

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/154936>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2.

I. Chevyrev\*    P. K. Friz†    A. Korepanov‡    I. Melbourne§    H. Zhang¶

10 June 2020. Updated 24 June 2021.

## Abstract

We consider deterministic homogenization for discrete-time fast-slow systems of the form

$$X_{k+1} = X_k + n^{-1}a_n(X_k, Y_k) + n^{-1/2}b_n(X_k, Y_k), \quad Y_{k+1} = T_n Y_k$$

and give conditions under which the dynamics of the slow equations converge weakly to an Itô diffusion  $X$  as  $n \rightarrow \infty$ . 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.

Combined with parallel developments on martingale approximations for families of nonuniformly expanding maps in Part 1 by Korepanov, Kosloff and Melbourne, we obtain optimal homogenization results when  $T_n$  is such a family of maps.

## 1 Introduction

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

$$X_{k+1}^{(n)} = X_k^{(n)} + n^{-1}a_n(X_k^{(n)}, Y_k^{(n)}) + n^{-1/2}b_n(X_k^{(n)}, Y_k^{(n)}), \quad Y_{k+1}^{(n)} = T_n Y_k^{(n)}, \quad (1.1)$$

where  $X_k^{(n)}$  takes values in  $\mathbb{R}^d$ ,  $Y_k^{(n)}$  takes values in a metric space  $\Lambda$ , and  $a_n, b_n : \mathbb{R}^d \times \Lambda \rightarrow \mathbb{R}^d$  and  $T_n : \Lambda \rightarrow \Lambda$  are suitable functions. The only source of randomness in the dynamics is the initial condition  $Y_0^{(n)}$  which we sample from a (not necessarily ergodic) probability measure  $\lambda_n$  on  $\Lambda$ .

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

$$dX = \tilde{a}(X) dt + \sigma(X) dB \quad (1.2)$$

---

\*School of Mathematics, University of Edinburgh, Edinburgh, EH9 3FD, UK.

†Institut für Mathematik, Technische Universität Berlin, and Weierstraß–Institut für Angewandte Analysis und Stochastik, Berlin, Germany

‡Mathematics Department, University of Exeter, Exeter, EX4 4QF, UK

§Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

¶Institute of Mathematics, Fudan University, Shanghai, 200433, China

with explicit formulae for the coefficients  $\tilde{a}$  and  $\sigma$ . Our assumptions on the system involve only moment bounds and a suitable (iterated) weak invariance principle on the fast dynamics  $T_n$ . In the companion paper, *Part 1* [19], it is shown how these assumptions can be verified for a large class of families  $T_n$  of nonuniformly hyperbolic dynamical systems. 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 (1.1) was initiated in [22], and has seen recent growth in a number of works, including [14, 16, 17]. See our survey paper [6] 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 should be compared to [16, 15] 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 [1, 17].

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 [16, 17, 1] in which only a subrange can be handled (specifically, in Assumptions 2.3(i) and 2.12(i) it now suffices that  $q > 1$  rather than  $q > 3$  as was the case previously; in particular the required control on ordinary moments is reduced from  $6 + \varepsilon$  to  $2 + \varepsilon$ ). In [6] we indicate a simplified version of this second contribution 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 [25], and to large classes of nonuniformly hyperbolic systems [27, 28]. A detailed account of discrete-time dynamical systems  $T$  for which our assumptions are verified can be found in [16, Sec. 10] and [17, 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 [18]. As mentioned above, in Part 1 [19], the assumptions in the current paper are verified for a range of families  $T_n$ .

The main tool in showing convergence of the system (1.1) is rough path theory [21], which we apply in the càdlàg setting in conjunction with the method in [17]. 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. [6, Sec. 3.2]). Our results employ the stability of “forward” (Itô) rough differential equations (RDEs) with jumps recently studied in [13], which we extend herein to the Banach space setting (though we restrict attention to the case of level-2 rough paths). The works [10, 4, 5] also study RDEs in the presence of jumps, but primarily focus on “geometric” (Marcus) notions of solution.

## 1.1 Illustrative example

Let  $\Lambda = [0, 1]$ . For  $\gamma \geq 0$ , we consider the intermittent map  $T: \Lambda \rightarrow \Lambda$ ,

$$Ty = \begin{cases} y(1 + 2^\gamma y^\gamma), & y \leq 1/2, \\ 2y - 1, & y > 1/2. \end{cases} \quad (1.3)$$

This is a prototypical example of a slowly mixing dynamical system [23]; the specific example is due to [20]. We describe in this subsection the homogenization results for the associated fast-slow systems which follow from this paper together with Part 1 [19], and compare these results with earlier works.

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  $v: \Lambda \rightarrow \mathbb{R}^m$  Hölder continuous with  $\int_{\Lambda} v \, d\mu = 0$ , the random variables  $n^{-1/2} \sum_{j=0}^{n-1} v \circ T^j$  defined on the probability space  $(\Lambda, \mu)$  converge in law to a (typically nondegenerate) normal distribution.

Consider a discrete-time fast-slow system of the form (1.1) with  $a_n \equiv a$ ,  $b_n \equiv b$ ,  $T_n \equiv T$  independent of  $n$ , and  $T$  such an intermittent map. Here  $a, b: \mathbb{R}^d \times \Lambda \rightarrow \mathbb{R}^d$  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  $x_n(t) = X_{[nt]}^{(n)}$ .

Prior results establish convergence  $x_n \rightarrow_{\mu} X$ , for  $X$  the solution of an SDE (1.2), provided that  $b$  is a product  $b(x, y) = h(x)v(y)$  with  $h: \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$  sufficiently smooth and  $v: \Lambda \rightarrow \mathbb{R}^m$  as above. It was proved first for  $\gamma < \frac{2}{11}$  in [16] using a discrete-time version of Hölder rough paths [15], and extended to the range  $\gamma < \frac{2}{5}$  in [6] using  $p$ -variation rough paths with jumps [13].

Part 1 by Korepanov *et al.* [19] develops the smooth ergodic theory side of things and together with [6] covers the optimal range  $\gamma \in (0, \frac{1}{2})$  in the case when  $b(x, y) = h(x)v(y)$  is a product. In the current paper, Theorem 2.10 enables two improvements to these prior results. First, the restriction that  $b$  is a product is now redundant. Second, we show that  $x_n \rightarrow_{\lambda} X$  for an enlarged class of measures  $\lambda$ . In particular, the hypotheses of Theorem 2.10 are verified in [19] for the most natural choice  $\lambda = \text{Leb}$  for all  $\gamma \in (0, \frac{1}{2})$ .

In addition, we consider the general setting (1.1) where  $T$ ,  $a$ ,  $b$  and  $\lambda$  are allowed to depend on  $n \in \mathbb{N} \cup \{\infty\}$ . This requires our main result Theorem 2.17. For example, consider the case where  $T_n$  is a family of intermittent maps with parameters  $\gamma_n$  limiting on  $\gamma_{\infty} \in (0, 1/2)$ . In [18], convergence results of the form  $x_n \rightarrow_{\mu_n} X$  and  $x_n \rightarrow_{\mu_{\infty}} X$  were obtained for the special case  $b_n(x, y) = h_n(x)v_n(y)$  with  $h_n$  exact. Theorem 2.17 combined with results in [19] yields the same convergence results without restrictions on  $b_n$ , and also shows that  $x_n \rightarrow_{\text{Leb}} X$ . The coefficients  $\tilde{a}$  and  $\sigma$  in (1.2) are given by

$$\tilde{a}(x) = \int_{\Lambda} a_{\infty}(x, y) \, d\mu_{\infty}(y) + \sum_{k=1}^d \sum_{\ell=1}^{\infty} \int_{\Lambda} b_{\infty}^k(x, y) \frac{\partial b_{\infty}}{\partial x_k}(x, T_{\infty}^{\ell} y) \, d\mu_{\infty}(y), \quad (1.4)$$

$$\begin{aligned} \sigma(x)^2 &= \int_{\Lambda} b_{\infty}(x, y) \otimes b_{\infty}(x, y) \, d\mu_{\infty}(y) \\ &\quad + \sum_{\ell=1}^{\infty} \int_{\Lambda} \{b_{\infty}(x, y) \otimes b_{\infty}(x, T_{\infty}^{\ell} y) + b_{\infty}(x, T_{\infty}^{\ell} y) \otimes b_{\infty}(x, y)\} \, d\mu_{\infty}(y), \end{aligned} \quad (1.5)$$

where  $b_{\infty}^k$  is the  $k$ th column of  $b_{\infty}$ . The details of how to apply Theorem 2.17 and the results in Part 1 [19] are given in Section 2.3.

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 (1.1) 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. We are grateful to the referees for helpful suggestions.

## 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 2.3 below is weaker than the naive restriction of Assumption 2.12). In Subsection 2.3, we show that our assumptions are satisfied for the intermittent maps considered in Section 1.1.

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

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

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

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

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

where  $D^k$  acts on the  $x$  component.

For the remainder of the section, we fix parameters  $q \in (1, \infty]$ ,  $\kappa, \bar{\kappa} \in (0, 1)$ , and  $\alpha > 2 + \frac{d}{q}$ . For  $T > 0$ , a metric space  $E$ , and a càdlàg function  $f : [0, 1] \rightarrow E$ , we define  $f^- : [0, 1] \rightarrow E$  by  $f^-(t) = \lim_{s \uparrow t} f(s)$  for  $t \in (0, T]$  and  $f^-(0) = f(0)$ .

