $|\Gamma_\alpha(y_1) - \Gamma_\alpha(y_2) - \Gamma_\alpha(z_1) + \Gamma_\alpha(z_2)| \leq |\Gamma_\beta(y_1) - \Gamma_\beta(y_2) - \Gamma_\beta(z_1) + \Gamma_\beta(z_2)|$ for every $y_1, y_2, z_1, z_2 \in \mathbb{R}^d$.
3. Γ_1 is the identity and $(\Gamma_\alpha)^{-1} = \Gamma_{\alpha^{-1}}$.
4. For every compact set $\mathcal{K} \subset C^\alpha([0, 1]; \mathbb{R}^d)$, $f \in \mathcal{K}$ and $\varepsilon > 0$, $\exists \delta > 0$, such that $|pq - 1| < \delta$ implies

$$\|\Gamma_p \circ \Gamma_q(f) - f\|_\alpha < \sqrt{\varepsilon}, \quad \forall p, q \in \mathbb{R}^+.$$

The simplest example of such a system of contractions is $\Gamma_\alpha(y) = \frac{y}{\alpha}$ centered at $x = 0$. Indeed this is the specific operator used when proving Strassen’s law for Brownian motion. Also note that we only really care about Γ_α for $\alpha > 1$. It is clear that for $\alpha < 1$, the operators Γ_α will not be contraction operators.

EXAMPLE 5.3. In fact, a linear contraction operator with a transformation will satisfy these conditions. Consider, for example, $\Gamma_\alpha(y) = \frac{(y-x)}{\alpha} + x$ and naturally, $\Gamma_\alpha(x) = x$. Similarly, for $\alpha \geq \beta$,

$$\begin{aligned} &\Gamma_\alpha(y_1) - \Gamma_\alpha(y_2) - \Gamma_\alpha(z_1) + \Gamma_\alpha(z_2) \\ &= \frac{y_1 - x}{\alpha} + x - \frac{y_2 - x}{\alpha} - x - \frac{z_1 - x}{\alpha} - x + \frac{z_2 - x}{\alpha} + x \\ &\leq \frac{y_1 - y_2 - z_1 + z_2}{\beta} = \Gamma_\beta(y_1) - \Gamma_\beta(y_2) - \Gamma_\beta(z_1) + \Gamma_\beta(z_2). \end{aligned}$$

Finally, for $|pq - 1| < \delta$, we have

$$\begin{aligned} \|\Gamma_p \circ \Gamma_q(f) - f\|_\alpha &= \sup_{s,t \in [0,1]} \frac{|\Gamma_p \circ \Gamma_q(f(t)) - f(t) - \Gamma_p \circ \Gamma_q(f(s)) + f(s)|}{|t - s|^\alpha} \\ &= \sup_{s,t \in [0,1]} \frac{|[\frac{f(t)}{pq} - f(t)] - [\frac{f(s)}{pq} - f(s)]|}{|t - s|^\alpha} \\ &= \left| \frac{1}{pq} - 1 \right| \sup_{s,t \in [0,1]} \frac{|f(t) - f(s)|}{|t - s|^\alpha} \leq \frac{\delta}{2} \|f\|_\alpha. \end{aligned}$$

These conditions are slightly stronger than those of [1] and are used in [21]. Condition 2. in Definition 5.2 needs to be strengthened to allow it to be applied to Hölder norms rather than just supremum norms. Observe that by choosing $y_2 = z_2 = x$, one gets

$$|\Gamma_\alpha(y_1) - \Gamma_\alpha(z_1)| \leq |\Gamma_\beta(y_1) - \Gamma_\beta(z_1)|.$$

This stronger condition still allows for the example of linear contractions up to a transformation. For $s \in \mathbb{R}^+$, define

$$\phi(s) = \sqrt{s \log(\log(s))}.$$

Let $b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times m}$ be progressively measurable functions such that there is a unique solution to

$$dY(t) = b(Y(t), \mathcal{L}_t^Y) dt + \sigma(Y(t), \mathcal{L}_t^Y) dW(t), \quad Y(0) = x \in \mathbb{R}^d.$$

DEFINITION 5.4. Let $u > 3$. Let $\hat{\sigma}_u : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d'}$ and $\hat{b}_u : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ be such that

$$\begin{aligned} \hat{\sigma}_u(y, \mu) &= \phi(u) \nabla[\Gamma_{\phi(u)}](\Gamma_{\phi(u)^{-1}}(y))^T \sigma(\Gamma_{\phi(u)^{-1}}(y), \mu \circ \Gamma_{\phi(u)}), \\ \hat{b}_u(y, \mu) &= u \mathbf{L}(y, \mu)[\Gamma_{\phi(u)}](\Gamma_{\phi(u)^{-1}}(y)), \end{aligned}$$

where for $\tilde{a} = \sigma^T \sigma$ the operator $\mathbf{L}(\cdot, \cdot)[\cdot]$ is given as

$$\begin{aligned} \mathbf{L}(y, \mu)[f](z) &= \sum_{i=1}^d \frac{\partial f}{\partial y_i}(\Gamma_{\phi(u)^{-1}}(z)) b_i(\Gamma_{\phi(u)^{-1}}(y), \mu \circ \Gamma_{\phi(u)}) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \tilde{a}_{i,j}(\Gamma_{\phi(u)^{-1}}(y), \mu \circ \Gamma_{\phi(u)}) \frac{\partial^2 f}{\partial y_i \partial y_j}(\Gamma_{\phi(u)^{-1}}(z)). \end{aligned}$$

ASSUMPTION 5.5. Throughout, we assume that Γ_u is twice differentiable for all $u > 3$ and that $\forall y \in \mathbb{R}^d, \forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$ we have for some $\hat{\sigma} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d'}$ and $\hat{b} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$,

$$\lim_{u \rightarrow \infty} \hat{\sigma}_u(y, \mu) = \hat{\sigma}(y, \mu) \quad \text{and} \quad \lim_{u \rightarrow \infty} \hat{b}_u(y, \mu) = \hat{b}(y, \mu),$$

where $\hat{\sigma}$ and \hat{b} satisfy Assumption 3.2 with the addition that $\hat{\sigma}$ is bounded by constant M .

