DETERMINISTIC HOMOGENIZATION FOR DISCRETE-TIME FAST-SLOW SYSTEMS UNDER OPTIMAL MOMENT ASSUMPTIONS

ILYA CHEVYREV, PETER K. FRIZ, ALEXEY KOREPANOV, IAN MELBOURNE, AND HUILIN ZHANG

ABSTRACT. We consider discrete-time fast-slow systems of the form

\[
X^{(n)}_{k+1} = X^{(n)}_k + n^{-1} a_n(X^{(n)}_k, Y^{(n)}_k) + n^{-1/2} b_n(X^{(n)}_k, Y^{(n)}_k), \quad Y^{(n)}_{k+1} = T_n Y^{(n)}_k.
\]

We give conditions under which the dynamics of the slow equations converge weakly to an Itô diffusion as \(n \to \infty\). The drift and diffusion coefficients of the limiting stochastic differential equation satisfied by \(X\) are given explicitly. This extends the results of [Kelly–Melbourne, J. Funct. Anal. 272 (2017) 4063–4102] from the continuous-time case to the discrete-time case. Moreover, our methods (\(p\)-variation rough paths) work under optimal moment assumptions.

1. INTRODUCTION

In this article, we are primarily concerned with homogenization of deterministic, discrete-time, fast-slow systems of the form

\[
(1) \quad X^{(n)}_{k+1} = X^{(n)}_k + n^{-1} a_n(X^{(n)}_k, Y^{(n)}_k) + n^{-1/2} b_n(X^{(n)}_k, Y^{(n)}_k), \quad Y^{(n)}_{k+1} = T_n Y^{(n)}_k,
\]

where \(X^{(n)}_k\) takes values in \(\mathbb{R}^d\), \(Y^{(n)}_k\) takes values in a metric space \(M\), and \(a_n, b_n : \mathbb{R}^d \times M \to \mathbb{R}^d\) and \(T_n : M \to M\) are suitable functions. The only source of randomness in the dynamics is the initial condition \(Y^{(n)}_0\) which we sample from a (not necessarily ergodic) probability measure \(\lambda_n\) on \(M\).

Our main result, Theorem 2.17, provides sufficient conditions for the dynamics \(x_n(t) = X^{(n)}_{[nt]}\) to converge in law (which we write in symbols as \(x_n \to \lambda_n X\)), with respect to the uniform topology, to the solution of a stochastic differential equation (SDE)

\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB
\]

with explicit formulae for the coefficients \(\tilde{a}, \sigma\). Our assumptions on the system involve only moment bounds and a suitable (iterated) weak invariance principle on the fast dynamics \(T_n\). See Section 1.1 for an illustrative example of a system to which our results apply.

The programme to study homogenization of deterministic systems of the form (1) was initiated in [18], and has seen recent growth in a number of works, including [11, 12, 13, 14, 15]. See our survey paper [3] for an overview. The contribution of this article is three-fold. The first two of these contributions are novel even when we suppose that \(a_n \equiv a, b_n \equiv b, T_n \equiv T\) are independent of \(n\). First, we are able to deal with discrete-time dynamics in the same way as continuous-time dynamics. This

2010 Mathematics Subject Classification. Primary 60H10; Secondary 37A50.

Key words and phrases. Fast-slow systems, homogenization, \(p\)-variation, rough paths.
should be compared to \[13,12\] in which results for discrete-time dynamics are only obtained in the special case \(a(x, y) = a(x), b(x, y) = b(x)v(y)\) and the case of general \(a, b\) is only handled for continuous-time dynamics in \[14\]. Second, we are able to work under optimal moment assumptions, and our results apply to the full range of systems in which one expects a weak invariance principle to hold for the fast dynamics. This extends (even for continuous-time dynamics) the results of \[13,14,11\] in which only a subrange can be handled (see Remark 2.4). In \[13\] we indicate a simplified version of the second of these contributions for the case \(a(x, y) = a(x), b(x, y) = b(x)v(y)\).

In particular, when \(T_n = T\) is independent of \(n\), our results apply to uniformly hyperbolic (Axiom A) systems \[20\], and to large classes of nonuniformly hyperbolic systems \[22, 23\]. A detailed account of discrete-time dynamical systems \(T\) for which our assumptions are verified can be found in \[13\] Sec. 10 and \[14\] Sec. 1; our results on homogenization apply to all the systems therein without restriction on the form of \(a\) and \(b\) and under optimal moment bounds.

Our third contribution is to incorporate families of fast dynamical systems \(T_n\) and measures \(\lambda_n\). Such fast-slow systems were studied in the situation of exact multiplicative noise (which does not require rough path theory) in \[13\]. In work in progress \[10\], it is shown how the assumptions in the current paper can be verified for a large class of families \(T_n\) of nonuniformly hyperbolic dynamical systems.

The main tool in showing convergence of the system (1) is rough path theory in the càdlàg setting, used in conjunction with the method in \[14\]. We note here that our second contribution outlined above (optimal moment assumptions) is due to switching from \(\alpha\)-Hölder to \(p\)-variation rough path topologies (which is analogous to the mode of convergence in the classical Donsker theorem, see e.g. \[5\] Sec. 3.2). Our results employ the stability of “forward” (Itô) rough differential equations (RDEs) with jumps recently studied in \[10\], which we extend herein to the Banach space setting (though we restrict attention to the case of level-2 rough paths). The works \[7,3,4\] also study RDEs in the presence of jumps, but primarily focus on “geometric” (Marcus) notions of solution.

1.1. Illustrative example. Let \(M = [0, 1]\). For \(\gamma \geq 0\), we consider the intermittent map \(T\): \(M \to M\),

\[
T y = \begin{cases} y(1 + 2^\gamma y^\gamma), & y \leq 1/2, \\ 2y - 1, & y > 1/2. \end{cases}
\]

This is a prototypical example of a slowly mixing dynamical system \[13\]; the specific example is due to \[14\]. For \(\gamma < 1\), there exists a unique \(T\)-invariant ergodic absolutely continuous probability measure \(\mu\).

We further restrict to \(\gamma < 1/2\), where the central limit theorem holds: for \(v: M \to \mathbb{R}^m\) Hölder with \(\int_M v \, d\mu = 0\), and \(\lambda\) any absolutely continuous probability measure (possibly \(\lambda = \mu\)), the random variables \(n^{-1/2} \sum_{j=0}^{n-1} v \circ T^j\) defined on the probability space \((M, \lambda)\) converge in law to a normal distribution.

Consider a discrete-time fast-slow system of the form (1) with \(a_n \equiv a, b_n \equiv b, T_n \equiv T\) independent of \(n\), and \(T\) such an intermittent map. Here \(a, b: \mathbb{R}^d \times M \to \mathbb{R}^d\) are suitably regular functions such that \(\int b(x, y) \, d\mu(y) = 0\) for all \(x \in \mathbb{R}^d\). Define the càdlàg random process \(x_n(t) = X_{[nt]}^{(n)}\). We prove that \(x_n \to X\) where \(X\) is the solution of an SDE. (The precise SDE is specified in Theorem \[2,10\] below.)
Prior results establish convergence of \( x_n \) when \( b \) is a product \( b(x, y) = b(x)v(y) \) with \( b : \mathbb{R}^d \to \mathbb{R}^{d \times m} \) sufficiently smooth and \( v : M \to \mathbb{R}^m \) as above. It was proved first for \( \gamma < \frac{1}{11} \) in [13] using a discrete-time version of Hölder rough paths [12], then improved to \( \gamma < \frac{1}{10} \) by obtaining optimal moment control in [16], and finally extended to the full range \( \gamma < \frac{1}{4} \) in [3] using \( p \)-variation rough paths with jumps [10]. See [5] for further history and discussion. The restriction that \( b \) is a product is now redundant by Theorem 2.10.

The forthcoming paper [16] considers the general setting (1) where \( \lambda \) are allowed to depend on \( n \), which applies to the case that \( \lambda \) do not depend on \( n \). This requires our main result Theorem 2.17.

The article is structured as follows. In Section 2 we state the main result of this article, Theorem 2.17, which gives precise conditions for the dynamics (1) to converge to the solution of an SDE. In Section 3 we collect the necessary material on càdlàg rough path theory in the Banach space setting. In Section 4 we prove Theorem 2.17. In Section 5 we give the version of Theorem 2.17 for the continuous-time dynamics. In Appendix A we give a Banach-space version of homogeneous Besov-variation and Besov-Hölder rough path embeddings.

Acknowledgements. I.C. is funded by a Junior Research Fellowship of St John’s College, Oxford. P.K.F. acknowledges partial support from the ERC, CoG-683164, the Einstein Foundation Berlin, and DFG research unit FOR2402. A.K. and I.M. acknowledge partial support from the European Advanced Grant StochExtHomog (ERC AdG 320977). A.K. is also supported by an Engineering and Physical Sciences Research Council grant EP/P034489/1. H.Z. is supported by the Chinese National Postdoctoral Program for Innovative Talents No: BX20180075. I.C., A.K., and H.Z. thank the Institute für Mathematik, TU Berlin, for its hospitality.

2. Discrete-time fast-slow systems. Statement of the main result

In this section we state our main result, Theorem 2.17. In fact, we first state a simplified version, Theorem 2.10, which applies to the case that \( a_n, b_n, T_n \) and \( \lambda_n \) do not depend on \( n \). We state the results separately not only because it eases our presentation, but also because Theorem 2.10 is slightly stronger than the naive restriction of Theorem 2.17 to the \( n \)-independent case (namely Assumption 2.6 below) is weaker than the naive restriction of Assumption 2.12.

For the remainder of this section, we fix a metric space \((M, \rho)\).

Definition 2.1. For \( \kappa \in [0,1) \) and \( m \geq 1 \), let \( C^\kappa(M, \mathbb{R}^m) \) denote the space of continuous \( \mathbb{R}^m \)-valued functions on \( M \) such that

\[
|v|_{C^\kappa} := \sup_{y \in M} |v(y)| + \sup_{y, y' \in M} \frac{|v(y) - v(y')|}{\rho(y, y')^\kappa} < \infty.
\]

We write \( C^\kappa(M) \) whenever \( m = 1 \). For \( \alpha \geq 0 \), define \( C^{\alpha, \kappa}(\mathbb{R}^d \times M, \mathbb{R}^d) \) to be the space of functions \( a = a(x, y) : \mathbb{R}^d \times M \to \mathbb{R}^d \) such that

\[
|a|_{C^{\alpha, \kappa}} := \sum_{|k| \leq \alpha} \sup_{x \in \mathbb{R}^d} |D^k a(x, \cdot)|_{C^\kappa} + \sum_{|k| = \alpha} \sup_{x, x' \in \mathbb{R}^d} \frac{|D^k a(x, \cdot) - D^k a(x', \cdot)|_{C^\kappa}}{|x - x'|^{\alpha - |\alpha|}} < \infty,
\]

where \( D^k \) acts on the \( x \) component.

For the remainder of the section, we fix parameters \( q \in (1, \infty], \kappa, \tilde{\kappa} \in (0,1) \), and \( \alpha > 2 + \frac{d}{q} \).
2.1. $n$-independent case. We now describe the assumptions and preliminary results required to state Theorem 2.10. We fix $a \in C^{1+\hat{c}}(\mathbb{R}^d \times M, \mathbb{R}^d)$ and $b \in C^{\alpha,\kappa}(\mathbb{R}^d \times M, \mathbb{R})$, and consider for every integer $n \geq 1$ the discrete-time dynamical system posed on $\mathbb{R}^d \times M$

\[ X_k^{(n)}(X_k^{(n)}, Y_k) + n^{-1/2}b(X_k^{(n)}, Y_k), \quad Y_{k+1} = TY_k, \]

where $T : M \to M$ is a Borel measurable map, $X_0^{(n)} = \xi_n \in \mathbb{R}^d$, and $Y_0$ is drawn randomly from a Borel probability measure $\lambda$ on $M$. Our first assumption deals with the function $a$.

**Assumption 2.2.** There exists $\bar{a} \in C^{1+\hat{c}}(\mathbb{R}^d, \mathbb{R})$ such that

\[ \sup_{t \in [0,1]} |V_n(t) - \bar{a}t|_{C^{1+\hat{c}}} \to 0 \quad \text{as} \quad n \to \infty, \]

where $V_n(t) = n^{-1} \sum_{k=0}^{[tn]} a(\cdot, Y_k)$.