### 2.1 $n$ -independent case

We now describe the assumptions and preliminary results required to state Theorem 2.10. We fix  $a \in C^{1+\bar{\kappa}, 0}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$  and  $b \in C^{\alpha, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$ , and consider for every integer  $n \geq 1$  the discrete-time dynamical system posed on  $\mathbb{R}^d \times \Lambda$

$$X_{k+1}^{(n)} = X_k^{(n)} + n^{-1}a(X_k^{(n)}, Y_k) + n^{-1/2}b(X_k^{(n)}, Y_k), \quad Y_{k+1} = TY_k, \quad (2.1)$$

where  $T : \Lambda \rightarrow \Lambda$  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  $\Lambda$ . Our first assumption deals with the function  $a$ .

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

$$\left| n^{-1} \sum_{k=0}^{n-1} a(x, Y_k) - \bar{a}(x) \right| \rightarrow_\lambda 0 \quad \text{as } n \rightarrow \infty.$$

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

$$W_{v,n}(t) = n^{-1/2} \sum_{0 \leq k < \lfloor nt \rfloor} v(Y_k), \quad \mathbb{W}_{v,w,n}(s, t) = \int_s^t (W_{v,n}^-(r) - W_{v,n}(s)) \otimes dW_{w,n}(r), \quad (2.2)$$

where we recall that  $W_{v,n}^-(r) = \lim_{s \uparrow r} W_{v,n}(s)$ . Note in particular that

$$\mathbb{W}_{v,w,n}(t) := \mathbb{W}_{v,w,n}(0, t) = n^{-1} \sum_{0 \leq k < \ell < \lfloor nt \rfloor} v(Y_k) \otimes w(Y_\ell). \quad (2.3)$$

Whenever  $v = w$ , we write simply  $\mathbb{W}_{v,n}$  for  $\mathbb{W}_{v,v,n}$ .

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

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

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

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

and

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

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

$$\mathbb{W}_v^{ij}(t) = \int_0^t W_v^i dW_v^j + \mathfrak{B}_0(v^i, v^j)t.$$

**Remark 2.4.** *One should compare Assumption 2.3(i) to [16, Thm. 9.1] and [17, 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.*

**Remark 2.5.** *Assumptions 2.2 and 2.3 are verified for a large class of dynamical systems in [16, Sec. 10] and [17, Sec. 1]. In these references, as in Subsection 1.1, there is a  $T$ -invariant ergodic Borel probability measure  $\mu$  on  $\Lambda$ , and we choose  $C_0^\kappa(\Lambda) = \{v \in C^\kappa(\Lambda) : \int_\Lambda v d\mu = 0\}$  and  $\bar{a} = \int_\Lambda a(\cdot, y) d\mu(y)$ .*

*The measure  $\mu$  plays no role in the proof of Theorem 2.10 and hence we do not mention it in our assumptions.*

**Remark 2.6.** *Under the assumption that  $\lambda$  is  $T$ -stationary, the simpler bounds*

$$|W_{v,n}(1)|_{L^{2q}(\lambda)} \leq K \quad \text{and} \quad |\mathbb{W}_{v,w,n}(1)|_{L^q(\lambda)} \leq K \quad \text{for all } n \geq 1$$

*imply Assumption 2.3(i).*

**Proposition 2.7.** *Suppose Assumption 2.3(i) holds. Then there exists  $K > 0$  such that for all  $n \geq 1$ ,  $0 \leq k, \ell \leq n$ , and  $v, w \in C_0^\kappa(\Lambda)$ ,*

$$\begin{aligned} \left| W_{v,n}(k/n) - W_{v,n}(\ell/n) \right|_{L^{2q}(\lambda)} &\leq K |v|_{C^\kappa} n^{-1/2} |k - \ell|^{1/2}, \\ \left| \mathbb{W}_{v,w,n}(k/n, \ell/n) \right|_{L^q(\lambda)} &\leq K |v|_{C^\kappa} |w|_{C^\kappa} n^{-1} |k - \ell|. \end{aligned} \quad (2.4)$$

*Proof.* As in [17, Prop. 2.7], the constants in Assumption 2.3(i) have the required dependence on  $|v|_{C^\kappa}$  and  $|w|_{C^\kappa}$  by the uniform boundedness principle.  $\square$

**Proposition 2.8.** *Suppose Assumption 2.3 holds. Then*

(a) *for all  $v \in C_0^\kappa(\Lambda, \mathbb{R}^m)$ , the limit  $\lim_{n \rightarrow \infty} n^{-1} \sum_{k=0}^{n-1} \mathbb{E}_\lambda(v^i v^j)(Y_k)$  exists and the covariance of  $W_v$  is given by*

$$\mathbb{E} W_v^i(1) W_v^j(1) = \mathfrak{B}(v^i, v^j) + \mathfrak{B}(v^j, v^i),$$

where

$$\mathfrak{B}(v^i, v^j) = \mathfrak{B}_0(v^i, v^j) + \frac{1}{2} \lim_{n \rightarrow \infty} n^{-1} \sum_{k=0}^{n-1} \mathbb{E}_\lambda(v^i v^j)(Y_k),$$

(b) *the bilinear operators  $\mathfrak{B}, \mathfrak{B}_0 : C_0^\kappa(\Lambda) \times C_0^\kappa(\Lambda) \rightarrow \mathbb{R}$  are bounded.*

*Proof.* (a) It follows from Assumption 2.3 that

$$\mathbb{E}_\lambda W_{v,n}^i(1) W_{v,n}^j(1) \rightarrow \mathbb{E} W_v^i(1) W_v^j(1),$$

and

$$\mathbb{E}_\lambda \mathbb{W}_{v,n}^{ij}(1) \rightarrow \mathbb{E} \mathbb{W}_v^{ij}(1) = \mathfrak{B}_0(v^i, v^j), \quad (2.5)$$

where we have used the fact that Itô integrals have zero mean. By (2.3), we have

$$W_{v,n}^i(1) W_{v,n}^j(1) = \mathbb{W}_{v,n}^{ij}(1) + \mathbb{W}_{v,n}^{ji}(1) + n^{-1} \sum_{k=0}^{n-1} (v^i v^j)(Y_k).$$

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

(b) Boundedness of  $\mathfrak{B}_0$  follows from (2.5) and (2.4) with  $k = 0$ ,  $\ell = n$ . By definition of  $\mathfrak{B}$ , we have  $|\mathfrak{B}(v, w)| \leq |\mathfrak{B}_0(v, w)| + \frac{1}{2} |v|_{C^0} |w|_{C^0}$ , yielding boundedness of  $\mathfrak{B}$ .  $\square$

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

$$\Sigma^{ij}(x) = \mathfrak{B}(b^i(x, \cdot), b^j(x, \cdot)) + \mathfrak{B}(b^j(x, \cdot), b^i(x, \cdot)), \quad i, j = 1, \dots, d,$$

*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 Proposition 2.8(a). Moreover,  $b$  lies in  $C^{\alpha, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$  with  $\alpha > 2 + \frac{d}{q} \geq 2$ , so  $\Sigma$  is  $C^2$  with globally bounded derivatives to second order. The conclusion now follows from [26, Thm. 5.2.3].  $\square$

As a consequence of Lemma 2.9 and [26, Cor. 5.1.2], for a Brownian motion  $B$  on  $\mathbb{R}^d$  and a Lipschitz function  $\tilde{a} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , there is a unique strong solution to the SDE

$$dX = \tilde{a}(X) dt + \sigma(X) dB, \quad X(0) = \xi. \quad (2.6)$$

In particular, the SDE (2.9) has uniqueness in law.

**Theorem 2.10.** *Suppose that Assumptions 2.2 and 2.3 hold and that  $\lim_{n \rightarrow \infty} \xi_n = \xi \in \mathbb{R}^d$ . Define the càdlàg path*

$$x_n : [0, 1] \rightarrow \mathbb{R}^d, \quad x_n(t) = X_{[nt]}^{(n)}. \quad (2.7)$$

*Then  $x_n \rightarrow_\lambda X$  in the uniform topology as  $n \rightarrow \infty$ , where  $X$  is a weak solution of the SDE (2.6), where  $B$  is a standard Brownian motion in  $\mathbb{R}^d$ ,  $\sigma$  is defined as in Lemma 2.9, and  $\bar{a}$  is the Lipschitz function given by*

$$\bar{a}^i(x) = \bar{a}^i(x) + \sum_{k=1}^d \mathfrak{B}_0(b^k(x, \cdot), \partial_k b^i(x, \cdot)), \quad i = 1, \dots, d.$$

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  $a_n \in C^{1+\bar{\kappa}, 0}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$ , and  $b_\infty, b_n \in C^{\alpha, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$  satisfying

$$\sup_{n \geq 1} |a_n|_{C^{1+\bar{\kappa}, 0}} + |b_n|_{C^{\alpha, \kappa}} < \infty, \quad \lim_{n \rightarrow \infty} |b_n - b_\infty|_{C^{\alpha, \kappa}} = 0.$$

For  $n \geq 1$ , we are interested in the discrete-time fast-slow system (1.1) where  $T_n : \Lambda \rightarrow \Lambda$  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  $\Lambda$ .

**Assumption 2.11.** *There exists  $\bar{a} \in C^{1+\bar{\kappa}}(\mathbb{R}^d, \mathbb{R}^d)$  such that, for all  $t \in [0, 1]$  and  $x \in \mathbb{R}^d$ ,*

$$|V_n(t)(x) - t\bar{a}(x)| \rightarrow_{\lambda_n} 0 \quad \text{as } n \rightarrow \infty,$$

where  $V_n(t) = n^{-1} \sum_{k=0}^{\lfloor tn \rfloor - 1} a_n(\cdot, Y_k^{(n)})$ .