For $t, u \in \mathbb{R}^+$ define

$$Z_u(t) = \Gamma_{\phi(u)}(Y(ut)),$$

and recall that since $Y(0) = x$ and $\Gamma_u(x) = x$, by assumption $Z_u(0) = x$. We use Itô's formula on $Z_u(t)$ by assuming twice differentiability of $\Gamma_{\phi(u)}(\cdot)$,

$$\begin{aligned} dZ_u(t) &= d(\Gamma_{\phi(u)}(Y(ut))) \\ &= \nabla[\Gamma_{\phi(u)}](Y(ut))^T dY(ut) + \frac{dY(ut)^T}{2} H[\Gamma_{\phi(u)}](Y(ut)) dY(ut). \end{aligned}$$

Rewriting $Y(ut) = \Gamma_{\phi(u)^{-1}}(Z_u(t))$ and substituting in gives

$$\begin{aligned} dZ_u(t) &= u \sum_{i=1}^d \frac{\partial \Gamma_{\phi(u)}}{\partial y_i}(\Gamma_{\phi(u)^{-1}}(Z_u(t))) b_i(\Gamma_{\phi(u)^{-1}}(Z_u(t)), \mathcal{L}_t^{Z_u}) \circ \Gamma_{\phi(u)} dt \\ &\quad + \frac{u}{2} \sum_{i,j=1}^d \frac{\partial^2 \Gamma_{\phi(u)}}{\partial y_i \partial y_j}(\Gamma_{\phi(u)^{-1}}(Z_u(t))) \\ &\quad \times \sum_{k=1}^{d'} \sigma_{k,i} \sigma_{j,k}(\Gamma_{\phi(u)^{-1}}(Z_u(t)), \mathcal{L}_t^{Z_u}) \circ \Gamma_{\phi(u)} dt \end{aligned}$$

$$\begin{aligned}
 &+ \sum_{i=1}^d \frac{\partial \Gamma_{\Phi(u)}}{\partial y_i} (\Gamma_{\phi(u)^{-1}}(Z_u(t))) \\
 &\times \sum_{k=1}^{d'} \sigma_{i,k} (\Gamma_{\phi(u)^{-1}}(Z_u(t)), \mathcal{L}_t^{Z_u}) \circ \Gamma_{\phi(u)} dW_k(ut).
 \end{aligned}$$

Next, using that $\mathcal{W}_u(t) = \frac{W(ut)}{\sqrt{u}}$ is a Brownian motion, we can rewrite all of this as the SDE with initial condition $Z_u(0) = x$,

$$dZ_u(t) = \frac{1}{\sqrt{\log \log(u)}} \hat{\sigma}_u(Z_u(t), \mathcal{L}_t^{Z_u}) d\mathcal{W}_u(t) + \hat{b}_u(Z_u(t), \mathcal{L}_t^{Z_u}) dt.$$

Under Assumption 5.5 and using Theorem 4.8, we get

$$\begin{aligned}
 (5.1) \quad -\Delta(\mathring{A}) &\leq \liminf_{u \rightarrow \infty} \frac{1}{\log \log(u)} \log \mathbb{P}(Z_u \in A) \\
 &\leq \limsup_{u \rightarrow \infty} \frac{1}{\log \log(u)} \log \mathbb{P}(Z_u \in A) \leq -\Delta(\bar{A})
 \end{aligned}$$

for every Borel set A induced by the α -Hölder topology with $\alpha < 1/2$. Recall the definition of the rate function $\Delta(A) := \inf\{\|\dot{h}\|_2^2/4; h \in H^{\otimes d'}, \Phi^x(h)(\cdot) \in A\}$ with the skeleton process,

$$\Phi^x(h)(t) = x + \int_0^t \hat{b}(\Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) ds + \int_0^t \hat{\sigma}(\Phi^x(h)(s), \delta_{\Phi^x(0)(s)}) \dot{h}(s) ds.$$

We can now state the main result of this section.

THEOREM 5.6. *With probability 1, the set of paths $\{Z_u; u > 3\}$ is relatively compact on the Hölder topology and its set of limit points coincides with $K = \{\Phi(h) : \frac{\|\dot{h}\|_2^2}{2} \leq 1\}$.*

We first prove some technical lemmas.