To state our assumption on $b$, we need to introduce further notation. For $v, w \in C_0^\infty(M, \mathbb{R}^m)$ and $0 \leq s \leq t \leq 1$, define $W_{v,w}(t) \in \mathbb{R}^m$ and $W_{v,w,n}(s, t) \in \mathbb{R}^{m \times m}$ by

\[ W_{v,w}(t) = n^{-1/2} \sum_{0 \leq k < \lfloor nt \rfloor} v(Y_k), \quad W_{v,w,n}(s, t) = \int_s^t (W_{v,w}(r) - W_{v,w}(s)) \otimes dW_{v,w}(r). \]

Note in particular that

\[ W_{v,w,n}(t) = W_{v,w,n}(0, t) = n^{-1} \sum_{0 \leq k < \lfloor nt \rfloor} v(Y_k) \otimes w(Y_k). \]

Whenever $v = w$, we write simply $W_{v,n}$ for $W_{v,v,n}$.

For a subspace $C_0^\alpha(M)$ of $C_0^\infty(M)$, we let $C_0^\alpha(M, \mathbb{R}^m)$ denote the space of all $v \in C_0^\infty(M, \mathbb{R}^m)$ such that $v^i \in C_0^\alpha(M)$ for all $i = 1, \ldots, m$, and we let $C_0^{\alpha,\kappa}(\mathbb{R}^d \times M, \mathbb{R})$ denote the subspace of all $f \in C_0^{\alpha,\kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$ for which $f(x, \cdot) \in C_0^\alpha(M, \mathbb{R}^d)$ for all $x \in \mathbb{R}^d$.

**Assumption 2.3.** There exists a closed subspace $C_0^\alpha(M)$ of $C_0^\infty(M)$ such that $b \in C_0^{\alpha,\kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$ and such that

(i) for all $v, w \in C_0^\alpha(M)$ there exists $K = K_{v,w,q} > 0$ such that for all $n \geq 1$ and $0 \leq k, \ell \leq n$

\[ |W_{v,n}(k/n) - W_{v,n}(\ell/n)|_{L^{2q}(\lambda)} \leq K n^{-1/2} |k - \ell|^{1/2} \]

and

\[ |W_{v,w,n}(k/n, \ell/n)|_{L^{2q}(\lambda)} \leq K n^{-1} |k - \ell|. \]

(ii) there exists a bilinear operator $\mathcal{B}_0 : C_0^\alpha(M) \times C_0^\alpha(M) \to \mathbb{R}$ such that for every $m \geq 1$ and every $v \in C_0^\infty(M, \mathbb{R}^m)$, it holds that $(W_{v,n}, W_{v,n}) \to \lambda (W_v, W_v)$ as $n \to \infty$ in the sense of finite-dimensional distributions, where $W_v$ is an $\mathbb{R}^m$-valued Brownian motion and

\[ W_{v,0}^j(t) = \int_0^t W_v^j dW_v^j + \mathcal{B}_0(v^i, v^j) t. \]

**Remark 2.4.** One should compare part (i) of Assumption 2.3 to [13, Thm. 9.1] and [14, Assump. 2.2] in which one imposes the restriction $q > 3$. As mentioned in the introduction, we are able to deal with the optimal moment condition $q > 1$ by working with $p$-variation rather than Hölder rough path topologies.
Proof. (a) It follows from Assumption 2.3 that boundedness of $B$

Suppose Assumption 2.3 holds. Then the quadratic form

Taking expectations on both sides and letting $n \to \infty$ yields the desired result.

(b) Boundedness of $\mathfrak{B}_0$ follows from (5) and (6) with $k = 0$, $\ell = n$. By definition of $\mathfrak{B}$, we have $|\mathfrak{B}(v, w)| \leq |\mathfrak{B}_0(v, w)| + \frac{1}{n}E(\kappa)\Sigma_{1}(\lambda)$, yielding boundedness of $\mathfrak{B}$.

Lemma 2.9. Suppose Assumption 2.3 holds. Then the quadratic form

is positive semi-definite and the unique positive semi-definite $\sigma$ satisfying $\sigma^2 = \Sigma$ is Lipschitz.
Proof. Positive semi-definiteness of $\Sigma$ follows from part (a) of Proposition 2.8. Moreover, $b$ lies in $C^{\alpha,\kappa}([0,1] \times M, \mathbb{R}^d)$ with $\alpha > 2 + \frac{\kappa}{2} > 2$, so $\Sigma$ is $C^2$ with globally bounded derivatives to second order. The conclusion now follows from [21, Thm. 5.2.3].

As a consequence of Lemma 2.11 and [21, Cor. 5.1.2], for a Brownian motion $B$ on $\mathbb{R}^d$ and a Lipschitz function $\tilde{a} : \mathbb{R}^d \to \mathbb{R}^d$, there is a unique strong solution to the SDE $dX = \tilde{a}(X) \, dt + \sigma(X) \, dB$, $X(0) = \xi$.

**Theorem 2.10.** Suppose that Assumptions 2.2 and 2.3 hold and that $\lim_{n \to \infty} \xi_n = \xi$. Define the càdlàg path
\[
x_n : [0,1] \to \mathbb{R}^d, \quad x_n(t) = X^{(n)}_{\lfloor nt \rfloor}.
\]
Then $x_n \to_{\text{in the uniform topology as } n \to \infty}$, where $X$ is the unique weak solution of the SDE
\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB, \quad X(0) = \xi.
\]
Here, $B$ is a standard Brownian motion in $\mathbb{R}^d$, $\alpha$ is defined as in Lemma 2.9, and $\tilde{a}$ is the Lipschitz function given by
\[
\tilde{a}^i(x) = \check{a}^i(x) + \sum_{k=1}^d \mathfrak{R}_0(b^k(x,\cdot), \partial_k b^i(x,\cdot)), \quad i = 1, \ldots, d.
\]

We omit the proof of Theorem 2.10, which follows from trivial modifications to the proof in Section 4 of Theorem 2.17.

### 2.2. General case.
We now state the assumptions and preliminary results required for our main result, Theorem 2.17. We fix functions $a_n \in C^{1+\bar{\kappa},0}([0,1] \times M, \mathbb{R}^d)$ and $b_n, \in C^{\alpha,\kappa}([0,1] \times M, \mathbb{R}^d)$ satisfying
\[
\sup_{n \geq 1} |a_n|_{C^{1+\bar{\kappa},0}} + |b_n|_{C^{\alpha,\kappa}} < \infty, \quad \lim_{n \to \infty} |b_n - b|_{C^{\alpha,\kappa}} = 0.
\]
For $n \geq 1$, we are interested in the discrete-time fast-slow system (11) where $T_n : M \to M$ is a measurable map, $X_0^{(n)} = \xi_n \in \mathbb{R}^d$, and $Y_0^{(n)}$ is drawn randomly from a Borel probability measure $\lambda_n$ on $M$.

**Assumption 2.11.** There exists $\check{a} \in C^{1+\bar{\kappa}}(\mathbb{R}^d, \mathbb{R}^d)$ such that
\[
\sup_{t \in [0,1]} |V_n(t) - \check{a}t|_{C^{1+\bar{\kappa}}} \to_{\lambda_n} 0 \quad \text{as} \quad n \to \infty,
\]
where $V_n(t) = n^{-1} \sum_{k=0}^{\lfloor nt \rfloor - 1} a_n(\cdot, Y_k^{(n)})$.

As in (3), for $v,w \in C^\kappa(M, \mathbb{R}^m)$ and $0 \leq s \leq t \leq 1$, define $W_{v,n}(t) : [0,1] \to \mathbb{R}^m$, and $\mathbb{W}_{v,w,n}(s,t) \in \mathbb{R}^{m \times m}$ by
\[
W_{v,n}(t) = n^{-1/2} \sum_{0 \leq k \leq \lfloor nt \rfloor} v(Y_k^{(n)}),
\]
\[
\mathbb{W}_{v,w,n}(s,t) = \int_s^t (W_{v,n}(r) - W_{v,n}(s)) \otimes dW_{w,n}(r).
\]
Whenever $v = w$, we again write $\mathbb{W}_{v,n}$ for $\mathbb{W}_{v,v,n}$.

Recall our notational convention about subspaces $C^\kappa_0(M)$ of $C^\kappa(M)$ introduced before Assumption 2.3. Given a family of subspaces $(C^\kappa_0(M))_{n \in \mathbb{N} \cup \{ \infty \}}$ of $C^\kappa(M)$,
we call \( v = (v_n)_{n \in \mathbb{N} \cup \{\infty\}} \) a \( C^\infty(M, \mathbb{R}^m) \)-family if \( v_n \in C^\infty_n(M, \mathbb{R}^m) \) and \( \lim_{n \to \infty} |v_n - v_\infty|_{C^\infty} = 0 \).

**Assumption 2.12.** There exists a closed subspace \( C^\infty_n(M) \) of \( C^\infty(M) \) for each \( n \in \mathbb{N} \cup \{\infty\} \) such that \( b_n \in C^\infty_n(R^d \times M, \mathbb{R}^d) \), and

(i) for all \( v = (v_1, \ldots, w = (w_1, \ldots) \in \prod_{n \in \mathbb{N}} C^\infty_n(M) \) with

\[
\sup_n |v_n|_{C^\infty} + |w_n|_{C^\infty} < \infty ,
\]

there exists \( K = K_{v,w,q} > 0 \) such that for all \( n \in \mathbb{N} \) and \( 0 \leq k, \ell \leq n \)

\[
|W_{v,n}(k/n) - W_{v,n}(\ell/n)|_{L^2(\lambda_n)} \leq Kn^{-1/2}|k - \ell|^{1/2}
\]

and

\[
|\mathbb{W}_{v,w}(k/n, \ell/n)|_{L^q(\lambda_n)} \leq Kn^{-1}|k - \ell| .
\]

(ii) there exist bounded bilinear operators \( \mathcal{B}_1, \mathcal{B}_2 : C^\infty(M) \times C^\infty(M) \to \mathbb{R} \) such that for every \( m \geq 1 \) and every \( C^\infty_n(M, \mathbb{R}^m) \)-family \( v = (v_n)_{n \in \mathbb{N} \cup \{\infty\}} \),

(a) \( \lim_{n \to \infty} n^{-1} \sum_{i=0}^{n-1} \mathbb{E}_{\lambda_n}(v_i, v'_{i+1})(Y_i(n)) = \mathcal{B}_1(v_\infty^q, v_\infty^q) \),

(b) \( (W_{v,n}, \mathbb{W}_{v,n}) \to_{\lambda_n} (W_v, \mathbb{W}_v) \) as \( n \to \infty \) in the sense of finite-dimensional distributions, where \( W_v \) is an \( \mathbb{R}^m \)-valued Brownian motion and

\[
\mathbb{W}^v (t) = \int_0^t W^v_s \, dW^v_s + \mathcal{B}_2(v_\infty^q, v_\infty^q) t .
\]

**Remark 2.13.** As in Remark 2.6, under the assumption that \( \lambda_n \) is \( T_n \)-stationary, the simpler bounds

\[
|W_{v,n}(1)|_{L^2(\lambda_n)} \leq K \quad \text{and} \quad |\mathbb{W}_{v,w,n}(0,1)|_{L^q(\lambda_n)} \leq K
\]

for all \( n, n' \geq 1 \) imply part (i) of Assumption 2.12. Also, part (ii)(a) of Assumption 2.12 reduces to \( \lim_{n \to \infty} \mathbb{E}_{\lambda_n}(v_i, v_{i+1})(Y_i(n)) = \mathcal{B}_1(v_\infty^q, v_\infty^q) \).