As in (2.2), for  $v, w \in C^\kappa(\Lambda, \mathbb{R}^m)$  and  $0 \leq s \leq t \leq 1$ , define  $W_{v,n}(t) \in \mathbb{R}^m$ , and  $\mathbb{W}_{v,w,n}(s, t) \in \mathbb{R}^{m \times m}$  by

$$\begin{aligned} W_{v,n}(t) &= n^{-1/2} \sum_{0 \leq k < \lfloor nt \rfloor} v(Y_k^{(n)}), \\ \mathbb{W}_{v,w,n}(s, t) &= \int_s^t (W_{v,n}^-(r) - W_{v,n}(s)) \otimes dW_{w,n}(r), \end{aligned} \quad (2.8)$$

where we recall that  $W_{v,n}^-(r) = \lim_{s \uparrow r} W_{v,n}(s)$ . Whenever  $v = w$ , we again write  $\mathbb{W}_{v,n}$  for  $\mathbb{W}_{v,v,n}$ .

Recall our notational convention about subspaces  $C_0^\kappa(\Lambda)$  of  $C^\kappa(\Lambda)$  introduced before Assumption 2.3.

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

(i) *for all  $v = (v_1, \dots), w = (w_1, \dots) \in \prod_{n \in \mathbb{N}} C_n^\kappa(\Lambda)$  with*

$$\sup_n |v_n|_{C^\kappa} + |w_n|_{C^\kappa} < \infty,$$

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

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

*and*

$$|\mathbb{W}_{v_n, w_n, n}(k/n, \ell/n)|_{L^q(\lambda_n)} \leq Kn^{-1} |k - \ell|.$$



(ii) there exist bounded bilinear operators  $\mathfrak{B}_1, \mathfrak{B}_2 : C_\infty^\kappa(\Lambda) \times C_\infty^\kappa(\Lambda) \rightarrow \mathbb{R}$  such that for every  $m \geq 1$  and all  $\mathbf{v} = (v_n)_{n \in \mathbb{N} \cup \{\infty\}}$  with  $v_n \in C_n^\kappa(\Lambda, \mathbb{R}^m)$  and  $\lim_{n \rightarrow \infty} |v_n - v_\infty|_{C^\kappa} = 0$ ,

(a)  $\lim_{n \rightarrow \infty} n^{-1} \sum_{k=0}^{n-1} \mathbb{E}_{\lambda_n} (v_n^i v_n^j) (Y_k^{(n)}) = \mathfrak{B}_1(v_\infty^i, v_\infty^j)$ ,

(b)  $(W_{v_n, n}, \mathbb{W}_{v_n, n}) \rightarrow_{\lambda_n} (W_{\mathbf{v}}, \mathbb{W}_{\mathbf{v}})$  as  $n \rightarrow \infty$  in the sense of finite-dimensional distributions, where  $W_{\mathbf{v}}$  is an  $\mathbb{R}^m$ -valued Brownian motion and

$$\mathbb{W}_{\mathbf{v}}^{ij}(t) = \int_0^t W_{\mathbf{v}}^i dW_{\mathbf{v}}^j + \mathfrak{B}_2(v_\infty^i, v_\infty^j)t .$$

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

$$|W_{v_n, n}(k/n)|_{L^{2q}(\lambda_n)} \leq K(k/n)^{1/2} \quad \text{and} \quad |\mathbb{W}_{v_n, w_n, n}(0, k/n)|_{L^q(\lambda_n)} \leq Kk/n$$

for all  $0 \leq k \leq n$ , imply Assumption 2.12(i). Also, Assumption 2.12(ii)(a) reduces to  $\lim_{n \rightarrow \infty} \mathbb{E}_{\lambda_n} (v_n^i w_n^j) = \mathfrak{B}_1(v_\infty^i, w_\infty^j)$ .

**Proposition 2.14.** Suppose that Assumption 2.12(i) holds. Then there exists  $K > 0$  such that for all  $n \in \mathbb{N}$ ,  $0 \leq k, \ell \leq n$ , and  $v, w \in C_n^\kappa(\Lambda)$ ,

$$\begin{aligned} \left| W_{v, n}(k/n) - W_{v, n}(\ell/n) \right|_{L^{2q}(\lambda_n)} &\leq K|v|_{C^\kappa} n^{-1/2} |k - \ell|^{1/2} , \\ \left| \mathbb{W}_{v, w, n}(k/n, \ell/n) \right|_{L^q(\lambda_n)} &\leq K|v|_{C^\kappa} |w|_{C^\kappa} n^{-1} |k - \ell| . \end{aligned}$$

*Proof.* Identical to Proposition 2.7. □

**Proposition 2.15.** Suppose that Assumption 2.12 holds. Let  $\mathfrak{B} = \frac{1}{2}\mathfrak{B}_1 + \mathfrak{B}_2$ . Then, for all  $\mathbf{v} = (v_n)_{n \in \mathbb{N} \cup \{\infty\}}$  with  $\lim_{n \rightarrow \infty} |v_n - v_\infty|_{C^\kappa} = 0$ , the covariance of  $W_{\mathbf{v}}$  is given by

$$\mathbb{E} W_{\mathbf{v}}^i(1) W_{\mathbf{v}}^j(1) = \mathfrak{B}(v_\infty^i, v_\infty^j) + \mathfrak{B}(v_\infty^j, v_\infty^i) .$$

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

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

$$\Sigma^{ij}(x) = \mathfrak{B}(b_\infty^i(x, \cdot), b_\infty^j(x, \cdot)) + \mathfrak{B}(b_\infty^j(x, \cdot), b_\infty^i(x, \cdot)) , \quad i, j = 1, \dots, d ,$$

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 2.16 and [26, Cor. 5.1.2], for a Brownian motion  $B$  on  $\mathbb{R}^d$  and a Lipschitz function  $\tilde{a} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , there is a unique strong solution to the SDE

$$dX = \tilde{a}(X) dt + \sigma(X) dB , \quad X(0) = \xi . \tag{2.9}$$

In particular, the SDE (2.9) has uniqueness in law.

**Theorem 2.17.** Suppose that Assumptions 2.11 and 2.12 hold, and that  $\lim_{n \rightarrow \infty} \xi_n = \xi \in \mathbb{R}^d$ . Define the càdlàg path

$$x_n : [0, 1] \rightarrow \mathbb{R}^d , \quad x_n(t) = X_{[nt]}^{(n)} . \tag{2.10}$$

Then  $x_n \rightarrow_{\lambda_n} X$  in the uniform topology as  $n \rightarrow \infty$ , where  $X$  is a weak solution of the SDE (2.9), 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

$$\tilde{a}^i(x) = \tilde{a}^i(x) + \sum_{k=1}^d \mathfrak{B}_2(b_\infty^k(x, \cdot), \partial_k b_\infty^i(x, \cdot)) , \quad i = 1, \dots, d .$$

### 2.3 Homogenization for the illustrative example

In this subsection, we apply our main result, Theorem 2.17 in the case where the fast dynamics  $T_n$  is a family of intermittent maps as in Section 1.1. Using the results from Part 1 [19], we verify the hypotheses of Theorem 2.17 and deduce convergence to an SDE (1.2) with coefficients  $\tilde{a}$  and  $\sigma$  as given in (1.4) and (1.5).

Recall that  $\Lambda = [0, 1]$  and  $T_n : \Lambda \rightarrow \Lambda$ ,  $n \in \mathbb{N} \cup \{\infty\}$ , is a family of intermittent maps as in (1.3) with parameters  $\gamma_n \in (0, \frac{1}{2})$  such that  $\lim_{n \rightarrow \infty} \gamma_n = \gamma_\infty$ . Let  $\mu_n$  be the corresponding family of  $T_n$ -invariant ergodic absolutely continuous probability measures. Let  $C_n^\kappa(\Lambda) = \{v \in C^\kappa(\Lambda) : \int_\Lambda v d\mu_n = 0\}$  and fix  $q \in (1, \gamma_\infty^{-1} - 1)$ . We consider fast-slow systems (1.1) where  $a_n, b_n$  satisfy the regularity conditions at the beginning of Subsection 2.2 and  $b_n \in C_n^{\alpha, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$ . We require further that  $\lim_{n \rightarrow \infty} |a_n - a_\infty|_\infty = 0$  and that  $a_\infty(x, \cdot) : \Lambda \rightarrow \mathbb{R}^d$  is Hölder continuous for each fixed  $x$ .

To apply Theorem 2.17, we verify Assumptions 2.11 and 2.12 for appropriate families of probability measures  $\lambda_n$ . We do this for the case  $\lambda_n \equiv \text{Leb}$  using the results in [19, Sec. 4.1]. The case  $\lambda_n = \mu_n$  works in the same way (indeed, this is the easier case in [19]).