LEMMA 5.7. *$\forall c > 1$ and $\forall \varepsilon > 0$ there exists a positive integer $j_0(\omega)$ almost surely finite such that $\forall j > j_0$,*

$$d_\alpha(Z_{c^j}, K) < \sqrt{\varepsilon}, \quad \text{where } d_\alpha(x, A) = \inf\{\|x - y\|_\alpha : y \in A\}.$$

PROOF. Start by considering the set of α -Hölder continuous paths $C_\varepsilon := \{g; d_\alpha(g, K) \geq \sqrt{\varepsilon}\}$. By definition, we have that $\Delta(C_\varepsilon) > 1$, so there exists a real number $\delta > 0$ such that $\Delta(C_\varepsilon) > 1 + \delta$. Using the LDP results in (5.1), we can rearrange this to get

$$\mathbb{P}[Z_{c^j} \in C_\varepsilon] \leq \exp(-(1 + \delta) \log \log(c^j)) \lesssim \frac{1}{j^{1+\delta}}.$$

Clearly, $\sum_{j=1}^\infty \mathbb{P}[Z_{c^j} \in C_\varepsilon] < \infty$ and by a direct application of Borel–Cantelli we have $\mathbb{P}[d_\alpha(Z_{c^j}, K) > \sqrt{\varepsilon} \text{ i.o.}] = 0$. \square

LEMMA 5.8. $\forall \varepsilon > 0 \exists c_\varepsilon > 1$ such that for $1 < c < c_\varepsilon$, there exists an almost surely finite integer $j_0(\omega)$ such that $\forall j > j_0, A_{j,c} \leq \sqrt{\varepsilon}$.

PROOF. For notational convenience define, for $c > 1$ and for every positive integer j , the quantity

$$A_{j,c} = \sup_{c^{j-1} \leq u \leq c^j} \|Z_u - \Gamma_{\phi(u)} \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j})\|_\alpha.$$

Start by observing that the set K is relatively compact in the α -topology, so it is bounded. Therefore, by Lemma 5.7, we have that $\forall j > j_0$ that $\|Z_{c^j}\|_\alpha < C$. We want to show that

$$\sum_{j \geq 1} \mathbb{P}[A_{c,j} > \sqrt{\varepsilon}] < \infty$$

$$\text{which is equivalent to } \sum_{j > j_0} \mathbb{P}[A_{c,j} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\alpha < C] < \infty.$$

Considering one of these sets, we see

$$\begin{aligned} & \{A_{j,c} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\infty \leq C\} \\ &= \left\{ \sup_{c^{j-1} \leq u \leq c^j} \sup_{0 \leq s, t \leq 1} (|\Gamma_{\phi(u)}(Y(ut)) - \Gamma_{\phi(u)}(Y(c^j t)) \right. \\ & \quad \left. - \Gamma_{\phi(u)}(Y(us) + \Gamma_{\phi(u)}(Y(c^j s)))|) / (|t - s|^\alpha) > \sqrt{\varepsilon}, \right. \\ & \quad \left. \|Z_{c^j}\|_\alpha < C \right\}. \end{aligned}$$

Using Definition 5.2, for $u \in [c^{j-1}, c^{j+1}]$,

$$\begin{aligned} & |\Gamma_{\phi(u)}(Y(ut)) - \Gamma_{\phi(u)}(Y(c^j t)) - \Gamma_{\phi(u)}(Y(us) + \Gamma_{\phi(u)}(Y(c^j s)))| \\ & \leq |\Gamma_{\phi(c^{j-1})}(Y(ut)) - \Gamma_{\phi(c^{j-1})}(Y(c^j t)) - \Gamma_{\phi(c^{j-1})}(Y(us) \\ & \quad + \Gamma_{\phi(c^{j-1})}(Y(c^j s)))|. \end{aligned}$$

Therefore,

$$\begin{aligned} & \{A_{c,j} > \sqrt{\varepsilon}\} \\ & \subseteq \left\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s, t \leq 1} \frac{1}{|t - s|^\alpha} |\Gamma_{\phi(c^{j-1})}(Y(c^j vt)) - \Gamma_{\phi(c^{j-1})}(Y(c^j t)) \right. \end{aligned}$$

$$\begin{aligned}
 & - \Gamma_{\phi(c^{j-1})}(Y(c^j vs) + \Gamma_{\phi(c^{j-1})}(Y(c^j s)))| > \sqrt{\varepsilon} \Big\} \\
 \subseteq & \left\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s, t \leq 1} \frac{|Z_{c^j}(vt) - Z_{c^j}(t) - Z_{c^j}(vs) + Z_{c^j}(s)|}{|t - s|^\alpha} > \frac{\sqrt{\varepsilon}}{2} \right\},
 \end{aligned}$$

using that $\exists j$ large enough so that for and $\delta > 0$,

$$\frac{\phi(c^{j-1})}{\phi(c^j)} = \frac{1}{\sqrt{c}} \sqrt{\frac{\log \log(c^j)}{\log \log(c^{j-1})}} \leq \frac{1}{\sqrt{c}}(1 - \delta),$$

and choosing c small enough we can make $\Gamma_{\frac{\phi(c^{j-1})}{\phi(c^j)}}$ within $\frac{\sqrt{\varepsilon}}{2}$ of the identity operator using properties from Definition 5.2. Therefore,

$$\begin{aligned}
 & \{A_{c,j} > \sqrt{\varepsilon}, \|Z_{c^j}\|_\infty \leq C\} \\
 \subseteq & \left\{ \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s, t \leq 1} \frac{|Z_{c^j}(vt) - Z_{c^j}(t) - Z_{c^j}(vs) + Z_{c^j}(s)|}{|t - s|^\alpha} > \frac{\sqrt{\varepsilon}}{2}, \right. \\
 & \left. \|Z_{c^j}\|_\alpha \leq C \right\} \\
 \subseteq & \{Z_{c^j} \in b_\varepsilon\},
 \end{aligned}$$