**Proposition 2.14.** Suppose that part (ii) of Assumption 2.12 holds. Then there exists \( K > 0 \) such that for all \( n \in \mathbb{N} \), \( 0 \leq k, \ell \leq n \), and \( v, w \in C^\infty_n(M) \),

\[
|W_{v,n}(k/n) - W_{v,n}(\ell/n)|_{L^2(\lambda_n)} \leq Kn^{-1/2}|k - \ell|^{1/2} ,
\]

\[
|\mathbb{W}_{v,w}(k/n, \ell/n)|_{L^q(\lambda_n)} \leq K|v|_{C^\infty} n^{-1/2}|k - \ell| .
\]

**Proof.** Identical to Proposition 2.14. \( \square \)

**Proposition 2.15.** Suppose that Assumption 2.12 holds. Let \( \mathcal{B} = \frac{1}{2} \mathcal{B}_1 + \mathcal{B}_2 \). Then the covariance of \( W_v \) is given by

\[
\mathbb{E}W^v(1)W^v(1) = \mathcal{B}(v_\infty^q, v_\infty^q) + \mathcal{B}(v_\infty^q, v_\infty^q) .
\]

**Proof.** Exactly the same as part (a) of Proposition 2.8 upon replacing \( W_{v,n} \) by \( W_{v,n} \) and \( W_v \) by \( W_v \), and using part (ii)(a) of Assumption 2.12. \( \square \)

**Lemma 2.16.** Suppose that Assumption 2.12 holds. Then the symmetric quadratic form

\[
\Sigma^{ij}(x) = \mathcal{B}(b^i_\infty(x, \cdot), b^j_\infty(x, \cdot)) + \mathcal{B}(b^i_\infty(x, \cdot), b^j_\infty(x, \cdot)) , \quad i, j = 1, \ldots, d ,
\]

is positive semi-definite and the unique positive semi-definite \( \sigma \) satisfying \( \sigma^2 = \Sigma \) is Lipschitz.
Proof. Identical to Lemma 2.9.

As before, by Lemma 2.10 and Cor. 5.1.2, for a Brownian motion \( B \) on \( \mathbb{R}^d \) and a Lipschitz function \( \tilde{a} : \mathbb{R}^d \to \mathbb{R}^d \), there is a unique strong solution to the SDE
\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB , \quad X(0) = \xi .
\]
Theorem 2.17. Suppose that Assumptions 2.11 and 2.12 hold, and that
\[
\lim_{n \to \infty} \xi_n = \xi \in \mathbb{R}^d .
\]
Define the càdlàg path
\[
x_n : [0, 1] \to \mathbb{R}^d , \quad x_n(t) = X_{[nt]}^{(n)} .
\]
Then \( x_n \to \lambda_n X \) in the uniform topology as \( n \to \infty \), where \( X \) is the unique weak solution of the SDE
\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB , \quad X(0) = \xi ,
\]
where \( B \) is a standard Brownian motion in \( \mathbb{R}^d \), \( \sigma \) is defined as in Lemma 2.16 and \( \tilde{a} \) is the Lipschitz function given by
\[
\tilde{a}^i(x) = \tilde{a}^i(x) + \sum_{k=1}^d \mathbb{E}_2(b^k_\infty(x, \cdot), \partial_k b^i_\infty(x, \cdot)) , \quad i = 1, \ldots, d .
\]

3. Banach space valued càdlàg rough paths

In this section, we collect all the necessary results on càdlàg rough path theory in Banach spaces which will be needed in the sequel.

For Banach spaces \( A, B \), we denote their algebraic tensor product by
\[
A \otimes_a B := \text{span} \{ a \otimes b : a \in A, b \in B \} .
\]
Given \( f \in A^* \) (the dual space of \( A \)), \( g \in B^* \), one may define an element on \((A \otimes_a B)^*\) by
\[
(f \otimes g)(\sum_{i=1}^N a_i \otimes b_i) := \sum_{i=1}^N f(a_i)g(b_i) .
\]
As a result, we consider \( A^* \otimes_a B^* \) as a subspace of \((A \otimes_a B)^*\). Generally, there are different (in)equivalent norms on \( A \otimes_a B \). We call a norm \(| \cdot |_{A \otimes B} \) on the vector space \( A \otimes_a B \) admissible (or reasonable), if for any \( a \in A, b \in B, f \in A^*, g \in B^* \),
\[
|a \otimes b|_{A \otimes B} \leq |a|_{A^*}|b|_B , \quad |f \otimes g|_{(A \otimes B)^*} \leq |f|_{A^*}|g|_{B^*} ,
\]
where \(| \cdot |_{(A \otimes B)^*}\) is defined as the dual norm on \((A \otimes B, | \cdot |_{A \otimes B})^*\). One may then complete \( A \otimes_a B \) under \(| \cdot |_{A \otimes B}\) to obtain a Banach space. All the tensor product spaces \( A \otimes B \) we consider in the sequel will implicitly be assumed to be Banach spaces completed from such an admissible norm.

Definition 3.1. A partition over an interval \([s, t]\) is a set \( \mathcal{P} \) of subintervals of \([s, t]\) of the form \( \mathcal{P} = \{ [t_0, t_1], [t_1, t_2], \ldots, [t_{k-1}, t_k] \} \) with \( t_i < t_{i+1} \) and \( t_0 = s, t_k = t \).
We define the mesh size of the partition as \( |\mathcal{P}| := \max_{[u, v] \in \mathcal{P}} |u - v| \).

For a Banach space \( B \) and \( p > 0 \), let \( \mathcal{V}^{p-\text{var}}([s, t], B) \) denote the space of all functions \( \Xi : \{(u, v) \in [s, t]^2 : u \leq v\} \to B \) such that \( \Xi(u, u) = 0 \) and
\[
\|\Xi\|_{p-\text{var};[s, t]} := \sup_{\mathcal{P}} \left( \sum_{[u, v] \in \mathcal{P}} \|\Xi(u, v)\|^p \right)^{1/p} < \infty ,
\]
where the supremum is over all partitions of \([s, t]\).
Note that if \( p \geq 1 \), then \( \mathcal{V}^{p\text{-var}}([s, t], \mathcal{B}) \) is a Banach space with norm \( \| \cdot \|_{p\text{-var};[s, t]} \). In the sequel, we will drop the reference to the interval \([s, t] \) whenever \([s, t] = [0, T] \).

We will also occasionally refer to \( p \)-variation over not necessarily closed intervals, i.e., \((s, t)\) or \([s, t] \) instead of \([s, t] \), with the obvious interpretation.

For a Banach space \( \mathcal{B} \), we equip \( \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) with the multiplication operation \((a, M)(b, N) := (a + b, M + a \otimes b + N) \). Note that the multiplicative identity in \( \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) is \((0, 0) \) and every element possesses an inverse given by \((a, M)^{-1} = (-a, -M + a \otimes a) \).

Hence \( \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) is a group.

**Definition 3.2.** Let \( \mathcal{B} \) be a Banach space. For a path \( X : [s, t] \rightarrow \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) and \( s \leq u \leq v \leq t \), define the increment \( X(u, v) := (X(u), \dot{X}(u, v)) := X(u)^{-1}X(v) \).

For \( p \geq 1 \), define the (homogeneous) \( p \)-variation of \( X \) by

\[
\| X \|_{p\text{-var};[s, t]} := \| X \|_{p\text{-var};[s, t]} + \| \dot{X} \|_{p/2\text{-var};[s, t]}^{1/2}.
\]

For \( p \in [2, 3) \), a \( p \)-rough path over \( \mathcal{B} \) is a càdlàg function \( X : [0, T] \rightarrow \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) such that \( X(0) = 0 \) and \( \| X \|_{p\text{-var}} < \infty \). For \( p \)-rough paths \( \tilde{X}, \hat{X} \), we define the (inhomogeneous) rough path metric by

\[
\| X, \tilde{X} \|_{p\text{-var}} := \| X - \tilde{X} \|_{p\text{-var}} + \| X - \hat{X} \|_{p/2\text{-var}},
\]

as well as the (Skorokhod-type) \( p \)-variation metric

\[
\sigma_{p\text{-var}}(X, \tilde{X}) := \inf_{\lambda \in \Lambda} \{ |\lambda| + \| X, \tilde{X} \circ \lambda \|_{p\text{-var}} \},
\]

where \( \Lambda \) denotes the set of all continuous increasing bijections \( \lambda : [0, T] \rightarrow [0, T] \), and

\[
|\lambda| := \sup_{t \in [0, T]} |t - \lambda(t)|.
\]

Let \( \mathcal{D}^{p\text{-var}}(\mathcal{B}) \) denote the space of all \( p \)-rough paths equipped with the metric \( \sigma_{p\text{-var}} \), and let \( \mathcal{D}^{0, p\text{-var}}(\mathcal{B}) \) denote the closure in \( \mathcal{D}^{p\text{-var}}(\mathcal{B}) \) of all \( \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \)-valued piecewise constant paths. For \( p \in [1, 2) \), define the \( p \)-variation \( \| \cdot \|_{p\text{-var}} \) of a path \( X : [0, T] \rightarrow \mathcal{B} \), as well as the metric \( \sigma_{p\text{-var}} \) and spaces \( \mathcal{D}^{p\text{-var}}(\mathcal{B}), \mathcal{D}^{0, p\text{-var}}(\mathcal{B}) \) in the exact same way as above but without the component \( X \).

**Remark 3.3.** The reason for introducing \( \mathcal{D}^{0, p\text{-var}} \) and \( \mathcal{D}^{0, p\text{-var}} \) is due to separability properties. Indeed, whenever \( \mathcal{B} \) is separable, \( \mathcal{D}^{0, p\text{-var}}(\mathcal{B}) \) and \( \mathcal{D}^{0, p\text{-var}}(\mathcal{B}) \) are Polish spaces, which is useful for our applications in probability theory.

We state a basic interpolation estimate which will be helpful in the sequel. Define

\[
\| X, \tilde{X} \|_{\infty} = \| X - \tilde{X} \|_{\infty} + \| X - \hat{X} \|_{\infty},
\]

where \( \| \Xi \|_{\infty} := \sup_{s,t} |\Xi(s, t)| \) (as usual, we treat \( X \) as a two parameter function by \( X(s, t) = X(t) - X(s) \)).

**Lemma 3.4.** For \( p' \geq p \geq 1 \) and \( X, \tilde{X} : [0, T] \rightarrow \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \), it holds that

\[
\| X, \tilde{X} \|_{p'\text{-var}} \leq \| X, \tilde{X} \|_{\infty}^{1-p/p'} \| X, \tilde{X} \|_{p\text{-var}}^{p/p'}.
\]

**Proof.** We readily see that

\[
\| X, \tilde{X} \|_{p'\text{-var}} \leq \| X - \tilde{X} \|_{\infty}^{1-p/p'} \| X - \tilde{X} \|_{p\text{-var}}^{p/p'} + \| X - \hat{X} \|_{\infty}^{1-p/p'} \| X - \hat{X} \|_{p/2\text{-var}}^{p/p'},
\]

and the conclusion follows by Hölder’s inequality

\[
a^{\theta}a^{1-\theta} + b^{\theta}b^{1-\theta} \leq (a+b)^{\theta}(a+b)^{1-\theta}
\]

for \( \theta \in [0, 1] \) and \( a, \bar{a}, b, \tilde{b} \geq 0 \).
We now introduce rough integration in the level-2 rough path case. Given Banach spaces $\mathcal{B}, \mathcal{E}$, let $L(\mathcal{B}, \mathcal{E})$ denote the space of bounded linear operators from $\mathcal{B}$ to $\mathcal{E}$. For $p \in [2, 3)$ and $X \in D^{p,\text{var}}(\mathcal{B})$, we call $(Y, Y')$ an $\mathcal{E}$-valued $X$-controlled rough path if

\[ Y \in D^{p,\text{var}}(\mathcal{E}) , \quad Y' \in D^{p,\text{var}}(L(\mathcal{B}, \mathcal{E})) , \]

and $R \in \mathcal{Y}_{\text{p/2-var}}(\mathcal{E})$, where

\[ R(s, t) := Y(s, t) - Y'(s)X(s, t) . \]

We denote the space of $X$-controlled rough paths as $\mathcal{D}^{p/2,\text{var}}(\mathcal{E})$. In the following, we are interested in $\mathbb{R}^{d}$-valued RDEs, i.e. $\mathcal{E} = \mathbb{R}^{d}$. In this case, one has the following stability of rough integration.