**Proposition 2.18.** *Assumption 2.11 holds with  $\bar{a}(x) = \int_\Lambda a_\infty(x, \cdot) d\mu_\infty$ .*

*Proof.* Fix  $x \in \mathbb{R}^d$  and define  $v_n = a_n(x, \cdot)$ . Then  $V_n(t)(x) = n^{-1} \sum_{j=0}^{\lfloor nt \rfloor - 1} v_n \circ T_n^j$  and it follows from [19, Prop. 4.3(a)] that  $V_n(t)(x) \xrightarrow{\text{Leb}} t \int_\Lambda v_\infty d\mu_\infty = t \int_\Lambda a_\infty(x, \cdot) d\mu_\infty$ .  $\square$

**Proposition 2.19.** *Assumption 2.12(i) holds.*

*Proof.* Let  $p = q + 1 \in (2, \gamma_\infty^{-1})$  and  $v = (v_1, \dots), w = (w_1, \dots) \in \prod_{n \in \mathbb{N}} C_n^\kappa(\Lambda)$ . By [19, Prop. 4.1], there is a constant  $C > 0$  such that for  $0 \leq \ell < k \leq n$ ,

$$\begin{aligned} |W_{v_n, n}(k/n) - W_{v_n, n}(\ell/n)|_{L^{2q}(\text{Leb})} &= n^{-1/2} \left| \sum_{\ell \leq j < k} v_n \circ T_n^j \right|_{L^{2(p-1)}(\text{Leb})} \\ &\leq C n^{-1/2} (k - \ell)^{1/2} |v_n|_{C^\kappa}, \\ |\mathbb{W}_{v_n, w_n, n}(k/n, \ell/n)|_{L^q(\text{Leb})} &= n^{-1} \left| \sum_{\ell \leq i < j < k} (v_n \circ T_n^i) \otimes (w_n \circ T_n^j) \right|_{L^{p-1}(\text{Leb})} \\ &\leq C n^{-1} (k - \ell) |v_n|_{C^\kappa} |w_n|_{C^\kappa}. \end{aligned}$$

These are the desired estimates.  $\square$

**Proposition 2.20.** *Assumption 2.12(ii) holds with*

$$\mathfrak{B}_1(v, w) = \int_\Lambda v w d\mu_\infty, \quad \mathfrak{B}_2(v, w) = \sum_{\ell=1}^{\infty} \int_\Lambda v \circ T_\infty^\ell w d\mu_\infty.$$

*Proof.* Assumption 2.12(ii)(a) is verified in [19, Prop. 4.3(b)]. By [19, Prop. 4.2] together with [19, Rem. 2.9],  $(W_{v_n, n}, \mathbb{W}_{v_n, n}) \xrightarrow{\text{Leb}} (W_\mathbf{v}, \mathbb{W}_\mathbf{v})$  where  $W_\mathbf{v}$  is an  $\mathbb{R}^m$ -valued Brownian motion and

$$\mathbb{W}_\mathbf{v}^{ij}(t) = \int_0^t W_\mathbf{v}^i dW_\mathbf{v}^j + E_\infty^{ij} t, \quad E_\infty^{ij} = \sum_{\ell=1}^{\infty} \int_\Lambda v_\infty^i v_\infty^j \circ T_\infty^\ell d\mu_\infty.$$

This verifies Assumption 2.12(ii)(b).  $\square$

We can now apply Theorem 2.17. Define  $\tilde{a}$  as in (1.4) and set

$$\begin{aligned} \Sigma(x) &= \int_{\Lambda} b_{\infty}(x, y) \otimes b_{\infty}(x, y) \, d\mu_{\infty}(y) \\ &\quad + \sum_{\ell=1}^{\infty} \int_{\Lambda} \{b_{\infty}(x, y) \otimes b_{\infty}(x, T_{\infty}^{\ell} y) + b_{\infty}(x, T_{\infty}^{\ell} y) \otimes b_{\infty}(x, y)\} \, d\mu_{\infty}(y). \end{aligned}$$

Let  $\sigma$  be the unique positive semidefinite square root of  $\Sigma$  as in Lemma 2.16. By Theorem 2.17,  $x_n \rightarrow_{\text{Leb}} X$  where  $X$  is the unique solution to the SDE (1.2) with coefficients  $\tilde{a}$  and  $\sigma$ .

### 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  $\mathcal{A}, \mathcal{B}$ , we denote their algebraic tensor product by

$$\mathcal{A} \otimes_a \mathcal{B} := \text{span} \{a \otimes b \mid a \in \mathcal{A}, b \in \mathcal{B}\}.$$

Given  $f \in \mathcal{A}^*$  (the dual space of  $\mathcal{A}$ ),  $g \in \mathcal{B}^*$ , one may define an element on  $(\mathcal{A} \otimes_a \mathcal{B})^*$  by

$$(f \otimes g) \left( \sum_{i=1}^N a_i \otimes b_i \right) := \sum_{i=1}^N f(a_i) g(b_i).$$

As a result, we consider  $\mathcal{A}^* \otimes_a \mathcal{B}^*$  as a subspace of  $(\mathcal{A} \otimes_a \mathcal{B})^*$ . Generally, there are different (inequivalent) norms on  $\mathcal{A} \otimes_a \mathcal{B}$ . We call a norm  $|\cdot|_{\mathcal{A} \otimes_a \mathcal{B}}$  on the vector space  $\mathcal{A} \otimes_a \mathcal{B}$  admissible (or reasonable), if for any  $a \in \mathcal{A}, b \in \mathcal{B}, f \in \mathcal{A}^*, g \in \mathcal{B}^*$ ,

$$|a \otimes b|_{\mathcal{A} \otimes_a \mathcal{B}} \leq |a|_{\mathcal{A}} |b|_{\mathcal{B}}, \quad |f \otimes g|_{(\mathcal{A} \otimes_a \mathcal{B})^*} \leq |f|_{\mathcal{A}^*} |g|_{\mathcal{B}^*}, \quad (3.1)$$

where  $|\cdot|_{(\mathcal{A} \otimes_a \mathcal{B})^*}$  is defined as the dual norm on  $(\mathcal{A} \otimes_a \mathcal{B}, |\cdot|_{\mathcal{A} \otimes_a \mathcal{B}})^*$ . Examples of admissible norms are the projective tensor norm and the injective tensor norm, see [24, Sec. 6.1]. One may then complete  $\mathcal{A} \otimes_a \mathcal{B}$  under  $|\cdot|_{\mathcal{A} \otimes_a \mathcal{B}}$  to obtain a Banach space. All the tensor product spaces  $\mathcal{A} \otimes \mathcal{B}$  we consider in the sequel will implicitly be assumed to be Banach spaces completed by 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  $\mathcal{P} = \{[t_0, t_1], [t_1, t_2], \dots, [t_{k-1}, t_k]\}$  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  $\mathcal{B}$  and  $p > 0$ , let  $\mathcal{V}^{p\text{-var}}([s, t], \mathcal{B})$  denote the space of all functions  $\Xi : \{(u, v) \in [s, t]^2 \mid u \leq v\} \rightarrow \mathcal{B}$  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  $\mathbf{X} : [s, t] \rightarrow \mathcal{B} \oplus \mathcal{B}^{\otimes 2}$  and  $s \leq u \leq v \leq t$ , define the increment  $\mathbf{X}(u, v) := (X(u, v), \mathbb{X}(u, v)) := \mathbf{X}(u)^{-1} \mathbf{X}(v)$ . For  $p \geq 1$ , define the (homogeneous)  $p$ -variation of  $\mathbf{X}$  by

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

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

$$\|\mathbf{X}; \tilde{\mathbf{X}}\|_{p\text{-var}} := \|X - \tilde{X}\|_{p\text{-var}} + \|\mathbb{X} - \tilde{\mathbb{X}}\|_{p/2\text{-var}}, \quad (3.2)$$

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

$$\sigma_{p\text{-var}}(\mathbf{X}, \tilde{\mathbf{X}}) := \inf_{\omega \in \Omega} \{|\omega| + \|\mathbf{X}; \tilde{\mathbf{X}} \circ \omega\|_{p\text{-var}}\}, \quad (3.3)$$

where  $\Omega$  denotes the set of all continuous increasing bijections  $\omega : [0, T] \rightarrow [0, T]$ , and

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

Let  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$  denote the space of all  $p$ -rough paths equipped with the metric  $\sigma_{p\text{-var}}$ . 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 space  $D^{p\text{-var}}(\mathcal{B})$  in the exact same way as above but without the component  $\mathbb{X}$ .

The purpose of the metric  $\sigma_{p\text{-var}}$  is to provide convenient tightness results. In short, tightness in the metric space  $(D^{p\text{-var}}, \sigma_{p\text{-var}})$  is implied by tightness of  $p'$ -variation for  $p' < p$  and tightness in the  $(J_1)$  Skorokhod space, with the latter two being simpler to check; see the proofs of Lemmas 4.5 and 4.7. Likewise for  $\mathcal{D}^{p\text{-var}}$ . The same is not true if we replace  $\sigma_{p\text{-var}}$  by  $\|\cdot\|_{p\text{-var}}$ .

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

$$\|\mathbf{X}; \tilde{\mathbf{X}}\|_{\infty} = \|X - \tilde{X}\|_{\infty} + \|\mathbb{X} - \tilde{\mathbb{X}}\|_{\infty},$$

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

$$\|\mathbf{X}; \tilde{\mathbf{X}}\|_{p'\text{-var}} \leq \|\mathbf{X}; \tilde{\mathbf{X}}\|_{\infty}^{1-p/p'} \|\mathbf{X}; \tilde{\mathbf{X}}\|_{p\text{-var}}^{p/p'}. \quad (3.4)$$

*Proof.* We readily see that

$$\|\mathbf{X}; \tilde{\mathbf{X}}\|_{p'\text{-var}} \leq \|X - \tilde{X}\|_{\infty}^{1-p/p'} \|X - \tilde{X}\|_{p\text{-var}}^{p/p'} + \|\mathbb{X} - \tilde{\mathbb{X}}\|_{\infty}^{1-p/p'} \|\mathbb{X} - \tilde{\mathbb{X}}\|_{p/2\text{-var}}^{p/p'},$$

and the conclusion follows by Hölder's inequality  $a^{\theta} \bar{a}^{1-\theta} + b^{\theta} \bar{b}^{1-\theta} \leq (a+b)^{\theta} (\bar{a} + \bar{b})^{1-\theta}$  for  $\theta \in [0, 1]$  and  $a, \bar{a}, b, \bar{b} \geq 0$ .  $\square$

We now introduce rough integration in the level-2 rough path case. Given Banach spaces  $\mathcal{B}, \mathcal{E}$ , let  $\mathbf{L}(\mathcal{B}, \mathcal{E})$  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

$$Y \in D^{p\text{-var}}(\mathcal{E}), \quad Y' \in D^{p\text{-var}}(\mathbf{L}(\mathcal{B}, \mathcal{E})),$$

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

$$R(s, t) := Y(s, t) - Y'(s)X(s, t). \quad (3.5)$$

We denote the space of  $X$ -controlled rough paths as  $\mathcal{D}_X^{p/2\text{-var}}(\mathcal{E})$ . 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. Remark that,  $\mathcal{B}^* \otimes_a \mathcal{B}^*$  is a subspace of  $(\mathcal{B} \otimes \mathcal{B})^*$  by admissibility of norms (3.1) and therefore  $f \otimes g \in (\mathcal{B} \otimes \mathcal{B})^*$  for every  $f, g \in \mathcal{B}^*$ . In particular,  $\Xi$  in the statement of the following lemma is well-defined.