where the set b_ε is given by

$$\begin{aligned}
 b_\varepsilon = & \left\{ g \in C^\alpha([0, 1]; \mathbb{R}^d) : \sup_{\frac{1}{c} \leq v \leq 1} \sup_{0 \leq s, t \leq 1} \frac{|g(vt) - g(t) - g(vs) + g(s)|}{|t - s|^\alpha} > \frac{\sqrt{\varepsilon}}{2}, \right. \\
 & \left. \|g\|_\alpha \leq C \right\},
 \end{aligned}$$

as we would expect. Let $h \in H^{\otimes d'}$ so $\|\dot{h}\|_2 < \infty$ such that $\Phi(h) \in b_\varepsilon$, then

$$\begin{aligned}
 (5.2) \quad \frac{\sqrt{\varepsilon}}{2} |t - s|^\alpha & \leq |[\Phi(h)(t) - \Phi(h)(vt)] - [\Phi(h)(s) - \Phi(h)(sv)]| \\
 & \leq \left| \int_{(vt) \vee s}^t d\Phi(h)(r) - \int_{vs}^{s \wedge (tv)} d\Phi(h)(r) \right|,
 \end{aligned}$$

for at least some choice of $v \in [\frac{1}{c}, 1]$ and $t, s \in [0, 1]$.

We know that a solution to the ODE $\Phi(h)$ exists uniquely and has finite supremum. Therefore, we can easily conclude that there exists constants M_1 and M_2 such that

$$\begin{aligned}
 \left| \int_s^t d\Phi(h)(r) \right| & \leq \left| \int_s^t b(\Phi(h)(r), \delta_{\Phi(h)(r)}) dr + \int_s^t \sigma(\Phi(h)(r), \delta_{\Phi(h)(r)}) dh(r) \right| \\
 & \leq M_1 \sqrt{|t - s|} \|\dot{h}\|_2 + M_2 |t - s|.
 \end{aligned}$$

It follows from (5.2) that

$$\|\dot{h}\|_2 \geq \frac{\frac{\sqrt{\varepsilon}}{2}|t-s|^\alpha - M_2(|t-s \vee (tv)| + |s \wedge (tv) - (sv)|)}{M_1(|t-s \vee (tv)|^{\frac{1}{2}} + |s \wedge (tv) - (sv)|^{\frac{1}{2}})}.$$

Let us consider first the case where $s < (tv)$.

$$\begin{aligned} \|\dot{h}\|_2 &\geq \frac{\frac{\sqrt{\varepsilon}}{2}|t-s|^\alpha - M_2|(t+s)(1-v)|}{M_1|(\sqrt{t} + \sqrt{s})\sqrt{1-v}|} \\ &\geq \frac{\sqrt{\varepsilon}|1-\frac{1}{c}|^\alpha}{4M_1|\sqrt{1-\frac{1}{c}}|} - \frac{M_2}{M_1}\sqrt{1-\frac{1}{c}}, \end{aligned}$$

so for c small enough we have $\|\dot{h}\|_2 \geq 1 + \delta$ for any choice of $\delta > 0$.

Second, consider the case where $s > (tv)$,

$$\begin{aligned} \|\dot{h}\|_2 &\geq \frac{\frac{\sqrt{\varepsilon}}{2}|t(1-\frac{s}{t})|^\alpha - M_2(|t(1-\frac{s}{t})|(1+v))}{2M_1(|t(1-\frac{s}{t})|^{\frac{1}{2}})} \\ &\geq \frac{\sqrt{\varepsilon}|1-\frac{1}{c}|^\alpha}{4M_1|\sqrt{1-\frac{1}{c}}|} - \frac{M_2}{M_1}\sqrt{1-\frac{1}{c}}, \end{aligned}$$

and taking $c > 1$ small enough as before gives $\|\dot{h}\|_2 \geq 1 + 2\delta$.

Therefore, using equation (5.1) we can get

$$\begin{aligned} \mathbb{P}[Z_{c^j} \in b_\varepsilon] &\leq \exp(-(\Delta(b_\varepsilon) - \delta) \log \log(c^j)) \\ &\leq \exp(-(1 + \delta) \log \log(c^j)) \lesssim \frac{1}{j^{1+\delta}}, \end{aligned}$$

and the conclusion of the proof is straightforward by Borel–Cantelli. \square

We are now able to prove the main theorem.

PROOF OF THEOREM 5.6. The proof is divided into two parts:

Step 1. Relative Compactness. For any $c > 1$, there will exist $j \in \mathbb{N}$ such that $c^{j-1} < u < c^j$,

$$\begin{aligned} (5.3) \quad d_\alpha(Z_u, K) &\leq d_\alpha(Z_{c^j}, K) \\ (5.4) \quad &+ \|\Gamma\phi(u) \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j}) - Z_{c^j}\|_\alpha \\ &+ \|Z_u - \Gamma\phi(u) \circ \Gamma_{\phi(c^j)^{-1}}(Z_{c^j})\|_\alpha, \end{aligned}$$

where j is chosen so that $c^{j-1} \leq u \leq c^j$.

Lemma 5.7 with j large enough ensures that (5.3) is bounded by $\frac{\sqrt{\varepsilon}}{3}$. From Lemma 5.7, we have that Z_{c^j} is bounded, since $\forall \delta > 0$,

$$1 \geq \frac{\phi(u)}{\phi(c^j)} \geq \frac{\phi(c^{j-1})}{\phi(c^j)} \geq \frac{(1 - \delta)}{\sqrt{c}},$$

for j large enough. Choosing $1 < c$ small enough, we can use the fourth part of Definition 5.2 to get that the first term in (5.4) is less than $\frac{\sqrt{\varepsilon}}{3}$. Lemma 5.8 bounds the second term of (5.4) by $\frac{\sqrt{\varepsilon}}{3}$.