**Lemma 3.5.** Let $X \in D^{p,\text{var}}(\mathcal{B})$, $(Y, Y') \in \mathcal{D}^{p/2,\text{var}}(\mathbb{R}^{d})$, and $H \in C^{2}(\mathbb{R}^{d}, L(\mathcal{B}, \mathbb{R}^{d}))$. Then, for every $t \in [0, T]$, the following integral (with values in $\mathbb{R}^{d}$) is well-defined

\[ \mathcal{I}_{X}(Y)(t) := \int_{0}^{t} H(Y(s))^{-} \text{d}X(s) := \lim_{|P| \to 0} \sum_{[u, v] \in P} \Xi(u, v) , \]

where $P$ are partitions of $[0, t]$ and, for $i = 1, \ldots, d$,

\[ \Xi(u, v)^{i} = H^{i}(Y(u))X(u, v) + \sum_{k=1}^{d}(\partial_{k}H^{i}(Y(u)) \otimes H^{k}(Y(u)))X(u, v) . \]

Furthermore, $(H(Y), DH(Y)Y')$ and $(\mathcal{I}_{X}(Y), H(Y))$ are $X$-controlled rough paths.

**Proof.** First, we check that $\Xi(u, v) \in \mathbb{R}^{d}$ is well-defined. Indeed, for $i, k = 1, \ldots, d$, $H^{k}(Y(u))$ and $\partial_{k}H^{i}(Y(u))$ are elements of $\mathcal{B}^{*}$. By admissibility of norms [12], $\mathcal{B}^{*} \otimes_{a} \mathcal{B}^{*}$ is a subspace of $(\mathcal{B} \otimes \mathcal{B})^{*}$, and thus $H^{k}(Y(u)) \otimes \partial_{k}H^{i}(Y(u)) \in (\mathcal{B} \otimes \mathcal{B})^{*}$. Hence $\Xi(u, v)$ is well-defined as claimed.

The claim that $(H(Y), DH(Y)Y')$ is an $X$-controlled rough path follows from Taylor expansion. Indeed, defining

\[ R^{H(Y)}(s, t) := H(Y(t)) - H(Y(s)) - DH(Y(s))Y'(s)X(s, t) , \]

one can check that $R^{H(Y)} \in \mathcal{Y}^{2}(L(\mathcal{B}, \mathbb{R}^{d}))$. Then one has the identity

\[ \Xi(s, t) - \Xi(s, u) = -R^{H(Y)}(s, u)X(u, t) - (DH(Y)Y')(s, u)X(u, t) . \]

According to the generalized sewing lemma [10, Thm. 2.5], the integral $\mathcal{I}_{X}(Y)$ is well-defined, and furthermore one has the local estimate

\[ |\mathcal{I}_{X}(Y)(s, t) - \Xi(s, t)| \leq C(\|R^{H(Y)}\|_{\mathcal{Y}^{p/2,\text{var}};[s, t]} \|X\|_{\mathcal{Y}^{p,\text{var}};[s, t]} + \|DH(Y)Y'\|_{\mathcal{Y}^{p,\text{var}};[s, t]} \|X\|_{\mathcal{Y}^{p,\text{var}};[s, t]} , \]

which implies that $(\mathcal{I}_{X}(Y), H(Y))$ is also an $X$-controlled rough path. \qed

**Remark 3.6.** Generally, to integrate $(Y, Y')$ against $X$, one needs $Y(t) \in L(\mathcal{B}, \mathcal{E})$ and $Y'(t) \in L(\mathcal{B}, L(\mathcal{B}, \mathcal{E}))$ to have the identity $Y(s, t) = Y'(s)X(s, t) + R(s, t)$. In this case, one further needs the embedding $L(\mathcal{B}, L(\mathcal{B}, \mathcal{E})) \hookrightarrow L(\mathcal{B} \otimes \mathcal{B}, \mathcal{E})$ to define $\Xi(s, t) := Y(s)X(s, t) + Y'(s)X(s, t)$. Luckily, in the above case where $\mathcal{E} = \mathbb{R}^{d}$, the embedding assumption is replaced by the fact $DH(Y)Y' \in L(\mathcal{B} \otimes \mathcal{B}, \mathbb{R}^{d})$ which follows by admissibility of norms.
The main convergence result for rough differential equations which we will require is the following. The proof, which we omit, is essentially the same as the finite dimensional case, i.e., [10, Thm. 3.8, 3.9], thanks to admissibility of norms.

**Theorem 3.7.** Let \( A, B \) be Banach spaces, \( q \in [1, 2), p \in [2, 3) \) with \( 1/p + 1/q > 1 \), and \( F \in C^\beta([\mathbb{R}^d, L(A, \mathbb{R}^d)]), H \in C^\gamma([\mathbb{R}^d, L(B, \mathbb{R}^d)]) \) for \( \beta > q, \gamma > p \). Then, for any \( V \in D^{p\text{-var}}(A) \), \( X \in D^{p\text{-var}}(B) \), and \( Y_0 \in \mathbb{R}^d \), there exists a unique \( X \)-controlled rough path \( (Y, Y') \in D^{p/2\text{-var}}(\mathbb{R}^d) \) solving the equation

\[
Y(t) = Y_0 + \int_0^t F(Y(s))^- dV(s) + \int_0^t H(Y(s))^- dX(s),
\]

where \( \int_0^t H(Y(s))^- dX(s) \) is defined by (15). Moreover, the solution map is locally Lipschitz in the sense that

\[
\|Y - \tilde{Y}\|_{p\text{-var}} \lesssim \|X; \tilde{X}\|_{p\text{-var}} + \|V - \tilde{V}\|_{q\text{-var}} + |Y_0 - \tilde{Y}_0|,
\]

where the proportionality constant is uniform over bounded classes of driving signals.

For our purposes, it will be useful to record the following corollary stated in terms of the metrics \( \sigma_{p\text{-var}} \) and \( \sigma_{q\text{-var}} \).

**Corollary 3.8.** Let notation be as in Theorem 3.7. Consider the solution map to equation (16)

\[
\Phi : D^{p\text{-var}}(A) \times D^{p\text{-var}}(B) \times \mathbb{R}^d \to D^{p\text{-var}}(\mathbb{R}^d),
\]

\[
\Phi : (V, X, Y_0) \mapsto Y.
\]

Equip \( D^{p\text{-var}}(\mathbb{R}^d) \) with the norm \( |Y(0)| + \|Y\|_{p\text{-var}} \) and \( D^{q\text{-var}}(A) \times D^{q\text{-var}}(B) \times \mathbb{R}^d \) with the product metric \( (\sigma_{q\text{-var}}, \sigma_{p\text{-var}}, |\cdot|) \). Then every point \( (V, X, Y_0) \), where \( V, X \) are continuous, is a continuity point of \( \Phi \).

**Proof:** Consider \( X \in D^{p\text{-var}}(B) \) continuous. Observe that, for \( p' > p \geq 1 \), if \( \sigma_{p\text{-var}}(X_n, X) \to 0 \), then \( \|X_n; X\|_{p'\text{-var}} \to 0 \). Indeed, it suffices to show that \( \|X; X \circ \lambda_n\|_{p'\text{-var}} \to 0 \) whenever \( \lambda_n \) is a sequence in \( \Lambda \) for which \( |\lambda_n|_\infty \to 0 \), which in turn follows from \( \|X; X \circ \lambda_n\|_{\infty} \to 0 \) (by continuity of \( X \)) and the interpolation estimate (14). The same considerations apply for continuous \( V \in D^{q\text{-var}}(A) \), and the result follows by applying Theorem 3.7 to any \( p' \in (p, \gamma) \) and \( q' \in (q, \beta) \). □

**Remark 3.9.** Recall that, for the classical \((J_1)\) Skorokhod space \( D \), a pair \((x, y) \in D^2\) is a continuity point of the addition map \( D^2 \to D \), \((x, y) \mapsto x + y\), whenever one of \( x \) or \( y \) is continuous. In a similar way, if one instead equips \( D^{p\text{-var}}(\mathbb{R}^d) \) with the metric \( |Y(0) - \tilde{Y}(0)| + \sigma_{p\text{-var}}(Y, \tilde{Y}) \), then one can show that \((V, X, Y_0)\) is a continuity point of \( \Phi \) whenever one of \( X \) or \( V \) is continuous.

We conclude this section with the following result which will be helpful in controlling the \( p \)-variation and càdlàg modulus of continuity of paths.

**Proposition 3.10.** Suppose \( X = (X, \mathbb{X}) : [0, T] \to \mathcal{B} \oplus \mathcal{B}^{\otimes 2} \) is a càdlàg piecewise constant random path with jump times contained in a deterministic set \( \{t_j\}_{0 \leq j \leq n} \subset [0, T] \) with \( t_0 < t_1 < \ldots < t_n \), such that, for some \( C_1, C_2 > 0 \), \( \beta \in (0, 1/2] \), and \( q \in [1, 1] \),

\[
|X(t_i, t_j)|_{L^{2q}} \leq C_1|t_j - t_i|^\beta, \quad |X(t_i, t_j)|_{L^q} \leq C_2|t_j - t_i|^{2\beta}.
\]
If \(2q > \frac{1}{\beta}\), then for any \(\alpha \in \left(\frac{1}{2q}, \beta\right)\)

\[
E\|X\|_{2q,\alpha-\var}^{2q} \leq CT^{\alpha-\beta}(C_1 + C_2^{1/2})
\]

and

\[
E\left[\left(\sup_{t_i \neq t_j} \frac{|X(t_i, t_j)| + |X(t_i, t_j)|^{1/2}}{|t_i - t_j|^{\alpha-\beta/2}}\right)^{2q}\right] \leq C(C_1 + C_2^{1/2})
\]

for a constant \(C > 0\) depending only on \(\alpha, \beta, q\).

For the proof, we require the following lemma.

**Lemma 3.11.** Let \(X\) be as in Proposition 3.10. Then there exists a continuous path \(\tilde{X} = (\tilde{X}, \tilde{X}): [0, T] \to B \oplus B^2\) such that \(\tilde{X}(t_i) = \tilde{X}(t_i)\), and

\[
|\tilde{X}(s, t)|_{L^{2q}} \leq 3^{1-\beta}C_1|t-s|^\beta, \quad |\tilde{X}(s, t)|_{L^q} \leq 3^{2-2\beta}(C_2 + C_3^2)|t-s|^{2\beta}.
\]

**Proof.** Let us define \((\tilde{X}, \tilde{X})\) for \(t \in [t_j, t_{j+1}]\) by

\[
\tilde{X}(t) := X(t_j) + \frac{t-t_j}{t_{j+1} - t_j}X(t_j, t_{j+1}),
\]

\[
\tilde{X}(0, t) := X(0, t_j) + \frac{t-t_j}{t_{j+1} - t_j}(X(0, t_{j+1}) - X(0, t_j)).
\]

To prove (20), consider \(s < t\) with \(s \in [t_j, t_{j+1}], t \in [t_k, t_{k+1}]\). Further we suppose that \(j < k\) (the case \(j = k\) is similar and simpler). Then

\[
|\tilde{X}(s, t)|_{L^{2q}} \leq |X(s, t_{j+1})|_{L^{2q}} + |X(t_{j+1}, t_k)|_{L^{2q}} + |X(t_k, t)|_{L^{2q}}
\]

\[
\leq C_1(|t_{j+1} - s|^{\beta} + |t_{j+1} - t_k|^{\beta} + |t - t_k|^{\beta})
\]

\[
\leq 3^{1-\beta}C_1|t-s|^{\beta}.
\]

Furthermore, one can check that

\[
|\tilde{X}(s, t_{j+1})|_{L^q} \leq (C_2 + C_3^2)|t_{j+1} - s|^{2\beta}.
\]

A similar estimate holds for \(\tilde{X}(t_k, t)\). Hence

\[
|\tilde{X}(s, t)|_{L^q} \leq |\tilde{X}(s, t_{j+1})|_{L^q} + |\tilde{X}(t_{j+1}, t_k)|_{L^q} + |\tilde{X}(t_k, t)|_{L^q}
\]

\[
\leq 3^{1-2\beta}(C_2 + C_3^2)|t-s|^{2\beta} + C_3^2|t-s|^{2\beta} + 2^{1-\beta}C_3^2|t-s|^{2\beta}
\]

\[
\leq 3^{2-2\beta}(C_2 + C_3^2)|t-s|^{2\beta}.
\]

**Proof of Proposition 3.10** Let \(\tilde{X}\) be as in Lemma 3.11 and suppose \(2q > \frac{1}{\beta}\) and \(\alpha \in \left(\frac{1}{2q}, \beta\right)\). Using the notation in Appendix A we have by Corollary A.3

\[
E\|\tilde{X}\|_{1,\alpha-\var}^{2q} \leq C(\alpha, q)T^{\alpha-\beta}E\|\tilde{X}\|_{W^{\alpha, 2q}}^{2q}
\]

\[
= C(\alpha, q)T^{\alpha-\beta}E\left[\int_{[0, 1]^2} \frac{|\tilde{X}(s, t)|_{B^q}^{2q} + |\tilde{X}(s, t)|_{S^{\alpha-\var}}^{2q}}{|t-s|^{2\alpha q+1}} \, ds \, dt\right]^{1/2q}.
\]
Using the estimate (20) and the condition $\alpha < \beta$, the final expectation is bounded by $\lambda(C_1 + C_2^{1/2})$, where $\lambda$ depends only on $\beta - \alpha$. In exactly the same way, using Corollary A.2, $E(||\tilde{X}||_{2q, \alpha < \beta}^{2q}) \leq C(C_1 + C_2^{1/2})$. The conclusion follows since $\tilde{X}(t_i) = X(t_i)$ and $X$ is constant on $[t_i, t_i]$.

\[ \Box \]

4. Proof of the main result

This section is devoted to the proof of Theorem 2.1. Throughout this section, we let notation be as in Section 2.2.