**Lemma 3.4.** *Let  $\mathbf{X} \in \mathcal{D}^{p\text{-var}}(\mathcal{B})$ ,  $(Y, Y') \in \mathcal{D}_X^{p/2\text{-var}}(\mathbb{R}^d)$ , and  $H \in C^2(\mathbb{R}^d, \mathbf{L}(\mathcal{B}, \mathbb{R}^d))$ . Then, for every  $t \in [0, T]$ , the following integral (with values in  $\mathbb{R}^d$ ) is well-defined*

$$\mathcal{I}_{\mathbf{X}}(Y)(t) := \int_0^t H(Y^-(s)) d\mathbf{X}(s) := \lim_{|\mathcal{P}| \rightarrow 0} \sum_{[u,v] \in \mathcal{P}} \Xi(u, v), \quad (3.6)$$

where  $\mathcal{P}$  are partitions of  $[0, t]$  and, for  $i = 1, \dots, d$  and  $0 \leq u \leq v \leq T$ ,

$$\Xi(u, v)^i = H^i(Y(u))X(u, v) + \sum_{k=1}^d \left( \partial_k H^i(Y(u)) \otimes (Y'(u))^k \right) \mathbb{X}(u, v).$$

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

*Proof.* 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  $R^{H(Y)} \in \mathcal{V}_2^{\frac{p}{2}}(\mathbf{L}(\mathcal{B}, \mathbb{R}^d))$ . Then one has the identity

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

According to the generalized sewing lemma [13, Thm. 2.5], the integral  $\mathcal{I}_{\mathbf{X}}(Y)$  is well-defined, and furthermore one has the local estimate

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

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

**Remark 3.5.** *Generally, to integrate  $(Y, Y')$  against  $\mathbf{X}$ , one needs  $Y(t) \in \mathbf{L}(\mathcal{B}, \mathcal{E})$  and  $Y'(t) \in \mathbf{L}(\mathcal{B}, \mathbf{L}(\mathcal{B}, \mathcal{E}))$  to have the identity  $Y(s, t) = Y'(s)X(s, t) + R(s, t)$ . In this case, one further needs the embedding  $\mathbf{L}(\mathcal{B}, \mathbf{L}(\mathcal{B}, \mathcal{E})) \hookrightarrow \mathbf{L}(\mathcal{B} \otimes \mathcal{B}, \mathcal{E})$  to define  $\Xi(s, t) := Y(s)X(s, t) + Y'(s)\mathbb{X}(s, t)$ . Luckily, in the above case where  $\mathcal{E} = \mathbb{R}^d$ , the embedding assumption is replaced by the fact  $DH(Y)Y' \in \mathbf{L}(\mathcal{B} \otimes \mathcal{B}, \mathbb{R}^d)$  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., [13, Thm. 3.8, 3.9], thanks to admissibility of norms.

**Theorem 3.6.** *Let  $\mathcal{A}, \mathcal{B}$  be Banach spaces,  $p \in [2, 3)$ ,  $q \in [1, p/2]$ , and  $F \in C^\beta(\mathbb{R}^d, \mathbf{L}(\mathcal{A}, \mathbb{R}^d))$ ,  $H \in C^\gamma(\mathbb{R}^d, \mathbf{L}(\mathcal{B}, \mathbb{R}^d))$  for  $\beta > q, \gamma > p$ . Then, for any  $V \in D^{q\text{-var}}(\mathcal{A})$ ,  $\mathbf{X} \in \mathcal{D}^{p\text{-var}}(\mathcal{B})$ , and  $Y_0 \in \mathbb{R}^d$ , there exists a unique  $X$ -controlled rough path  $(Y, Y') \in \mathcal{D}_X^{p/2\text{-var}}(\mathbb{R}^d)$  solving the equation*

$$Y(t) = Y_0 + \int_0^t F(Y^-(s)) dV(s) + \int_0^t H(Y^-(s)) d\mathbf{X}(s). \quad (3.7)$$

Moreover, the solution map is locally Lipschitz in the sense that

$$\|Y - \tilde{Y}\|_{p\text{-var}} \lesssim \|\mathbf{X}; \tilde{\mathbf{X}}\|_{p\text{-var}} + \|V - \tilde{V}\|_{q\text{-var}} + |Y_0 - \tilde{Y}_0|, \quad (3.8)$$

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

Recall that in (3.7),  $\int_0^t H(Y^-(s)) d\mathbf{X}(s)$  is defined by (3.6) and that  $\int_0^t F(Y^-(s)) dV(s)$  is the classical Young integral [13, Prop. 2.4]

$$\int_0^t F(Y^-(s)) dV(s) = \lim_{|\mathcal{P}| \rightarrow 0} \sum_{[u,v] \in \mathcal{P}} F(Y(u))V(u,v),$$

where  $\mathcal{P}$  are partitions of  $[0, t]$ , which is well-defined since  $1/q + 1/p > 1$ . Note that the restriction  $q \leq p/2$  arises from the Young estimate  $\|\int_0^t (Z^-(s) - Z(0)) dV(s)\|_{q\text{-var}} \lesssim \|Z\|_{p\text{-var}} \|V\|_{q\text{-var}}$  and the requirement that  $R \in \mathcal{V}^{p/2\text{-var}}(\mathbb{R}^d)$  in the definition (3.5).

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 3.7.** *Let notation be as in Theorem 3.6. Consider the solution map to equation (3.7)*

$$\begin{aligned} \Phi &: D^{q\text{-var}}(\mathcal{A}) \times \mathcal{D}^{p\text{-var}}(\mathcal{B}) \times \mathbb{R}^d \rightarrow D^{p\text{-var}}(\mathbb{R}^d), \\ \Phi &: (V, \mathbf{X}, Y_0) \mapsto Y. \end{aligned}$$

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

*Proof.* It suffices to consider  $p' \in (p, \gamma)$ . Let  $\mathbf{X} \in \mathcal{D}^{p\text{-var}}(\mathcal{B})$  be continuous. We claim that  $\sigma_{p\text{-var}}(\mathbf{X}_n, \mathbf{X}) \rightarrow 0$  implies  $\|\mathbf{X}_n; \mathbf{X}\|_{p'\text{-var}} \rightarrow 0$ . Indeed,  $\sigma_{p\text{-var}}(\mathbf{X}_n, \mathbf{X}) \rightarrow 0$  implies the existence of  $\{\omega_n\}_{n \geq 1} \subset \Omega$  satisfying both  $|\omega_n| \rightarrow 0$  and  $\|\mathbf{X}_n; \mathbf{X} \circ \omega_n\|_{p\text{-var}} \rightarrow 0$ . Observe that  $\|\mathbf{X}; \mathbf{X} \circ \omega_n\|_\infty \rightarrow 0$  by continuity of  $\mathbf{X}$ , and therefore, combining with the interpolation estimate (3.4),  $\|\mathbf{X}; \mathbf{X} \circ \omega_n\|_{p'\text{-var}} \rightarrow 0$ . Since  $\|\mathbf{X}_n; \mathbf{X}\|_{p'\text{-var}} \leq \|\mathbf{X}_n; \mathbf{X} \circ \omega_n\|_{p'\text{-var}} + \|\mathbf{X} \circ \omega_n; \mathbf{X}\|_{p'\text{-var}}$ , this proves the claim. The same considerations apply to continuous  $V \in D^{q\text{-var}}(\mathcal{A})$  and  $q' > q$ . The result follows from Theorem 3.6 by taking  $q' \in (q, \beta \wedge p')$ .  $\square$

**Remark 3.8.** *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  $D^2 \rightarrow D$ ,  $(x, y) \mapsto x + y$ , 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  $(V, \mathbf{X}, Y_0)$  is a continuity point of  $\Phi$  whenever one of  $\mathbf{X}$  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 3.9.** *Let  $(\mathcal{O}, \mathcal{F}, \mathbb{P})$  be a probability space and let  $\{\mathbf{X}_t\}_{t \in [0, T]} = \{(X_t, \mathbb{X}_t)\}_{t \in [0, T]}$  be a  $\mathcal{B} \oplus \mathcal{B}^{\otimes 2}$ -valued stochastic process defined on  $(\mathcal{O}, \mathcal{F}, \mathbb{P})$ . Suppose further that, for  $\mathbb{P}$ -a.e.  $\omega \in \mathcal{O}$ ,  $t \mapsto \mathbf{X}_t^\omega$  is càdlàg, piecewise constant, and has jump times contained in a deterministic set  $\{t_j\}_{0 \leq j \leq n} \subset [0, T]$  with  $0 = t_0 < t_1 < \dots < t_n = T$ , such that, for some  $C_1, C_2 > 0$ ,  $\beta \in (0, \frac{1}{2}]$ , and  $q \in [1, \infty]$ ,*

$$|X(t_i, t_j)|_{L^{2q}(\mathbb{P})} \leq C_1 |t_j - t_i|^\beta, \quad |\mathbb{X}(t_i, t_j)|_{L^q(\mathbb{P})} \leq C_2 |t_j - t_i|^{2\beta}.$$

If  $2q > \frac{1}{\beta}$ , then for any  $\alpha \in (\frac{1}{2q}, \beta)$

$$\mathbb{E}[\|\mathbf{X}\|_{1/\alpha\text{-var}}^{2q}]^{\frac{1}{2q}} \leq CT^{\alpha - \frac{1}{2q}}(C_1 + C_2^{1/2}) \quad (3.9)$$

and

$$\mathbb{E}\left[\left|\sup_{t_i \neq t_j} \frac{|X(t_i, t_j)| + |\mathbb{X}(t_i, t_j)|^{1/2}}{|t_i - t_j|^{\alpha - \frac{1}{2q}}}\right|^{2q}\right]^{\frac{1}{2q}} \leq C(C_1 + C_2^{1/2}) \quad (3.10)$$

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

For the proof, we require the following lemma.