Therefore, we conclude that the set $\{Z_u : u > 3\}$ is relatively compact (and hence we have convergence in probability).

Step 2. The set of limit points. Let $\Phi(h) \in K$ so that $\frac{\|h\|_2^2}{2} < 1$. Then for $\varepsilon > 0$ and $\beta > 0$, we define the sets:

$$E_j = \left\{ \left\| \frac{\mathcal{W}_{c^j}(t)}{\sqrt{\log \log(c^j)}} - h \right\|_\infty \leq \beta \right\} \quad \text{and} \quad F_j = \{ \|Z_{c^j} - \Phi(h)\|_\alpha \leq \sqrt{\varepsilon} \}.$$

Using Proposition 4.9, we have that for j large enough and α small enough that

$$(5.5) \quad \mathbb{P}[E_j] - \mathbb{P}[F_j] = \mathbb{P}[E_j \cap F_j^c] \leq \exp(-2 \log \log(c^j)) \lesssim \frac{1}{j^2}.$$

Strassen’s law tells us that $\mathbb{P}(\limsup E_j) = 1$; see [36]. Therefore, $\sum_j \mathbb{P}[E_j] = \infty$. However, by equation (5.5) we also have

$$\begin{aligned} \sum_j (\mathbb{P}[E_j] - \mathbb{P}[F_j]) < \infty &\Rightarrow \sum_j \mathbb{P}[F_j] = \infty \\ &\Rightarrow \mathbb{P}[\|Z_{c^j} - \Phi(h)\|_\alpha < \sqrt{\varepsilon} \text{ i.o.}] = 1, \end{aligned}$$

the latter following from Borel–Cantelli.

Finally, since $(c_{j \in \mathbb{N}}^j)$ is just a subsequence of $(m)_{m \in \mathbb{N}}$, the result can be extended to the conclusion. \square

APPENDIX: A COLLECTION OF AUXILIARY RESULTS

A.1. Classical large deviation principles. The following lemma corresponds to [16], Lemma 5.6.18.

LEMMA A.1. *Let $(b_t)_t, (\sigma_t)_t$ be progressively measurable processes. Let*

$$(A.1) \quad dz_t = b_t dt + \sqrt{\varepsilon} \sigma_t dW(t) \quad \text{where } z_0 \text{ is deterministic.}$$

Let $\tau \in [0, 1]$ be a stopping time with respect to the filtration of $(W(t))_{t \in [0, 1]}$. Suppose that the coefficients of the diffusion matrix are uniformly bounded, and for some constants M, B, ρ and any $t \in [0, \tau_1]$,

$$|\sigma_t| \leq M(|z_t|^2 + \rho^2)^{\frac{1}{2}}, \quad \text{and} \quad |b_t| \leq B(|z_t|^2 + \rho^2)^{\frac{1}{2}}.$$

Then, for any $\delta > 0$ and any $\varepsilon \leq 1$,

$$\varepsilon \log\left(\mathbb{P}\left[\sup_{[0, \tau_1]} |z_t| \geq \delta\right]\right) \leq B + M^2\left(1 + \frac{d}{2}\right) + \log\left(\frac{\rho^2 + |z_0|^2}{\rho^2 + \delta^2}\right).$$

In particular, if $z_0 = 0$:

$$\varepsilon \log\left(\mathbb{P}\left[\sup_{[0, \tau_1]} |z_t| \geq \delta\right]\right) \leq B + M^2\left(1 + \frac{d}{2}\right) + \log\left(\frac{\rho^2}{\rho^2 + \delta^2}\right).$$

We also need [16], Lemma 5.2.1.

LEMMA A.2. For any dimension d' , and any $\tau, \varepsilon, \delta$,

$$\mathbb{P}\left[\sup_{0 \leq t \leq \tau} |\sqrt{\varepsilon}W(t)| \geq \delta\right] \leq 4d' \exp\left(-\frac{\delta^2}{2d\tau\varepsilon}\right).$$

We also need [16], Theorem 4.2.23.

PROPOSITION A.3. Let $\{\mu_\varepsilon\}$ be a family of probability measures that satisfies the LDP with a good rate function I on a Hausdorff topological space \mathcal{X} , and for $m \in \mathbb{N}$, let f_m be continuous functions from \mathcal{X} to \mathcal{Y} , where (\mathcal{Y}, d) is a metric space. Assume there exists a measurable map f from \mathcal{X} to \mathcal{Y} such that for every $\alpha < \infty$:

$$\limsup_{m \rightarrow \infty} \sup_{x: I(x) \leq \alpha} d(f_m(x), f(x)) = 0.$$

Then any family of probability measures $\{\tilde{\mu}_\varepsilon\}$ for which $\{\mu_\varepsilon \circ f_m^{-1}\}$ are exponentially good approximations satisfies the LDP in \mathcal{Y} with the good rate function $I'(y) := \inf\{I(x) : y = f(x)\}$.

A.2. Large deviation principles in path space topologies. The following results are of their own independent interest and can be found in [3], Lemme 1, page 196, with token proofs. We provide a full proof for the benefit of the reader. The extension to $[0, T]$ is straightforward.