The first step is to reformulate the system (1) as a càdlàg controlled ODE. We introduce the Banach spaces

\[ A = C^{1+\bar{\kappa}}(\mathbb{R}^d, \mathbb{R}^d) \quad \text{and} \quad B = \tilde{C}^{\bar{\theta}}(\mathbb{R}^d, \mathbb{R}^d), \]

where $\theta \in (2, \alpha - \frac{4}{q})$ is fixed and $\tilde{C}^{\theta}$ denotes the closure in $C^{\theta}$ of smooth functions. Note that the space $B$ is separable and contains $C^{\theta} (\mathbb{R}^d, \mathbb{R}^d)$ for all $\theta > \theta$. We furthermore equip $B^{\otimes 2}$ with the admissible norm as specified in [14, Prop. 4.5].

For any $\eta > 0$, it holds for the point evaluation map $F : \mathbb{R}^d \to L(C^{\eta}(\mathbb{R}^d, \mathbb{R}^d), \mathbb{R}^d)$, given by $F(x) : u \mapsto u(x)$, that $F \in C^{\eta}(\mathbb{R}^d, L(C^{\eta}(\mathbb{R}^d, \mathbb{R}^d), \mathbb{R}^d))$. We let $F : \mathbb{R}^d \to L(A, \mathbb{R}^d)$ and $H : \mathbb{R}^d \to L(B, \mathbb{R}^d)$ denote the corresponding point evaluation maps.

The following lemma is now immediate from Theorem 4.1.

**Lemma 4.1.** The càdlàg RDE

\[ dx(t) = F(x^-(t)) \, dV(t) + H(x^-(t)) \, dW(t), \quad x(0) = \xi \in \mathbb{R}^d \]

is well-posed for any $(V, W) \in D^{3+\vartheta}(\mathbb{R}^d, \mathbb{R}^d) \times D^{p,\vartheta}(\mathbb{R}^d)$ with $\beta \in [1, 1 + \bar{\kappa})$ and $p \in [1, \theta)$.

We introduce the $A$-valued and $B$-valued paths

\[ V_n(t) = n^{-1} \sum_{k=0}^{[tn]} a_n(\cdot, Y^{(n)}_k), \quad W_n(t) = n^{-1/2} \sum_{k=0}^{[tn]} b_n(\cdot, Y^{(n)}_k), \]

and let $W_n = (W_n, W_n)$ be the canonical level-2 lift of $W_n$.

**Lemma 4.2.** The path $x_n$, given by (10), is the unique solution of the càdlàg ODE

\[ dx_n = F(x^-_n) \, dV_n + H(x^-_n) \, dW_n, \quad x_n(0) = \xi_n \in \mathbb{R}^d. \]

**Proof.** Observe that $x_n$ given by (10) satisfies for all $1 \leq k \leq n$

\[ x_n(k/n) - x_n((k - 1)/n) = n^{-1}a(x_n((k - 1)/n), Y^{(n)}_{k-1}) \]
\[ + n^{-1/2}b(x_n((k - 1)/n), Y^{(n)}_{k-1}) \]
\[ = \int_{(k-1)/n}^{k/n} F(x^-_n(s)) \, dV_n(s) + H(x^-_n(s)) \, dW_n(s). \]

\[ \Box \]

Following Lemmas 4.1 and 4.2, we are reduced to showing convergence in law for $V_n$ and $W_n$ in suitable rough path topologies and identifying the solution of the limiting RDE with an SDE. We first establish this result for the case that the support of $b_n$ is uniformly bounded, i.e., there exists a compact set $\mathfrak{K} \subset \mathbb{R}^d$ such that the support of $b_n$ is contained in $\mathfrak{K} \times M$ for all $n \in \mathbb{N} \cup \{\infty\}$. 






Proof. Lemma 4.5. Suppose that Assumption 2.11 holds. For every $p \in (2, 3)$ and $\beta \in (1, 2)$, there exists a random variable $(V, W)$ in $D^{0, \beta \text{-var}}(A) \times D^{0, p \text{-var}}(B)$ such that $(V_n, W_n) \to \lambda_n (V, W)$, and such that $(V, W)$ is a.s. continuous. Moreover, if $\beta \in (1, 1 + \bar{\kappa})$, $p \in (2, \theta)$, then the RDE (22) driven by $(V, W)$ along the vector fields $(F,H)$ is the unique weak solution of the SDE (11).

Before proving Theorem 4.3 we first state an immediate consequence of Corollary 3.8, Lemma 4.2, Theorem 4.3, and the continuous mapping theorem.

Corollary 4.4. Suppose we are in the setting of Theorem 2.17 and that the support of $b_n$ is uniformly bounded. Then, for any $p > 2$, $x_n \to \lambda_n X$ in the $p$-variation norm $|x(0)| + \|x\|_{p \text{-var}}$, where $X$ is the unique weak solution of the SDE (11).

We break the proof of Theorem 4.3 into several lemmas.

Lemma 4.5. Suppose that Assumption 2.11 holds. For every $\beta > 1$, it holds that $|V_n - V|_{\beta \text{-var}} \to \lambda_n 0$ as $n \to \infty$, where $V(t) = \bar{a} t$.

Proof. For $|t - s| \geq n^{-1}$, observe that $|V_n(t) - V_n(s)|_A \leq 2 |t - s| a_n |C^{\kappa, \kappa}\beta|$. Thus $V_n$, as paths in $A$, have 1-variation uniformly bounded in $n \geq 1$ and $Y_0^{(n)} \in M$. The conclusion follows from Assumption 2.11 and interpolation (13).

Showing convergence of $W_n$ is more involved.

Lemma 4.6. Suppose that part (i) of Assumption 2.12 holds and that the support of $b_n$ is uniformly bounded. Then

$$
\mathbb{E}_{\lambda_n} \left[ |W_n(k/n) - W_n(\ell/n)|_{\dot{B}^q_2}^{2q} \right]^{1/(2q)} \lesssim n^{-1/2} |k - \ell|^{1/2},
$$

$$
\mathbb{E}_{\lambda_n} \left[ |W_n(k/n, \ell/n)|_{\dot{B}^q_{2/q}}^q \right]^{1/q} \lesssim n^{-1/2} |k - \ell|,
$$

uniformly in $n \geq 1$ and $0 \leq k, \ell \leq n$.

Proof. For a function $u : \mathbb{R}^d \to \mathbb{R}$, let us introduce the notation

$$
\Delta_{\sigma} u(x) = u(x + \sigma) - u(x), \quad \text{and} \quad \Delta_{\sigma}^{m+1} = \Delta_{\sigma} \circ \Delta_{\sigma}^m.
$$

For $s > 0$ and $p \geq 1$, recall the Besov space $B^s_p$ consisting of all $L^p$ functions $u : \mathbb{R}^d \to \mathbb{R}$ such that

$$
|u|_{B^s_p}^p = |u|_{L^p} + \int_{|\sigma| \leq 1} \sigma^{-sp - d} |\Delta_{\sigma}^{[s]+1} u|_{L^p}^p \, d\sigma < \infty.
$$

Let us further introduce the notation

$$
\Delta_{\sigma}^m W_n(k, \ell; x) = \sum_{r=k}^{\ell} \Delta_{\sigma}^m b_n(x, Y^{(n)}_r).
$$

Denote in the sequel $s = k/n$ and $t = \ell/n$. Proposition 2.14 implies that for each $m \geq 1$ (cf. [14, p. 4088])

$$
\mathbb{E}_{\lambda_n} \left[ |\Delta_{\sigma}^m W_n(s, t; x)|_{\dot{B}^q_2} \right]^{1/(2q)} \lesssim |\Delta_{\sigma}^m b_n(x, \cdot)|_{C^w} |t - s|^{1/2}.
$$

\begin{align}
\mathbb{E}_{\lambda_n} \left[ |\Delta_{\sigma}^m W_n(s, t; x)|_{\dot{B}^q_2} \right]^{1/(2q)} & \lesssim |\Delta_{\sigma}^m b_n(x, \cdot)|_{C^w} |t - s|^{1/2}.
\end{align}
Setting $m = \lceil \theta + \frac{d}{2q} \rceil + 1$, it follows that

$$
\mathbb{E}_{\lambda_n} \left[ |W_n(s, t; \cdot)|_{B_n}^{2q} \right] \lesssim \mathbb{E}_{\lambda_n} \left[ |W_n(s, t; \cdot)|_{B_n^{\theta + d/(2q)}}^{2q} \right]
$$

$$
= \mathbb{E}_{\lambda_n} \left[ \int |W_n(s, t; x)|^{2q} \, dx + \int_{|\sigma| \leq 1} |\sigma|^{-2q - d} \int |\Delta_{\sigma} m W_n(s, t; x)|^{2q} \, d\sigma \right]
$$

$$
\lesssim \int |b_n(x, \cdot)|_{C_\alpha}^{2q} |t - s|^q \, dx + \int_{|\sigma| \leq 1} |\Delta_{\sigma} m b_n(x, \cdot)|_{C_\alpha}^{2q} |t - s|^q \, d\sigma
$$

$$
\lesssim |t - s|^q |b_n|_{B_n^{\theta + d/(2q)}}^{2q} \lesssim |t - s|^q,
$$

where the first estimate follows from the embedding $B_n^{\theta + d/(2q)} \hookrightarrow C^0$, the second from (23), the third from the definition of $|\cdot|_{B_n^{\theta + d/(2q)}}$, and the fourth from Lemma 5.5 since $b_n$ has uniformly bounded support and $\sup_{n \geq 1} |b_n|_{C^{0, \kappa}} < \infty$ with $\alpha > \theta + d/(2q)$.

The second estimate follows in a similar way from Proposition 2.14 upon using the bound

$$
\mathbb{E}_{\lambda_n} \left[ |\Delta_{x, \theta} m_{\alpha} W_n(s, t; x, x')|^{q} \right]^{1/q} \lesssim |\Delta_{x, \theta} m b_n(x, \cdot)|_{C_\alpha} |\Delta_{\theta} m b_n(x', \cdot)|_{C_\alpha} |t - s|
$$

and the argument from [14, p. 4089-4090] (note that this is where we require $\sup_{n \geq 1} |b_n|_{C^{0, \kappa}} < \infty$ for $\alpha > \theta + d/q$, so that $\sup_{n \geq 1} |b_n|_{B_n^{\theta + d/q}} < \infty$).

**Lemma 4.7 (Tightness).** Suppose that part (f) of Assumption 2.12 holds and that the support of $b_n$ is uniformly bounded. Then, for any $p > 2$, it holds that

$$
\sup_{n \geq 1} \mathbb{E}_{\lambda_n}[\|W_n\|_{p, \text{var}}^{q}] < \infty
$$

and that $(W_n)_{n \geq 1}$ is a family of tight random variables in $D^{0, p, \text{var}}(B)$.