**Lemma 3.10.** *Let  $\mathbf{X}$  be as in Proposition 3.9. Then there exists a  $\mathbb{P}$ -a.s. continuous  $\mathcal{B} \oplus \mathcal{B}^{\otimes 2}$ -valued process  $\{\tilde{\mathbf{X}}_t\}_{t \in [0, T]} = \{(\tilde{X}_t, \tilde{\mathbb{X}}_t)\}_{t \in [0, T]}$  such that  $\mathbf{X}(t_i) = \tilde{\mathbf{X}}(t_i)$  for all  $i = 0, \dots, n$ , and*

$$|\tilde{X}(s, t)|_{L^{2q}(\mathbb{P})} \leq 3^{1-\beta} C_1 |t - s|^\beta, \quad |\tilde{\mathbb{X}}(s, t)|_{L^q(\mathbb{P})} \leq 3^{2-2\beta} (C_2 + C_1^2) |t - s|^{2\beta}. \quad (3.11)$$

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

$$\begin{aligned} \tilde{X}(t) &:= X(t_j) + \frac{t - t_j}{t_{j+1} - t_j} X(t_j, t_{j+1}), \\ \tilde{\mathbb{X}}(0, t) &:= \mathbb{X}(0, t_j) + \frac{t - t_j}{t_{j+1} - t_j} (\mathbb{X}(0, t_{j+1}) - \mathbb{X}(0, t_j)). \end{aligned}$$

To prove (3.11), 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

$$\begin{aligned} |\tilde{X}(s, t)|_{L^{2q}(\mathbb{P})} &\leq |\tilde{X}(s, t_{j+1})|_{L^{2q}(\mathbb{P})} + |\tilde{X}(t_{j+1}, t_k)|_{L^{2q}(\mathbb{P})} + |\tilde{X}(t_k, t)|_{L^{2q}(\mathbb{P})} \\ &\leq C_1 (|t_{j+1} - s|^\beta + |t_{j+1} - t_k|^\beta + |t - t_k|^\beta) \\ &\leq 3^{1-\beta} C_1 |t - s|^\beta. \end{aligned}$$

Furthermore, one can check that

$$\tilde{\mathbb{X}}(s, t_{j+1}) = \frac{t_{j+1} - s}{t_{j+1} - t_j} \mathbb{X}(t_j, t_{j+1}) + \frac{(t_{j+1} - s)(s - t_j)}{(t_{j+1} - t_j)^2} X_{t_j, t_{j+1}}^{\otimes 2},$$

from which it follows that

$$|\tilde{\mathbb{X}}(s, t_{j+1})|_{L^q(\mathbb{P})} \leq (C_2 + C_1^2) |t_{j+1} - s|^{2\beta}.$$

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

$$\begin{aligned} |\tilde{\mathbb{X}}(s, t)|_{L^q(\mathbb{P})} &\leq |\tilde{\mathbb{X}}(s, t_{j+1})|_{L^q(\mathbb{P})} + |\tilde{\mathbb{X}}(t_{j+1}, t_k)|_{L^q(\mathbb{P})} + |\tilde{\mathbb{X}}(t_k, t)|_{L^q(\mathbb{P})} \\ &\quad + |\tilde{X}(s, t_{j+1}) \otimes \tilde{X}(t_{j+1}, t_k)|_{L^q(\mathbb{P})} + |\tilde{X}(s, t_k) \otimes X(t_k, t)|_{L^q(\mathbb{P})} \\ &\leq 3^{1-2\beta} (C_2 + C_1^2) |t - s|^{2\beta} + C_1^2 |t - s|^{2\beta} + 2^{1-\beta} C_1^2 |t - s|^{2\beta} \\ &\leq 3^{2-2\beta} (C_2 + C_1^2) |t - s|^{2\beta}. \end{aligned}$$

□

*Proof of Proposition 3.9.* Let  $\tilde{\mathbf{X}}$  be as in Lemma 3.10 and suppose  $2q > \frac{1}{\beta}$  and  $\alpha \in (\frac{1}{2q}, \beta)$ . Using the notation in Appendix A, we have by Corollary A.3

$$\begin{aligned} \mathbb{E}[\|\tilde{\mathbf{X}}\|_{1/\alpha\text{-var}}^{2q}]^{\frac{1}{2q}} &\leq C(\alpha, q) T^{\alpha - \frac{1}{2q}} \mathbb{E}[\|\tilde{\mathbf{X}}\|_{W^{\alpha, 2q}}^{2q}]^{\frac{1}{2q}} \\ &= C(\alpha, q) T^{\alpha - \frac{1}{2q}} \mathbb{E}\left[\iint_{[0, T]^2} \frac{|\tilde{X}(s, t)|_{\mathcal{B}}^{2q} + |\tilde{X}(s, t)|_{\mathcal{B}^{\otimes 2}}^q}{|t - s|^{2\alpha q + 1}} ds dt\right]^{\frac{1}{2q}}. \end{aligned}$$

Using the estimate (3.11) 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,  $\mathbb{E}[\|\tilde{\mathbf{X}}\|_{(\alpha - \frac{1}{2q})\text{-Hölder}}^{2q}]^{\frac{1}{2q}} \leq C(C_1 + C_2^{1/2})$ . The conclusion follows since  $\tilde{\mathbf{X}}(t_i) = \mathbf{X}(t_i)$  and  $\mathbf{X}$  is constant on  $[t_i, t_{i+1})$ .  $\square$

## 4 Proof of the main result

This section is devoted to the proof of Theorem 2.17. Throughout this section, we let notation be as in Section 2.2.

The first step is to reformulate the system (1.1) as a càdlàg controlled ODE. Let us fix  $\kappa' \in (0, \bar{\kappa})$  and  $\theta \in (2, \alpha - \frac{d}{q})$ , and introduce the Banach spaces

$$\mathcal{A} = C^{1+\kappa'}(\mathbb{R}^d, \mathbb{R}^d) \quad \text{and} \quad \mathcal{B} = C^\theta(\mathbb{R}^d, \mathbb{R}^d).$$

We furthermore equip  $\mathcal{B}^{\otimes 2}$  with the admissible norm as specified in [17, Prop. 4.5].

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

The following lemma is now immediate from Theorem 3.6.

**Lemma 4.1.** *The càdlàg RDE*

$$dx(t) = F(x^-(t)) dV(t) + H(x^-(t)) d\mathbf{W}(t), \quad x(0) = \xi \in \mathbb{R}^d \quad (4.1)$$

is well-posed for any  $(V, \mathbf{W}) \in D^{\beta\text{-var}}([0, 1], \mathcal{A}) \times \mathcal{D}^{p\text{-var}}(\mathcal{B})$  with  $\beta \in [1, 1 + \kappa')$  and  $p \in [2, \theta)$  such that  $\beta \leq p/2$ .

We introduce the  $\mathcal{A}$ -valued and  $\mathcal{B}$ -valued paths

$$V_n(t) = n^{-1} \sum_{k=0}^{\lfloor tn \rfloor - 1} a_n(\cdot, Y_k^{(n)}), \quad W_n(t) = n^{-1/2} \sum_{k=0}^{\lfloor tn \rfloor - 1} b_n(\cdot, Y_k^{(n)}),$$

and let  $\mathbf{W}_n = (W_n, \mathbb{W}_n)$  be the canonical level-2 lift of  $W_n$  given by

$$\mathbb{W}_n(t) = \int_0^t W_n(r) \otimes dW_n(r).$$

Remark that  $\mathbf{W}_n$  is a  $p$ -rough path over  $\mathcal{B}$  for any  $p \in [2, 3)$  in the sense of Definition 3.2.

**Lemma 4.2.** *The path  $x_n$  given by (2.10) is the unique solution of the càdlàg ODE*

$$dx_n = F(x_n^-) dV_n + H(x_n^-) d\mathbf{W}_n, \quad x_n(0) = \xi_n \in \mathbb{R}^d. \quad (4.2)$$



*Proof.* Observe that  $x_n$  given by (2.10) satisfies for all  $1 \leq k \leq n$

$$\begin{aligned} x_n(k/n) - x_n((k-1)/n) &= n^{-1}a(x_n((k-1)/n), Y_{k-1}^{(n)}) \\ &\quad + n^{-1/2}b(x_n((k-1)/n), Y_{k-1}^{(n)}) \\ &= \int_{(k-1)/n}^{k/n} F(x_n^-(s)) dV_n(s) + H(x_n^-(s)) dW_n(s). \end{aligned}$$

□

Following Lemmas 4.1 and 4.2, we are reduced to showing convergence in law for  $V_n$  and  $\mathbf{W}_n$  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  $a_n$  and  $b_n$  is uniformly bounded, i.e., there exists a compact set  $\mathfrak{K} \subset \mathbb{R}^d$  such that the support of  $a_n$  and  $b_n$  is contained in  $\mathfrak{K} \times \Lambda$  for all  $n \in \mathbb{N} \cup \{\infty\}$ .