LEMMA A.4 ([3]). Let $(W(t))_{t \in [0, 1]}$ be a d' -dimensional Brownian motion. Then there exists a constant $C > 0$ which is independent of m such that $\forall u, v > 0$,

$$\mathbb{P}[\|W\|_\alpha \geq u, \|W\|_\infty \leq v] \leq C \max\left(1, \left(\frac{u}{v}\right)^{1/\alpha}\right) \exp\left(\frac{-1}{C} \frac{u^{1/\alpha}}{v^{(1/\alpha)-2}}\right).$$

PROOF. Consider a Brownian motion W satisfying the constraint $\|W\|_\infty \leq v$. We use methods from [26] to represent the α -Hölder norm in terms of a supremum

of Fourier coefficients generated by Schauder functions. By direct calculation, one can dominate the Fourier coefficients

$$|W_{pm}| = 2^{p/2} \left| 2W\left(\frac{2m-1}{2^{p+1}}\right) - W\left(\frac{m}{2^p}\right) - W\left(\frac{m-1}{2^p}\right) \right| \leq 2^{p/2} 4v,$$

if we also restrict that $\|W\|_\alpha = \sup_{p,m} |W_{pm}| 2^{p(\alpha-1/2)} \geq u$ and search for values of p and m which do not yield a contradiction. Observe that we require $u \geq 4v2^{\alpha p}$. If we consider a p where this was not true, we would have that $W_{pm} < u$. The supremum of all W_{pm} is still be greater than u , but this value of p could be removed from the collection over which the supremum is taken over without affecting the measure of the event. Let p_0 be the least such relevant p , defined as $p_0 := \inf\{p \in \mathbb{N}; 2^{\alpha p} \geq u/(4v)\}$. Then for an arbitrary choice of $\lambda > 0$, we have

$$\begin{aligned} &\mathbb{P}[\|W\|_\alpha \geq u, \|W\|_\infty \leq v] \\ &= \mathbb{P}\left[\sup_{p \geq p_0, m} 2^{p(\alpha-1/2)} |W_{pm}| \geq u\right] \\ &\leq \frac{\sup_{p \geq p_0} \mathbb{E}[\exp(\lambda 2^{p(\alpha-1/2)} |W_{pm}|)]}{\exp(\lambda u)} \\ &\leq \sup_{p \geq p_0} 2 \exp\left(\frac{\lambda^2 2^{p(2\alpha-1)}}{2} - \lambda u\right) \\ &\leq 2 \exp\left(\frac{-u^2 2^{p_0(1-2\alpha)}}{2}\right), \end{aligned}$$

where for the last line we choose $\lambda = u2^{p(1-2\alpha)}$ to minimize the expression (since λ is arbitrary). From the definition of p_0 , we have

$$2^{p_0(1-2\alpha)} \geq \left(\frac{u}{4v}\right)^{\frac{1}{\alpha}-2},$$

and substituting this in yields the final result. \square

The next lemma iterates on the first; see [3], Lemme 2, page 196.

LEMMA A.5. *Let $(W(t))_{t \in [0,1]}$ be a d' -dimensional Brownian motion. There exists a constant $C' > 0$ which is independent of d' and α such that $\forall u > 0$ and $\forall K \in C([0, 1])$ such that $\|K\|_\infty \leq 1$,*

$$\mathbb{P}\left[\left\|\int_0^\cdot K(s) dW(s)\right\|_\alpha \geq u, \|K\|_\infty \leq 1\right] \leq C' \exp\left(\frac{-u^2}{C'}\right).$$

PROOF. Let $\|K\|_\infty \leq 1$. In the case where K is deterministic, the stochastic integral of K is clearly normally distributed and the result is clear. For K not deterministic, it hard to say anything about the probability distribution of the stochastic integral.

Using the equivalent definition of Hölder norms in terms of a Schauder expansion, we have that

$$\begin{aligned} & \mathbb{P}\left[\left\|\int_0^1 K(s) dW(s)\right\|_\alpha \geq u, \|K\|_\infty \leq 1\right] \\ &= \mathbb{P}\left[\sup_{p,m} \left|\int_0^1 H_{pm}(s)K(s) dW(s)\right| \geq u, \|K\|_\infty \leq 1\right] \\ &\leq \frac{\mathbb{E}[\exp(\lambda \sup_{p,m} |\int_0^1 H_{pm}(s)K(s) dW(s)|)]}{\exp(\lambda u)} \\ &\leq \sup_{p,m} \frac{\mathbb{E}[\exp(\lambda |\int_0^1 H_{pm}(s)K(s) dW(s)|)]}{\exp(\lambda u)}, \end{aligned}$$

where the supremum can come outside the expectation by the Beppo–Levi theorem since the random variables are all positive. Temporarily, consider the process $Y(t) = \int_0^t H_{pm}(s)K(s) dW(s)$. Using Itô’s formula, we get that

$$\mathbb{E}[|Y(t)|^n] = \int_0^t \mathbb{E}\left[\frac{n(n-1)}{2} |Y(s)|^{n-2} H_{pm}(s)^2 K(s)^2\right] ds.$$

$Y(t)$ is a martingale, since H_{pm} and K are bounded, with $Y(0) = 0$ so $\mathbb{E}[Y(t)] = 0$. By the Itô isometry, the second moment of $Y(t)$ is

$$\begin{aligned} \mathbb{E}[Y(t)^2] &= \mathbb{E}\left[\int_0^t H_{pm}(s)^2 K(s)^2 ds\right] \\ &\leq \begin{cases} 0, & 0 \leq t \leq \frac{m-1}{2^p}, \\ 2^p \left(t - \frac{m-1}{2^p}\right), & \frac{m-1}{2^p} < t < \frac{m}{2^p}, \\ 1, & \frac{m}{2^p} \leq t \leq 1. \end{cases} \end{aligned}$$