**Proof.** Consider $\theta' \in (\theta, \alpha - \frac{d}{q})$ and the space $B' = C^{0, \theta'}(\mathbb{R} \times \mathbb{R}^d)$, where $\mathbb{R} \times \mathbb{R}$ contains the support of all $b_n$. Considering $W_n$ as an element of $D^{0, p', \text{var}}(B')$ for some $p' \in (2, p)$, it follows from Lemma 1.6 and 1.8 in Proposition 3.10 (applied to these new parameters) that $\sup_{n \geq 1} \mathbb{E}_{\lambda_n}[\|W_n\|_{p', \text{var}}^{q}] < \infty$. Observe that, for $R > 0$ and a compact subset $K$ of the classical $(J_1)$ Skorokhod space $D([0, 1], B \oplus B^{\otimes 2})$, by the interpolation estimate [14], the set

$$
\{ X \in K \mid \|X\|_{p', \text{var}} \leq R \}
$$

is compact in $D^{p, \text{var}}(B)$. Since the embedding $B' \hookrightarrow B$ is compact, we can furthermore use the bound [10] and the Arzelà–Ascoli theorem for the Skorokhod space to conclude that $(W_n)_{n \geq 1}$ is a family of tight random variables in $D^{0, p, \text{var}}(B)$. 

For an element $\pi \in L(B, \mathbb{R}^m)$ and $b \in C^{0, \kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$, write $\pi b : M \to \mathbb{R}^m$ for the function $y \mapsto \pi(b(\cdot, y))$. A direct verification shows that $|\pi b|_{C^0} \leq |\pi|_{L(B, \mathbb{R}^m)} |b|_{C^{0, \kappa}}$ (see, e.g., the proof of [14, Lem. 5.12]).
Consider the subspace of \( L(B, \mathbb{R}) \)
\[
\widetilde{L}(B, \mathbb{R}) := \text{span} \{ b \mapsto D^k b^j(x) \mid x \in \mathbb{R}^d, k \in \mathbb{N}^d, |k| \leq 1, j \in \{1, \ldots, d\} \}.
\]
For \( m \geq 1 \), we denote by \( \tilde{L}(B, \mathbb{R}^m) \) the subspace of \( \pi \in L(B, \mathbb{R}^m) \) such that \( \pi^i \in \tilde{L}(B, \mathbb{R}) \) for every \( i = 1, \ldots, m \). We note that \( \tilde{L}(B, \mathbb{R}) \) does not appear in the work [14], however, due to the generality of our setting, we find it more convenient to work with than the full space \( L(B, \mathbb{R}) \).

Observe that, for \( b \in C^0_n(\mathbb{R}^d) \), the map \( x \mapsto b(x, \cdot) \) is a \( C^0 \) map from \( \mathbb{R}^d \) into the closed subspace \( C^0_n(M) \), and thus \( \pi b \in C^0_n(M) \) for all \( \pi \in \tilde{L}(B, \mathbb{R}) \).

**Lemma 4.8** (Finite-dimensional projections). Let \( \pi \in \tilde{L}(B, \mathbb{R}^m) \) for some \( m \geq 1 \) and suppose that Assumption 2.11 holds. Let \( \mathfrak{B} \) be defined as in Proposition 2.13 and let \( W_\pi \) be an \( \mathbb{R}^m \)-valued Brownian motion with covariance
\[
\mathbb{E}[W^i(1)W^j(1)] = \mathfrak{B}(\pi^i \beta_\infty, \pi^j \beta_\infty) + \mathfrak{B}(\pi^i \beta_\infty, \pi^j \beta_\infty).
\]
Define further
\[
\tilde{W}^i_\pi(t) = \int_0^t W^i_\pi \, dW^j_\pi + \mathfrak{B}(\pi^i \beta_\infty, \pi^j \beta_\infty)t.
\]
Then, as \( n \to \infty \),
\[
(\pi W_n, (\pi \circ \pi) \tilde{W}_n) \to_{\lambda_n} (W_\pi, \tilde{W}_\pi)
\]
in the sense of finite-dimensional distributions.

**Proof.** By the preceding remarks, \( (\pi b_n)_{n \in \mathbb{N} \cup \{\infty\}} \) is a \( C^0_n(M, \mathbb{R}^m) \)-family, and the conclusion follows by Assumption 2.12 and Proposition 2.15. □

The convergence of finite-dimensional distributions, together with tightness, allows us to establish uniqueness of weak limit points (which we note settles a point of ambiguity in [14, Rem. 5.14]).

**Proposition 4.9.** Suppose that Assumptions 2.11 and 2.12 hold. Then there exists a unique random variable in \( D^{0, \beta_{\text{var}}}(A) \times D^{0, \rho_{\text{var}}}(B) \) such that \( (V_n, W_n) \to_{\lambda_n} (V, W) \). Furthermore, \( (V, W) \) is a.s. continuous.

**Proof.** By Lemma 4.3, \( \|V_n - V\|_{\beta_{\text{var}}} \to_{\lambda_n} 0 \), where \( V \) is deterministic and continuous. It remains to show that \( W_n \) converges to a unique weak limit point \( W \) which is a.s. continuous. By Lemma 4.7, \( (W_n)_{n \geq 1} \) is tight. Hence, by Prokhorov’s theorem and the fact that \( D^{0, \rho_{\text{var}}}(B) \) is Polish, \( (W_n)_{n \geq 1} \) is relatively compact. Let \( W \) and \( \tilde{W} \) be weak limit points of \( W_n \). Since the largest jump of \( W_n \) is of the order \( n^{-1/2} \) and the largest jump of \( t \mapsto \tilde{W}_n(0, t) \) is of the order \( n^{-1/2} \sup_{t \in [0, 1]} |W_n(t)| \),
\[
\begin{align*}
\text{it follows that} & \quad W \text{ is a.s. continuous (and likewise for } \tilde{W}). \\
\text{We now show that} & \quad W \text{ and } \tilde{W} \text{ have the same law. Consider the collection of } \mathbb{R}\text{-valued functions on } D^{0, \rho_{\text{var}}}(B) \\
& \quad \mathcal{F} := \left\{ w \mapsto \sum_{j=1}^k \pi_j w(t_j) \mid k \geq 1, \pi_1, \ldots, \pi_k \in \tilde{L}(B, \mathbb{R}) , t_1, \ldots, t_k \in [0, 1] \right\}.
\end{align*}
\]
For any \( f \in \mathcal{F} \), it follows from Lemma 4.3 that \( f(W) \) and \( f(\tilde{W}) \) have the same law. In particular, \( \mathbb{E}[e^{i f(W)}] = \mathbb{E}[e^{i f(\tilde{W})}] \) for all \( f \in \mathcal{F} \). However, note that the collection of \( \mathbb{C}\text{-valued functions } \bar{F} := \{ w \mapsto e^{i f(w)} \mid f \in \mathcal{F} \} \) is a unital algebra of bounded functions on \( D^{0, \rho_{\text{var}}}(B) \) which separates points and is closed under conjugation. Moreover, every \( f \in \bar{F} \) is continuous on the subspace of continuous
paths in $D^{0,p\text{-var}}(B)$, and in particular on the support of $W$ and $\tilde{W}$. It follows by the Stone–Weierstrass theorem and a compactification argument (see, e.g., [2 Ex. 7.14.79]) that $W$ and $\tilde{W}$ have the same law.

Finally, we have the characterization of the RDE driven by $(V, W)$ as the solution to an SDE. We flesh out the abstract statement in the following lemma.

**Lemma 4.10.** Let $X$ be the solution to the RDE
\[ dX = F(X)\bar{a} dt + H(X)\,dW, \quad X(0) = \xi \in \mathbb{R}^d, \]
where $\bar{a} \in A$ is fixed and $W = (W, \mathcal{W})$ is a random $p$-rough path over $B$, $p < \theta$. Suppose that, for all $m \geq 1$ and $\pi \in \mathcal{L}(B, \mathbb{R}^m)$,
\[ (\pi W, (\pi \otimes \pi)\mathcal{W}) \sim (W_\pi, \mathcal{W}_\pi) \]
in the sense of finite dimensional distributions, where $W_\pi$ is an $\mathbb{R}^m$-valued Brownian motion with covariance
\[ \Sigma_{\pi}^{ij} := \mathbb{E}[W_\pi^i(1)W_\pi^j(1)] \]
and
\[ \mathcal{W}_\pi(t) = \int_0^t W_\pi^i \, dW_\pi^j + \Gamma_\pi^{ij}t. \]

For every $x \in \mathbb{R}^d$, let us define $\Sigma(x) := \Sigma_{H(x)}$ and, for $i = 1, \ldots, d$,
\[ \Gamma^i(x) := \sum_{k=1}^d H^{k(i)}_{H(x) \otimes DH(x)} , \]
where $H(x) \oplus DH(x) \in \mathcal{L}(B, \mathbb{R}^d \oplus (\mathbb{R}^d)^* \otimes \mathbb{R}^d)$. Suppose further that
\[ \sup_{x \in \mathbb{R}^d} \sum_{i=1}^d |\Sigma_{\pi}^{ii}(x)| + |\Gamma^i(x)| < \infty . \]

Then $X$ solves the martingale problem associated with $\mathcal{L} = (\bar{a} + \Gamma)D + \frac{1}{2}\Sigma D^2$.

**Proof.** Let $\{\mathcal{F}_t\}_{t \in [0,1]}$ denote the filtration generated by the finite-dimensional projections of $W$. We first show that $M : [0,1] \to \mathbb{R}^d$ is a martingale with respect to $\mathcal{F}$, where
\[ M(t) := X(t) - \int_0^t \bar{a}(X(s)) \, ds - \int_0^t \sum_{k=1}^d \Gamma(X(s)) \, ds , \]
with quadratic variation
\[ [M^i, M^j]_t = \int_0^t \Sigma_{\pi}^{ij}(X(s)) \, ds . \]

Indeed, the definition of the rough integral readily implies that $X$ and $M$ are adapted to $\mathcal{F}$ (cf. [14, Lem 6.3]). Furthermore, for fixed $0 \leq s < t \leq 1$, we have
\[ M(t) - M(s) = \int_s^t H(X(u)) \, dW(u) - \int_s^t \Gamma(X(u)) \, du \]
\[ \quad = \lim_{|\mathcal{P}| \to 0} \sum_{[u,v] \in \mathcal{P}} M^P_{[u,v]} , \]
where the limit is taken over partitions $\mathcal{P}$ of $[s,t]$, and
\[ M^P_{[u,v]} := H(X(u))W(u,v) + (H \otimes DH)(X(u))\mathcal{W}(u,v) - \Gamma(X(u))(v - u) . \]
Note that the same argument as in [14, Lem. 6.2] implies that $\pi W(u, v)$ and $(\pi \otimes \pi)\mathbb{W}(u, v)$ are independent of $F_u$ for any $\pi \in \overline{L}(\mathcal{B}, \mathbb{R}^m)$. Taking $\pi(x) = H(x) \otimes DH(x)$ in (24), it follows that
$$
\mathbb{E}[M_{\pi, v}^P] | F_u = 0.
$$
Furthermore, for $i, j = 1, \ldots, d$,
$$
\mathbb{E}[H^i(X(u))W(u, v)H^j(X(u))W(u, v) | F_u] = \Sigma^{ij}(X(u))(v - u)
$$
and, by Itô’s isometry,
$$
\mathbb{E}[(\Gamma \otimes \Gamma)(X(u))\mathbb{W}(u, v) - \Gamma(X(u))(v - u)^2 | F_u] \lesssim |v - u|^2,
$$
where the proportionality constant depends only on $\Sigma(X(u))$. Using the bound (25), it follows that $M$ is a martingale with quadratic variation (26) as claimed.

Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a smooth, compactly supported function. Since $[X] = [M]$, by Itô's formula,
$$
\varphi(X(t)) = \varphi(X(s)) + \int_s^t D\varphi(X(u)) \, dX(u) + \frac{1}{2} \int_s^t D^2\varphi(X(u)) \, d[M](u),
$$
from which it follows that
$$
\varphi(X(t)) - \varphi(X(s)) - \int_s^t \left[ D\varphi(\bar{\alpha} + \Gamma) + \frac{1}{2} D^2\varphi \Sigma \right] (X(u)) \, du
$$
is a martingale.