**Theorem 4.3.** *Suppose that Assumptions 2.11 and 2.12 hold and that the support of  $a_n$  and  $b_n$  is uniformly bounded. Then, for any  $p \in (2, 3)$  and  $\beta \in (1, 2)$ , there exists a random variable  $(V, \mathbf{W})$  in  $D^{\beta\text{-var}}(\mathcal{A}) \times \mathcal{D}^{p\text{-var}}(\mathcal{B})$  such that  $(V_n, \mathbf{W}_n) \rightarrow_{\lambda_n} (V, \mathbf{W})$ , and such that  $(V, \mathbf{W})$  is a.s. continuous. Moreover, if  $\beta \in (1, 1 + \kappa')$ ,  $p \in (2, \theta)$ , and  $\beta \leq p/2$ , then the RDE (4.1) driven by  $(V, \mathbf{W})$  along the vector fields  $(F, H)$  is a weak solution of the SDE (2.9).*

Before proving Theorem 4.3, we first state an immediate consequence of Corollary 3.7, 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  $a_n$  and  $b_n$  is uniformly bounded. Then, for any  $p > 2$ ,  $x_n \rightarrow_{\lambda_n} X$  in the  $p$ -variation norm  $|x(0)| + \|x\|_{p\text{-var}}$ , where  $X$  is a weak solution of the SDE (2.9).*

We break the proof of Theorem 4.3 into several lemmas.

**Lemma 4.5.** *Suppose that Assumption 2.11 holds and that the support of  $a_n$  is uniformly bounded. Then for every  $\beta > 1$*

$$\|V_n - V\|_{\beta\text{-var}} \rightarrow_{\lambda_n} 0,$$

where  $V: [0, 1] \rightarrow \mathcal{A}$  is the deterministic path  $V(t) = t\bar{a}$ .

*Proof.* Let  $\mathfrak{K} \subset \mathbb{R}^d$  be compact such that  $\mathfrak{K} \times \Lambda$  contains the support of  $a_n$ . Then the embedding  $C^{1+\bar{\kappa}}(\mathfrak{K}, \mathbb{R}^d) \hookrightarrow \mathcal{A}$  is compact. Observe further that, for all  $s, t \in [0, 1]$  with  $|t - s| > n^{-1}$ ,

$$|V_n(t) - V_n(s)|_{C^{1+\bar{\kappa}}} \leq 2|t - s| \|a_n\|_{C^{1+\bar{\kappa}, 0}}. \quad (4.3)$$

It follows that  $(V_n)_{n \geq 1}$  satisfies the compact containment condition [7, Rem. 3.7.3] and condition [7, Thm. 3.7.2(b)]. Hence, by the tightness criterion [7, Thm. 3.7.6],  $(V_n)_{n \geq 1}$  is tight in the  $(J_1)$  Skorokhod space  $D([0, 1], \mathcal{A})$  (note that [7, Thm. 3.7.6] implies only relative compactness, but tightness is a consequence of the proof).

Observe next that, for  $\beta > 1$ , the interpolation estimate (3.4) implies  $\sigma_{\beta\text{-var}}(X, Y) \leq \sigma_{\infty}(X, Y)^{1-1/\beta} (\|X\|_{1\text{-var}} + \|Y\|_{1\text{-var}})^{1/\beta}$ , where  $\sigma_{\infty}$  is the  $(J_1)$  Skorokhod metric defined as in (3.3) with  $\|\cdot\|_{p\text{-var}}$  replaced by  $\|\cdot\|_{\infty}$ . Moreover, the map  $X \mapsto \|X\|_{1\text{-var}}$  is invariant under reparametrizations and is lower semi-continuous under the metric  $\|\cdot\|_{\infty}$  and thus under the metric  $\sigma_{\infty}$ . It follows that, for every  $R \geq 0$  and every compact subset  $K$  of the classical  $(J_1)$  Skorokhod space  $D([0, 1], \mathcal{A})$ , the set

$$\{X \in K \mid \|X\|_{1\text{-var}} \leq R\}$$

is compact in  $D^{\beta\text{-var}}(\mathcal{A})$ . Furthermore, (4.3) implies that, as a càdlàg path with values in  $C^{1+\bar{\kappa}}(\mathbb{R}^d, \mathbb{R}^d)$ ,  $V_n$  has 1-variation uniformly bounded in  $n \geq 1$  and  $Y_0^{(n)} \in \Lambda$ . It follows that  $(V_n)_{n \geq 1}$  is tight in  $D^{\beta\text{-var}}(\mathcal{A})$ . Hence, by Prokhorov's theorem [2, Thm. 5.1],  $(V_n)_{n \geq 1}$  is weakly relatively compact in the space of probability measures on  $D^{\beta\text{-var}}(\mathcal{A})$ , and by Assumption 2.11, the only possible limit point is  $V$ . Finally, since  $V$  is continuous and deterministic, the same argument as in the proof of Corollary 3.7 implies that  $\|V_n - V\|_{\beta\text{-var}} \rightarrow_{\lambda_n} 0$  for all  $\beta > 1$ .  $\square$

Showing convergence of  $W_n$  is more involved.

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

$$\begin{aligned} \mathbb{E}_{\lambda_n} \left[ |W_n(k/n) - W_n(\ell/n)|_{\mathcal{B}}^{2q} \right]^{1/(2q)} &\lesssim n^{-1/2} |k - \ell|^{1/2}, \\ \mathbb{E}_{\lambda_n} \left[ |\mathbb{W}_n(k/n, \ell/n)|_{\mathcal{B}^{\otimes 2}}^q \right]^{1/q} &\lesssim n^{-1} |k - \ell|, \end{aligned}$$

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

*Proof.* For a function  $u : \mathbb{R}^d \rightarrow \mathbb{R}$ , 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  $B_p^s$  consisting of all  $L^p$  functions  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  such that

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

Let us further introduce the notation

$$\Delta_\sigma^m W_n(k/n, \ell/n; x) = n^{-1/2} \sum_{r=k}^{\ell} \Delta_\sigma^m b_n(x, Y_r^{(n)}).$$

Denote in the sequel  $s = k/n$  and  $t = \ell/n$ . Proposition 2.14 implies that for each  $m \geq 1$  (cf. [17, p. 4088])

$$\mathbb{E}_{\lambda_n} \left[ |\Delta_\sigma^m W_n(s, t; x)|_{\mathcal{B}}^{2q} \right]^{1/(2q)} \lesssim |\Delta_\sigma^m b_n(x, \cdot)|_{C^\kappa} |t - s|^{1/2}. \quad (4.4)$$

Setting  $m = \lceil \theta + \frac{d}{2q} \rceil + 1$ , it follows that

$$\begin{aligned} \mathbb{E}_{\lambda_n} \left[ |W_n(s, t; \cdot)|_{\mathcal{B}}^{2q} \right] &\lesssim \mathbb{E}_{\lambda_n} \left[ |W_n(s, t; \cdot)|_{B_{2q}^{\theta+d/(2q)}}^{2q} \right] \\ &= \mathbb{E}_{\lambda_n} \left[ \int |W_n(s, t; x)|^{2q} dx \right. \\ &\quad \left. + \int_{|\sigma| \leq 1} |\sigma|^{-2\theta q - 2d} \int |\Delta_\sigma^m W_n(s, t; x)|^{2q} dx d\sigma \right] \\ &\lesssim \int |b_n(x, \cdot)|_{C^\kappa}^{2q} |t - s|^q dx \\ &\quad + \int_{|\sigma| \leq 1} \int |\Delta_\sigma^m b_n(x, \cdot)|_{C^\kappa}^{2q} |t - s|^q dx d\sigma \\ &\lesssim |t - s|^q |b_n|_{B_{2q}^{\theta+d/(2q); C^\kappa}}^{2q} \lesssim |t - s|^q, \end{aligned}$$

where the first estimate follows from the embedding  $B_{2q}^{\theta+d/(2q)} \hookrightarrow C^\theta$ , the second from (4.4), the third from the definition of  $|\cdot|_{B_{2q}^{\theta+d/(2q)}; C^\kappa}$  [17, p. 4086], and the fourth from [17, Lemma 5.5] since  $b_n$  has uniformly bounded support and  $\sup_{n \geq 1} |b_n|_{C^{\alpha, \kappa}} < \infty$  with  $\alpha > \theta + d/(2q)$ .

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

$$\mathbb{E}_{\lambda_n} \left[ |\Delta_{x, \sigma}^m \Delta_{x', \sigma'}^{m'} \mathbb{W}_n(s, t; x, x')|^q \right]^{1/q} \lesssim |\Delta_\sigma^m b_n(x, \cdot)|_{C^\kappa} |\Delta_{\sigma'}^{m'} b_n(x', \cdot)|_{C^\kappa} |t - s|$$

and the argument from [17, p. 4089-4090] (note that this is where we require  $\sup_{n \geq 1} |b_n|_{C^{\alpha, \kappa}} < \infty$  for  $\alpha > \theta + d/q$ , so that  $\sup_{n \geq 1} |b_n|_{B_q^{\theta+d/q}} < \infty$ ).  $\square$

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

$$\sup_{n \geq 1} \mathbb{E}_{\lambda_n} [\|\mathbf{W}_n\|_{p\text{-var}}^{2q}] < \infty$$

and that  $(\mathbf{W}_n)_{n \geq 1}$  is a family of tight random variables in  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$ .