Therefore, by induction on n we see

$$\mathbb{E}[Y(t)^{2n}] \leq \begin{cases} 0, & 0 \leq t \leq \frac{m-1}{2^p}, \\ \frac{(2n)!}{n!2^n} (2^p)^n \left(t - \frac{m-1}{2^p}\right)^n, & \frac{m-1}{2^p} < t < \frac{m}{2^p}, \\ \frac{(2n)!}{n!2^n}, & \frac{m}{2^p} \leq t \leq 1. \end{cases}$$

For the odd moments of $|Y(t)|$, we first use the Burkholder–Davies–Gundy inequality to say

$$\begin{aligned} \mathbb{E}[|Y(t)|] &\leq C_1 \mathbb{E} \left[\left(\int_0^t H_{pm}(s)^2 K(s)^2 ds \right)^{\frac{1}{2}} \right] \\ &\leq \begin{cases} 0, & 0 \leq t \leq \frac{m-1}{2^p}, \\ C_1 2^{p/2} \left(t - \frac{m-1}{2^p} \right)^{1/2}, & \frac{m-1}{2^p} < t < \frac{m}{2^p}, \\ C_1, & \frac{m}{2^p} \leq t \leq 1, \end{cases} \end{aligned}$$

and by induction on n again we see that

$$\mathbb{E}[|Y(t)|^{2n+1}] \leq \begin{cases} 0, & 0 \leq t \leq \frac{m-1}{2^p}, \\ C_1 n! 2^n \left(t - \frac{m-1}{2^p} \right)^{\frac{2n+1}{2}} 2^{\frac{2n+1}{2}}, & \frac{m-1}{2^p} < t < \frac{m}{2^p}, \\ C_1 n! 2^n, & \frac{m}{2^p} \leq t \leq 1. \end{cases}$$

Hence $\mathbb{E}[|Y(t)|^{2n}] \leq (C_1 \vee 1) \frac{(2n)!}{n! 2^n}$ and $\mathbb{E}[|Y(t)|^{2n+1}] \leq (C_1 \vee 1) n! 2^n$. The upper bounds for these moments are the same as the moments of a half normal distribution with variance 1 up to a multiplicative constant. Therefore, we can upper bound the moment generating function of the RV $|Y(1)|$ using the moment generating function of a half normal random variable. If Z is half normally distributed with variance a , we have

$$\mathbb{E}[\exp(\lambda Z)] = \int_0^\infty \frac{2}{\sqrt{2\pi a^2}} e^{\lambda x} \exp\left(\frac{-x^2}{2a^2}\right) dx \leq 4 \exp\left(\frac{\lambda a^2}{2}\right).$$

Therefore, $\mathbb{E}[\exp(\lambda |\int_0^1 H_{pm}(s) K(s) dW(s)|)] \lesssim \exp(\frac{\lambda^2}{2})$, and hence

$$\mathbb{P} \left[\left\| \int_0^\cdot K(s) dW(s) \right\|_\alpha \geq u, \|K\|_\infty \leq 1 \right] \lesssim \exp\left(\frac{\lambda^2}{2} - \lambda u\right) \lesssim \exp\left(\frac{-u^2}{2}\right),$$

by choosing λ to minimize the equation since the choice of λ was arbitrary ($\lambda = u$). □

LEMMA A.6. *Let $\psi \in C^\alpha([0, 1])$ with $\psi(0) = 0$. Then $\|\psi\|_\infty \leq \|\psi\|_\alpha$.*

PROOF. Using that $t \in [0, 1]$, one easily computes

$$\begin{aligned} \|\psi\|_\infty &= \sup_{t \in [0, 1]} |\psi(t)| = \sup_{t \in [0, 1]} \frac{|\psi(t) - \psi(0)| \cdot |t - 0|}{|t - 0|} \\ &\leq \sup_{t \in [0, 1]} \frac{|\psi(t) - \psi(0)|}{|t - 0|} \cdot \sup_{t \in [0, 1]} |t - 0| \leq \|\psi\|_\alpha. \end{aligned} \quad \square$$

Acknowledgements. The authors thank the referee for an in-depth review of the initial version of our manuscript which lead to several improvements.