**Proof of Theorem 4.3.** The fact that $(V_n, W_n) \to_{\lambda_n} (V, W)$, where $(V, W)$ is a.s. continuous, follows from Proposition 4.9. By Lemma 4.8 $W$ satisfies assumption (24) of Lemma 4.10 with $\Sigma^{ij} = \mathcal{B}(\pi^ib, \pi^ib) + \mathcal{B}(\pi^ib, \pi^ib)$ and $\Gamma^{ij} = \mathcal{B}_2(\pi^ib, \pi^ib)$. In particular, $\Gamma$ in Lemma 4.10 is given by $\Gamma(x) = \sum_{k=1}^{\infty} \mathcal{B}_2(b^k(x, \cdot), \partial_0 b^k(x, \cdot))$. Furthermore, $\mathcal{B} = \frac{1}{2} \mathcal{B}_1 + \mathcal{B}_2$ is bounded by part (ii) of Assumption 2.12 so $\Sigma^{ii}(x) \lesssim |b(x, \cdot)|_{C^{\alpha}} \leq |b|_{C^{\alpha, \kappa}} \lesssim |b|_{C^{\alpha, \kappa}}$. Hence all the assumptions of Lemma 4.10 are verified, and the conclusion follows from [21] Thm. 4.5.3 by the equivalence of weak solutions to SDEs and the martingale problem.

**Proof of Theorem 2.17.** This follows from Corollary 1.4 and the exact same localization argument as in [14, Sec. 7].

### 5. Continuous-time dynamics revisited

In this section, we show how the results of the Section 2 extend to the case of continuous-time dynamics. In particular, we extend the results of [14] to include optimal moment assumptions and families of dynamical systems. Since the arguments are very similar to those of the discrete-time case (and the setting is similar to that of [14]), we omit the proofs and only state the main results.

Consider a compact Riemannian manifold $M$ with Riemannian distance $\rho$. Recall the function spaces defined in Definition 2.4 and fix parameters $q > 1$, $\kappa, \bar{\kappa} \in (0, 1)$, and $\alpha > 2 + \frac{q}{\kappa}$. Let $a_\varepsilon \in C^{1+\kappa, \bar{\kappa}}(\mathbb{R}^d \times M, \mathbb{R}^d)$ and $b_\varepsilon, b_0 \in C^{\alpha, \kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$, for $\varepsilon \in (0, 1)$, such that
$$
\sup_{\varepsilon \in (0, 1)} |a_\varepsilon|_{C^{1+\kappa, 0}} + |b_\varepsilon|_{C^{\alpha, \kappa}} < \infty, \quad \lim_{\varepsilon \to 0} |b_\varepsilon - b_0|_{C^{\alpha, \kappa}} = 0.
$$
We consider the fast-slow systems of ODEs posed on $\mathbb{R}^d \times M$

$$\frac{d}{dt} x_\varepsilon = a_\varepsilon(x_\varepsilon, y_\varepsilon) + \varepsilon^{-1} b_\varepsilon(x_\varepsilon, y_\varepsilon), \quad \frac{d}{dt} y_\varepsilon = \varepsilon^{-2} g_\varepsilon(y_\varepsilon),$$

where $g_\varepsilon : M \to TM$ is a Lipschitz vector field. As before, the initial condition $x_\varepsilon(0) = \xi_\varepsilon \in \mathbb{R}^d$ is deterministic, and $y_\varepsilon(0)$ is drawn randomly from a Borel probability measure $\lambda_\varepsilon$ on $M$.

We now give the analogues of Assumptions 2.11 and 2.12 for the current setting.

**Assumption 5.1.** There exists $a \in C^{1+\kappa}(\mathbb{R}^d, \mathbb{R}^d)$ such that

$$\sup_{t \in [0,1]} |V_\varepsilon(t) - \dot{a}|_{C^{1+\kappa}} \to 0 \quad \text{as} \quad \varepsilon \to 0,$$

where $V_\varepsilon(t) = \int_0^t a_\varepsilon(\cdot, y_\varepsilon(s)) \, ds$.

Let $g_{\varepsilon,t}$ denote the flow generated by the vector field $g_\varepsilon$. Given $v, w \in C^\kappa(M, \mathbb{R}^m)$ and $0 \leq s \leq t \leq 1$, we define $W_{v,\varepsilon}(t) \in \mathbb{R}^m$ and $\mathbb{W}_{v,w,\varepsilon}(s, t) \in \mathbb{R}^{m \times m}$ by

$$W_{v,\varepsilon}(t) = \varepsilon \int_0^{t^\varepsilon} v \circ g_{\varepsilon,s} \, ds, \quad \mathbb{W}_{v,w,\varepsilon}(s, t) = \int_s^t (W_{v,\varepsilon}(r) - W_{v,\varepsilon}(s)) \otimes dW_{w,\varepsilon}(r).$$

As before, we write simply $\mathbb{W}_{v,\varepsilon}$ for $\mathbb{W}_{v,v,\varepsilon}$.

Recall our notational convention about subspaces $C^\kappa_\varepsilon(M)$ of $C^\kappa(M)$ introduced before Assumption 2.3. As in Section 2.2, given a family of subspaces $(C^\kappa_\varepsilon(M))_{\varepsilon \in (0,1]}$ of $C^\kappa(M)$, we call $v = (v_\varepsilon)_{\varepsilon \in [0,1]}$ a $C^\kappa_\varepsilon(M, \mathbb{R}^m)$-family if $v_\varepsilon \in C^\kappa_\varepsilon(M, \mathbb{R}^m)$ and $\lim_{\varepsilon \to 0} |v_\varepsilon - v_0|_{C^\kappa} = 0$.

**Assumption 5.2.** There exists a closed subspace $C^\kappa_\varepsilon(M)$ of $C^\kappa(M)$ for each $\varepsilon \in [0,1]$ such that $b_\varepsilon \in C^\kappa_{\varepsilon \kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$ and such that

(i) for all $v = (v_\varepsilon), w = (w_\varepsilon) \in \prod_{\varepsilon \in (0,1]} C^\kappa_\varepsilon(M)$ with

$$\sup_{\varepsilon \in [0,1]} |v_\varepsilon|_{C^\kappa} + |w_\varepsilon|_{C^\kappa} < \infty,$$

there exists $K = K_{v,w,q} > 0$ such that for all $0 \leq s \leq t \leq 1$ and $\varepsilon > 0$,

$$|W_{v,\varepsilon}(s, t)|_{L^2(\lambda_\varepsilon)} \leq K |t - s|^{1/2}, \quad |\mathbb{W}_{v,w,\varepsilon}(s, t)|_{L^2(\lambda_\varepsilon)} \leq K |t - s|.$$

(ii) There exists a bounded bilinear operator $\mathfrak{B} : C^\kappa_0(M) \times C^\kappa_0(M) \to \mathbb{R}$ such that for every $m \geq 1$ and every $C^\kappa_\varepsilon(M, \mathbb{R}^m)$-family $v = (v_\varepsilon)_{\varepsilon \in (0,1]}$, it holds that $\langle W_{v_\varepsilon,\varepsilon}, \mathbb{W}_{v,\varepsilon} \rangle \to \lambda_\varepsilon \langle W_\varepsilon, \mathbb{W}_\varepsilon \rangle$ as $\varepsilon \to 0$ in the sense of finite-dimensional distributions, where $W_\varepsilon$ is an $\mathbb{R}^m$-valued Brownian motion and

$$\mathbb{W}_\varepsilon^{ij}(t) = \int_0^t W_\varepsilon^i \, dW_\varepsilon^j + \mathfrak{B}(v_0^i, v_0^j) t.$$

**Remark 5.3.** As in Remark 2.13 under the assumption that $\lambda_\varepsilon = g_{\varepsilon,t}$-stationary, the simpler bounds

$$|W_{v,\varepsilon'}(1)|_{L^2(\lambda_\varepsilon)} \leq K \quad \text{and} \quad |\mathbb{W}_{v,w,\varepsilon'}(0,1)|_{L^2(\lambda_\varepsilon)} \leq K \quad \text{for all} \quad \varepsilon, \varepsilon' \in (0,1]$$

imply part (i) of Assumption 5.2.

**Remark 5.4.** As in Proposition 2.17 one can show that Assumption 5.2 implies that the covariance of $W_\varepsilon$ is given by

$$\mathbb{E}[W_\varepsilon^i(1)W_\varepsilon^j(1)] = \mathfrak{B}(v_0^i, v_0^j) + \mathfrak{B}(v_0^i, v_0^j).$$
Furthermore, as in Section 2.1, if \( a_\varepsilon, b_\varepsilon, T_\varepsilon, \lambda_\varepsilon \) do not depend on \( \varepsilon \), then one can drop the condition that \( B \) is bounded in Assumption 5.2 since this follows automatically (see [14, Prop. 2.8]).

Consider the quadratic form
\[
\Sigma^{ij}(x) = \mathfrak{A}(b^j_\varepsilon(x, \cdot), b^i_\varepsilon(x, \cdot)) + \mathfrak{A}(b^i_\varepsilon(x, \cdot), b^j_\varepsilon(x, \cdot)), \quad i, j = 1, \ldots, d.
\]
By the same argument as Lemma 2.9, \( \Sigma \) is positive semi-definite and the unique positive semi-definite \( \sigma \) satisfying \( \sigma^2 = \Sigma \) is Lipschitz. In particular, as before, there is a unique (strong) solution to the SDE
\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB
\]
for any Lipschitz \( \tilde{a} : \mathbb{R}^d \to \mathbb{R}^d \).

The following is the main result of this section, the proof of which we omit since it requires only minor changes to that of Theorem 2.17.

**Theorem 5.5.** Suppose that Assumptions 5.1 and 5.2 hold, and that \( \xi_\varepsilon \xrightarrow{\varepsilon} \xi \in \mathbb{R}^d \). Then \( x_\varepsilon \xrightarrow{\varepsilon} \lambda_\varepsilon X \) in the uniform topology as \( \varepsilon \to 0 \), where \( X \) is the unique weak solution of the SDE
\[
dX = \tilde{a}(X) \, dt + \sigma(X) \, dB, \quad X(0) = \xi.
\]
Here, \( B \) is a standard Brownian motion in \( \mathbb{R}^d \), \( \sigma \) is the unique positive semi-definite square root of \( \Sigma \) given by (27), and \( \tilde{a} \) is the Lipschitz function given by
\[
\tilde{a}^i(x) = \bar{a}^i(x) + \sum_{k=1}^d \mathfrak{A}(b^k_\varepsilon(x, \cdot), \partial_k b^i_\varepsilon(x, \cdot)), \quad i = 1, \ldots, d.
\]

**Appendix A. Rough path Besov-variation embedding**

We adapt Friz–Victoir [8, 9] in proving some variants of a Besov-variation embedding, applicable in an infinite-dimensional rough path setting. Let \( \mathcal{B} \) be a Banach space and equip \( \mathcal{B}^{\otimes 2}, \ldots, \mathcal{B}^{\otimes N} \) with a system of admissible tensor norms. For a continuous multiplicative function \( W = (1, W^1, \ldots, W^N) : [0, T]^2 \to \bigoplus_{k=0}^N \mathcal{B}^{\otimes k} \) define the homogeneous Besov norm
\[
\|W\|_{W^{\alpha,q};[s,t]} := \sum_{k=1}^N \int_{[s,t]^2} \frac{|W^{q/k}_{s,u} W^{q/k}_{u,v}|^{q/k}}{|u-v|^{q\alpha+1}} \, du \, dv.
\]

**Proposition A.1.** Suppose \( q > 1 \) and \( \alpha \in \left(\frac{1}{q}, 1\right) \). There exists a constant \( C = C(\alpha, q, N) \) such that
\[
\sum_{k=1}^N |W^{k}_{s,t}|^{q/k} \leq C|t-s|^{q\alpha-1} \|W\|_{W^{\alpha,q};[s,t]}^q.
\]