REFERENCES

- [1] BALDI, P. (1986). Large deviations and functional iterated logarithm law for diffusion processes. *Probab. Theory Related Fields* **71** 435–453. [MR0824713](#)
- [2] BALDI, P., BEN AROUS, G. and KERKYACHARIAN, G. (1992). Large deviations and the Strassen theorem in Hölder norm. *Stochastic Process. Appl.* **42** 171–180. [MR1172514](#)
- [3] BEN AROUS, G. and LEDOUX, M. (1994). Grandes déviations de Freidlin–Wentzell en norme hölderienne. In *Séminaire de Probabilités, XXVIII. Lecture Notes in Math.* **1583** 293–299. Springer, Berlin. [MR1329119](#)
- [4] BENACHOUR, S., ROYNETTE, B., TALAY, D. and VALLOIS, P. (1998). Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* **75** 173–201. [MR1632193](#)
- [5] BUCKDAHN, R., LI, J., PENG, S., RAINER, C. et al. (2017). Mean-field stochastic differential equations and associated PDEs. *Ann. Probab.* **45** 824–878. [MR3630288](#)
- [6] BUDHIRAJA, A., DUPUIS, P. and FISCHER, M. (2012). Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* **40** 74–102. [MR2917767](#)
- [7] CARMONA, R. (2016). *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications. Financial Mathematics* **1**. SIAM, Philadelphia, PA. [MR3629171](#)
- [8] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. I. Probability Theory and Stochastic Modelling* **83**. Springer, Cham. [MR3752669](#)
- [9] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. II. Probability Theory and Stochastic Modelling* **84**. Springer, Cham. [MR3753660](#)
- [10] CATTIAUX, P., GUILLIN, A. and MALRIEU, F. (2008). Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** 19–40. [MR2357669](#)
- [11] COLLET, F., DAI PRA, P. and SARTORI, E. (2010). A simple mean field model for social interactions: Dynamics, fluctuations, criticality. *J. Stat. Phys.* **139** 820–858. [MR2639892](#)
- [12] CRISAN, D. and MCMURRAY, E. (2018). Smoothing properties of McKean–Vlasov SDEs. *Probab. Theory Related Fields* **171** 97–148. [MR3800831](#)
- [13] CRISAN, D. and XIONG, J. (2010). Approximate McKean–Vlasov representations for a class of SPDEs. *Stochastics* **82** 53–68. [MR2677539](#)
- [14] DAI PRA, P. and DEN HOLLANDER, F. (1996). McKean–Vlasov limit for interacting random processes in random media. *J. Stat. Phys.* **84** 735–772. [MR1400186](#)
- [15] DAWSON, D. A. and GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. [MR0885876](#)
- [16] DEMBO, A. and ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability* **38**. Springer, Berlin. [MR2571413](#)
- [17] DEUSCHEL, J.-D., FRIZ, P. K., MAURELLI, M. and SLOWIK, M. (2018). The enhanced Sanov theorem and propagation of chaos. *Stochastic Process. Appl.* **128** 2228–2269. [MR3804792](#)
- [18] DOS REIS, G., SMITH, G. and TANKOV, P. (2018). Importance sampling for McKean–Vlasov SDEs. ArXiv preprint. Available at [ArXiv:1803.09320](#).

- [19] DUNFORD, N. and SCHWARTZ, J. T. (1988). *Linear Operators. Part I*. Wiley, New York. [MR1009162](#)
- [20] EDDAHBI, M. (2000). Large deviation estimates and functional law for diffusions in modulus spaces. *Random Oper. Stoch. Equ.* **8** 373–396. [MR1797027](#)
- [21] EDDAHBI, M. and N’ZI, M. (2002). Strassen’s local law for diffusion processes under strong topologies. *Acta Math. Vietnam.* **27** 151–163. [MR1939953](#)
- [22] FREIDLIN, M. I. and WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer, Heidelberg. Translated from the 1979 Russian original by Joseph Szücs. [MR2953753](#)
- [23] GANTERT, N. (1993). An inversion of Strassen’s law of the iterated logarithm for small time. *Ann. Probab.* **21** 1045–1049. [MR1217579](#)
- [24] GÄRTNER, J. (1988). On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137** 197–248. [MR0968996](#)
- [25] GUASONI, P. and ROBERTSON, S. (2008). Optimal importance sampling with explicit formulas in continuous time. *Finance Stoch.* **12** 1–19. [MR2362149](#)
- [26] HERRMANN, S., IMKELLER, P., PAVLYUKOVICH, I. and PEITHMANN, D. (2014). *Stochastic Resonance: A Mathematical Approach in the Small Noise Limit. Mathematical Surveys and Monographs* **194**. Amer. Math. Soc., Providence, RI. [MR3155413](#)
- [27] HERRMANN, S., IMKELLER, P. and PEITHMANN, D. (2008). Large deviations and a Kramers’ type law for self-stabilizing diffusions. *Ann. Appl. Probab.* **18** 1379–1423. [MR2434175](#)
- [28] MAO, X. (2008). *Stochastic Differential Equations and Applications*, 2nd ed. Horwood Publishing Limited, Chichester. [MR2380366](#)
- [29] MCKEAN, H. P. JR. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. [MR0221595](#)
- [30] MCKEAN, H. P. JR. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)* 41–57. Air Force Office Sci. Res., Arlington, VA. [MR0233437](#)
- [31] MÉLÉARD, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995). Lecture Notes in Math.* **1627** 42–95. Springer, Berlin. [MR1431299](#)
- [32] MISHURA, Y. S. and VERETENNIKOV, A. Y. (2016). Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. ArXiv preprint. Available at [ArXiv:1603.02212](#).
- [33] ØKSENDAL, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*. Springer, Berlin.
- [34] POPE, S. B. (2000). *Turbulent Flows*. Cambridge Univ. Press, Cambridge.
- [35] SCHEUTZOW, M. (1987). Uniqueness and nonuniqueness of solutions of Vlasov–McKean equations. *J. Aust. Math. Soc. A* **43** 246–256. [MR0896631](#)
- [36] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3** 211–226. [MR0175194](#)
- [37] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D’Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. [MR1108185](#)
- [38] VILLANI, C. (2009). *Optimal Transport. Old and New. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Springer, Berlin. [MR2459454](#)

G. DOS REIS
W. SALKELD
SCHOOL OF MATHEMATICS
UNIVERSITY OF EDINBURGH
PETER GUTHRIE TAIT ROAD
EDINBURGH, EH9 3FD
UNITED KINGDOM
E-MAIL: G.dosReis@ed.ac.uk
w.j.salkeld@sms.ed.ac.uk

J. TUGAUT
INSTITUT CAMILLE JORDAN
UNIVERSITÉ JEAN MONNET
23 RUE DU DOCTEUR PAUL MICHELON
42023 SAINT-ÉTIENNE
FRANCE
E-MAIL: tugaut@math.cnrs.fr
URL: <http://tugaut.perso.math.cnrs.fr>