**Proof.** We follow a similar strategy to [9, Proposition A.9]. We proceed by induction on \( N \). The case \( N = 1 \) follows directly from the GRR lemma [9, Corollary A.2]. Suppose the result is true for \( N - 1 \). Since both sides scale homogeneously with dilations, we may suppose that \( \|W\|_{W^{\alpha,q};[s,t]}^q \leq 1 \). Let us write \( \alpha - \frac{1}{q} =: 1/p \). All double integrals in the sequel are taken over \([s,t]^2\), and \( C \) denotes an unimportant positive constant which may change from line to line.
Define \( \Upsilon_{s,t} = \sup_{u,v \in [s,t]} |W_{u,v}|_{\frac{p}{q \alpha + 1} \cdot \frac{q}{p}} \), and observe that it suffices to show \( \Upsilon_{s,t} \leq C \). We have

\[
W_{s,v}^N - W_{s,u}^N = W_{u,v}^N + \sum_{j=1}^{N-1} W_{s,u}^{N-j} \otimes W_{u,v}^j ,
\]

and thus

\[
\left( \int \int \frac{|W_{s,u}^N|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q} \leq \Delta_1 + \Delta_2 ,
\]

where

\[
\Delta_1 = \sum_{j=1}^{N-1} \left( \int \int \frac{|W_{s,u}^{N-j}|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q},
\]

\[
\Delta_2 = \left( \int \int \frac{|W_{u,v}^j|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q} .
\]

For \( \Delta_1 \), by the inductive hypothesis, we have

\[
|W_{s,u}^{(N-j)}|^q \leq |t-s|^{q(N-j)/p},
\]

so that

\[
\Delta_1 \leq \sum_{j=1}^{N-1} |t-s|^{(N-j)/p} \left( \int \int \frac{|W_{u,v}^j|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q}.
\]

Again by the inductive hypothesis, we have

\[
|W_{u,v}^j|^q(1-1/j) \leq |t-s|^{q(j-1)/p},
\]

so that

\[
\Delta_1 \leq \sum_{j=1}^{N} |t-s|^{(N-1)/p} \left( \int \int \frac{|W_{u,v}^j|^q/j}{|u-v|^{q\alpha + 1}} \, du \, dv \right) \leq \sum_{j=1}^{N} |t-s|^{(N-1)/p} .
\]

For \( \Delta_2 \), we have

\[
\Delta_2 \leq \left( \int \int \Upsilon_{s,t}^{q(1-1/N)} |t-s|^{q(N-1)/p} \frac{|W_{s,u}^N|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q} \leq \Upsilon_{s,t}^{q(1-1/N)} |t-s|^{(N-1)/p} .
\]

Combining the above two estimates, we have

\[
\left( \int \int \frac{|W_{s,u}^N|^q}{|u-v|^{q\alpha + 1}} \, du \, dv \right)^{1/q} \leq C |t-s|^{(N-1)/p}(1 + \Upsilon_{s,t}^{1-1/N}) .
\]

Applying the GRR lemma to the continuous path \( W_{s,t}^N : [s,t] \to B_{\gamma}^\oplus \) we have

\[
|W_{s,t}^N| \leq C |t-s|^{1/p} |t-s|^{(N-1)/p}(1 + \Upsilon_{s,t}^{1-1/N}) \leq C |t-s|^{N/p}(1 + \Upsilon_{s,t}^{1-1/N}) .
\]

Finally, note that the above argument applies to any interval \([s',t'] \subset [s,t]\). It follows that

\[
\Upsilon_{s,t} \leq C(1 + \Upsilon_{s,t}^{1-1/N}) ,
\]

and thus \( \Upsilon_{s,t} \leq C \) as desired. \( \square \)

Recall the homogeneous \( \gamma \)-Hölder “norm” for \( \gamma \in (0,1] \)

\[
\|W\|_{\gamma,\text{Hölder}} := \sum_{k=1}^{N} \sup_{u,v \in [s,t]} \frac{|W_{u,v}^k|^{1/k}}{|u-v|^\gamma} .
\]
Corollary A.2. Let $q > 1$ and $\alpha \in (\frac{1}{q}, 1)$. There exists a constant $C = C(\alpha, q, N)$ such that
\[
\|W\|_{(\alpha-1/q)\text{-Hölder};[s,t]} \leq C\|W\|_{W^{\alpha,q};[s,t]}.
\]
Proof. Immediate from Proposition [A.1] \qed

Recall the homogeneous $p$-variation “norm” for $p \geq 1$
\[
\|W\|_{p\text{-var};[s,t]} := \sup_P \sum_{u,v} \sum_{k=1}^N |W^k_{u,v}|^{p/k},
\]
where the supremum runs over all partitions $P$ of $[s,t]$.

Corollary A.3. Let $q > 1$ and $\alpha \in (\frac{1}{q}, 1)$. There exists a constant $C = C(\alpha, q, N)$ such that
\[
\|W\|_{1/\alpha\text{-var};[s,t]} \leq C|t-s|^{(\alpha-1)/q}\|W\|_{W^{\alpha,q};[s,t]}.
\]
Proof. By Proposition [A.1] we have for all $u,v \in [s,t]$ and $k = 1,\ldots,N$
\[
|W^k_{u,v}|^{1/\alpha} = \left(\frac{|W^k_{u,v}|^{q/k}}{|u-v|^{\alpha-1}}\right)^{1/\alpha} \leq C \left(\frac{|u-v|^{\alpha-1}}{|u-v|^{\alpha-1}}\right)^{1/\alpha} \left(\|W\|_{W^{\alpha,q};[u,v]}\right)^{1/\alpha}.
\]
Note however that $\omega_1(u,v) = |u-v|$ and $\omega_2(u,v) := \|W\|_{W^{\alpha,q};[u,v]}$ are controls, and thus so is $\omega := \omega_1^{1-\frac{1}{\alpha}}\omega_2^{\frac{1}{\alpha}}$. Hence
\[
\|W\|_{1/\alpha\text{-var};[s,t]} \leq \omega(s,t),
\]
from which the conclusion follows. \qed

Remark A.4. Besov (rough path) regularity effectively interpolates between the well-known Hölder- and $p$-variation cases, see [6] for a discussion.

References

[1] I. Bailleul and R. Catellier. “Rough flows and homogenization in stochastic turbulence”. J. Differential Equations 263.8 (2017), pp. 4894–4928. url: https://doi.org/10.1016/j.jde.2017.06.006

[2] V. I. Bogachev. Measure theory. Vol. I, II. Berlin: Springer-Verlag, 2007, Vol. I: xviii+500 pp., Vol. II: xiv+575. url: http://dx.doi.org/10.1007/978-3-540-34514-5

[3] I. Chevyrev. “Random walks and Lévy processes as rough paths”. Probab. Theory Related Fields 170.3-4 (2018), pp. 891–932. url: https://doi.org/10.1007/s00440-017-0781-1

[4] I. Chevyrev and P. K. Friz. “Canonical RDEs and general semimartingales as rough paths”. Ann. Probab. 47.1 (2019), pp. 420–463.

[5] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. “Multiscale systems, homogenization, and rough paths”. ArXiv e-prints (Dec. 2017). To appear in Proceedings of the Conference in Honor of the 75th Birthday of S.R.S. Varadhan. arXiv: [1712.01343 [math.DS]]

[6] P. K. Friz and D. J. Prömel. “Rough path metrics on a Besov-Nikolskii-type scale”. Trans. Amer. Math. Soc. 370.12 (2018), pp. 8521–8550. url: https://doi.org/10.1090/tran/7264
[7] P. K. Friz and A. Shekhar. “General rough integration, Lévy rough paths and a Lévy–Kintchine-type formula”. *Ann. Probab.* 45.4 (July 2017), pp. 2707–2765. URL: [http://dx.doi.org/10.1214/16-AOP1123](http://dx.doi.org/10.1214/16-AOP1123)

[8] P. K. Friz and N. Victoir. “A variation embedding theorem and applications”. *J. Funct. Anal.* 239.2 (2006), pp. 631–637. URL: [https://doi.org/10.1016/j.jfa.2005.12.021](https://doi.org/10.1016/j.jfa.2005.12.021)

[9] P. K. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths*. Vol. 120. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 2010, pp. xiv+670.

[10] P. K. Friz and H. Zhang. “Differential equations driven by rough paths with jumps”. *J. Differential Equations* 264.10 (2018), pp. 6226–6301. URL: [https://doi.org/10.1016/j.jde.2018.01.031](https://doi.org/10.1016/j.jde.2018.01.031)

[11] G. A. Gottwald and I. Melbourne. “Homogenization for deterministic maps and multiplicative noise”. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 469.2156 (2013), pp. 20130201, 16. URL: [http://dx.doi.org/10.1098/rspa.2013.0201](http://dx.doi.org/10.1098/rspa.2013.0201)

[12] D. Kelly. “Rough path recursions and diffusion approximations”. *Ann. Appl. Probab.* 26.1 (2016), pp. 425–461. URL: [http://dx.doi.org/10.1214/15-AAP1096](http://dx.doi.org/10.1214/15-AAP1096)

[13] D. Kelly and I. Melbourne. “Smooth approximation of stochastic differential equations”. *Ann. Probab.* 44.1 (Jan. 2016), pp. 479–520. URL: [http://dx.doi.org/10.1214/14-AOP979](http://dx.doi.org/10.1214/14-AOP979)

[14] D. Kelly and I. Melbourne. “Deterministic homogenization for fast-slow systems with chaotic noise”. *Journal of Functional Analysis* 272.10 (2017), pp. 4063–4102. URL: [http://www.sciencedirect.com/science/article/pii/S0022123617300162](http://www.sciencedirect.com/science/article/pii/S0022123617300162)

[15] A. Korepanov, Z. Kosloff, and I. Melbourne. “Martingale-coboundary decomposition for families of dynamical systems”. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 35.4 (2018), pp. 859–885. URL: [https://doi.org/10.1016/j.anihpc.2017.08.005](https://doi.org/10.1016/j.anihpc.2017.08.005)

[16] A. Korepanov, Z. Kosloff, and I. Melbourne. In preparation.

[17] C. Liverani, B. Saussol, and S. Vaienti. “A probabilistic approach to intermittency”. *Ergodic Theory Dynam. Systems* 19.3 (1999), pp. 671–685. URL: [https://doi.org/10.1017/S0143385799133856](https://doi.org/10.1017/S0143385799133856)

[18] I. Melbourne and A. M. Stuart. “A note on diffusion limits of chaotic skew-product flows”. *Nonlinearity* 24.4 (2011), pp. 1361–1367. URL: [http://dx.doi.org/10.1088/0951-7715/24/4/018](http://dx.doi.org/10.1088/0951-7715/24/4/018)

[19] Y. Pomeau and P. Manneville. “Intermittent transition to turbulence in dissipative dynamical systems”. *Comm. Math. Phys.* 74.2 (1980), pp. 189–197. URL: [http://projecteuclid.org/euclid.cmp/1103907981](http://projecteuclid.org/euclid.cmp/1103907981)

[20] S. Smale. “Differentiable dynamical systems”. *Bull. Amer. Math. Soc.* 73 (1967), pp. 747–817. URL: [https://doi.org/10.1090/S0002-9904-1967-11798-1](https://doi.org/10.1090/S0002-9904-1967-11798-1)

[21] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Reprint of the 1997 edition. Springer-Verlag, Berlin, 2006, pp. xii+338.
[22] L-S. Young. “Statistical properties of dynamical systems with some hyperbolicity”. *Ann. of Math. (2)* 147.3 (1998), pp. 585–650. URL: https://doi.org/10.2307/120960

[23] L-S. Young. “Recurrence times and rates of mixing”. *Israel J. Math.* 110 (1999), pp. 153–188. URL: https://doi.org/10.1007/BF02808180

I. Chevyrev, Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, United Kingdom
E-mail address: chevyrev@maths.ox.ac.uk

P.K. Friz, Institut für Mathematik, Technische Universität Berlin, and Weierstrass–Institut für Angewandte Analysis und Stochastik, Berlin, Germany
E-mail address: friz@math.tu-berlin.de

A. Korepanov, Department of Mathematics, University of Exeter, Exeter, EX4 4QF, United Kingdom
E-mail address: a.korepanov@exeter.ac.uk

I. Melbourne, Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom
E-mail address: i.melbourne@warwick.ac.uk

H. Zhang, Institute of Mathematics, Fudan University, Shanghai, 200433, China
E-mail address: huilinzhang2014@gmail.com