*Proof.* Consider  $\theta' \in (\theta, \alpha - \frac{d}{q})$  and the space  $\mathcal{B}' = C^{\theta'}(\mathfrak{K}, \mathbb{R}^d)$ , where  $\mathfrak{K} \times \Lambda$  contains the support of all  $b_n$ . We first show, using a similar argument as in the proof of Lemma 4.5, that  $(\mathbf{W}_n)_{n \geq 1}$  is tight in the  $(J_1)$  Skorokhod space  $D([0, 1], \mathcal{B} \oplus \mathcal{B}^{\otimes 2})$ . Indeed, Lemma 4.6 and (3.10) imply that  $(\mathbf{W}_n)_{n \geq 1}$  satisfies condition [7, Thm. 3.7.2(b)], and, since the embedding  $\mathcal{B}' \hookrightarrow \mathcal{B}$  is compact, that  $(\mathbf{W}_n)_{n \geq 1}$  satisfies the compact containment condition [7, Rem. 3.7.3]. Hence, by [7, Thm. 3.7.6],  $(\mathbf{W}_n)_{n \geq 1}$  is tight in  $D([0, 1], \mathcal{B} \oplus \mathcal{B}^{\otimes 2})$ .

Consider now  $p' \in (2, p)$ . The interpolation estimate (3.4) implies  $\sigma_{p\text{-var}}(\mathbf{X}, \mathbf{Y}) \leq \sigma_\infty(\mathbf{X}, \mathbf{Y})^{1-p'/p} (\|\mathbf{X}\|_{p'\text{-var}} + \|\mathbf{Y}\|_{p'\text{-var}})^{p'/p}$  with  $\sigma_\infty$  defined as in the proof of Lemma 4.5. Moreover, the map  $\mathbf{X} \mapsto \|\mathbf{X}\|_{p'\text{-var}}$  is invariant under reparametrizations and is lower semi-continuous under  $\|\cdot\|_\infty$  and thus under  $\sigma_\infty$ . Hence, for  $R > 0$  and a compact subset  $K$  of the  $(J_1)$  Skorokhod space  $D([0, 1], \mathcal{B} \oplus \mathcal{B}^{\otimes 2})$ , the set

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

is compact in  $\mathcal{D}^{p'\text{-var}}(\mathcal{B}')$ . Considering  $\mathbf{W}_n$  as an element of  $\mathcal{D}^{p'\text{-var}}(\mathcal{B}')$ , it follows from Lemma 4.6 and (3.9) (applied to these new parameters) that  $\sup_{n \geq 1} \mathbb{E}_{\lambda_n} [\|\mathbf{W}_n\|_{p'\text{-var}}^{2q}] < \infty$ . Consequently  $(\mathbf{W}_n)_{n \geq 1}$  is tight in  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$ .  $\square$

For an element  $\pi \in \mathbf{L}(\mathcal{B}, \mathbb{R}^m)$  and  $b \in C^{\theta, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$ , write  $\pi b : M \rightarrow \mathbb{R}^m$  for the function  $y \mapsto \pi(b(\cdot, y))$ . A direct verification shows that  $|\pi b|_{C^\kappa} \leq |\pi|_{\mathbf{L}(\mathcal{B}, \mathbb{R}^m)} |b|_{C^{\theta, \kappa}}$  (see, e.g., the proof of [17, Lem. 5.12]).

Consider the subspace of  $\mathbf{L}(\mathcal{B}, \mathbb{R})$

$$\tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}) = \text{span} \{b \mapsto D^k b^j(x) \mid x \in \mathbb{R}^d, k \in \mathbb{N}^d, |k| \leq 1, j \in \{1, \dots, d\}\}.$$

For  $m \geq 1$ , we denote by  $\tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^m)$  the subspace of  $\pi \in \mathbf{L}(\mathcal{B}, \mathbb{R}^m)$  such that  $\pi^i \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R})$  for every  $i = 1, \dots, m$ . We note that  $\tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R})$  does not appear in the work [17], however, due to the generality of our setting, we find it more convenient to work with than the full space  $\mathbf{L}(\mathcal{B}, \mathbb{R})$ .

Observe that, for  $b \in C_n^{\theta, \kappa}(\mathbb{R}^d \times \Lambda, \mathbb{R}^d)$ , the map  $x \mapsto b(x, \cdot)$  is a  $C^\theta$  map from  $\mathbb{R}^d$  into the closed subspace  $C_n^\kappa(\Lambda)$ , and thus  $\pi b \in C_n^\kappa(\Lambda)$  for all  $\pi \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R})$ .

**Lemma 4.8** (Finite-dimensional projections). *Let  $\pi \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^m)$  for some  $m \geq 1$  and suppose that Assumption 2.12 holds. Let  $\mathfrak{B}$  be defined as in Proposition 2.15, and let  $W_\pi$  be an  $\mathbb{R}^m$ -valued Brownian motion with covariance*

$$\mathbb{E}[W_\pi^i(1)W_\pi^j(1)] = \mathfrak{B}(\pi^i b_\infty, \pi^j b_\infty) + \mathfrak{B}(\pi^j b_\infty, \pi^i b_\infty).$$

Define further

$$\mathbb{W}_\pi^{i,j}(t) = \int_0^t W_\pi^i dW_\pi^j + \mathfrak{B}_2(\pi^i b_\infty, \pi^j b_\infty)t.$$

Then, as  $n \rightarrow \infty$ ,

$$(\pi W_n, (\pi \otimes \pi)\mathbb{W}_n) \rightarrow_{\lambda_n} (W_\pi, \mathbb{W}_\pi)$$

in the sense of finite-dimensional distributions.

*Proof.* By the preceding remarks,  $\pi b_n \in C_n^\kappa(\Lambda, \mathbb{R}^m)$  for  $n \in \mathbb{N} \cup \{\infty\}$  and  $\lim_{n \rightarrow \infty} |\pi b_n - \pi b_\infty|_{C^\kappa} = 0$ , so the conclusion follows by Assumption 2.12 and Proposition 2.15.  $\square$

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 [17, Rem. 5.14]).

**Proposition 4.9.** *Suppose that Assumptions 2.11 and 2.12 hold and that the support of  $a_n$  and  $b_n$  is uniformly bounded. Let  $\beta > 1$  and  $p > 2$ . Then there exists a random variable  $(V, \mathbf{W})$  in  $D^{\beta\text{-var}}(\mathcal{A}) \times \mathcal{D}^{p\text{-var}}(\mathcal{B})$  such that  $(V_n, \mathbf{W}_n) \rightarrow_{\lambda_n} (V, \mathbf{W})$ . Furthermore,  $(V, \mathbf{W})$  is a.s. continuous.*

*Proof.* By Lemma 4.5,  $\|V_n - V\|_{\beta\text{-var}} \rightarrow_{\lambda_n} 0$ , where  $V$  is deterministic and continuous. It remains to show that  $\mathbf{W}_n$  converges weakly to a limit point  $\mathbf{W}$  which is a.s. continuous. By Lemma 4.7,  $(\mathbf{W}_n)_{n \geq 1}$  is tight and thus weakly relatively compact by Prokhorov's theorem [2, Thm. 5.1]. Let  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$  be weak limit points of two subsequences of  $\mathbf{W}_n$ . Since the largest jump of  $W_n$  is of the order  $n^{-1/2}$  and the largest jump of  $t \mapsto \mathbb{W}_n(0, t)$  is of the order  $n^{-1/2} \sup_{t \in [0,1]} |W_n(t)|_{\mathcal{B}}$ , it follows that  $\mathbf{W}$  is a.s. continuous (and likewise for  $\tilde{\mathbf{W}}$ ).

We now show that  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$  have the same law. Consider the collection of  $\mathbb{R}$ -valued functions on  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$

$$\mathcal{F} := \left\{ (X, \mathbb{X}) \mapsto \sum_{j=1}^k \tau_j (\pi_j X(t_j), (\pi_j \otimes \pi_j)\mathbb{X}(t_j)) \right\},$$

where the parameters range over all  $k \geq 1$ ,  $\tau_j \in \mathbf{L}(\mathbb{R}^{m_j} \oplus (\mathbb{R}^{m_j})^{\otimes 2}, \mathbb{R})$ ,  $\pi_j \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^{m_j})$ ,  $m_j \geq 1$ , and  $t_j \in [0, 1]$ . For any  $f \in \mathcal{F}$ , it follows from Lemma 4.8 that  $f(\mathbf{W})$  and  $f(\tilde{\mathbf{W}})$  have the same law. In particular,  $\mathbb{E}[e^{if(\mathbf{W})}] = \mathbb{E}[e^{if(\tilde{\mathbf{W}})}]$  for all  $f \in \mathcal{F}$ . However, the collection of  $\mathbb{C}$ -valued functions  $\tilde{\mathcal{F}} := \{w \mapsto e^{if(w)} \mid f \in \mathcal{F}\}$  is a unital algebra of bounded functions on  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$  which separates points and is closed under conjugation. Moreover, every  $f \in \tilde{\mathcal{F}}$  is continuous on the subspace of continuous paths in  $\mathcal{D}^{p\text{-var}}(\mathcal{B})$ , and in particular on the support of  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$ . The laws of  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$  are Radon measures since they are obtained as weak limit points of tight sequences, hence, by the Stone–Weierstrass theorem and a compactification argument (see, e.g., [3, Ex. 7.14.79]),  $\mathbf{W}$  and  $\tilde{\mathbf{W}}$  have the same law.  $\square$

It remains to characterize the RDE driven by  $(V, \mathbf{W})$  as the solution to an SDE. We flesh out the abstract statement in the following lemma, which is a slight simplification of [17, Lem. 6.1].

**Lemma 4.10.** *Let  $X$  be the solution to the RDE*

$$dX = F(X)\bar{a} dt + H(X) d\mathbf{W}, \quad X(0) = \xi \in \mathbb{R}^d,$$

where  $\bar{a} \in \mathcal{A}$  is fixed and  $\mathbf{W} = (W, \mathbb{W})$  is a random  $p$ -rough path over  $\mathcal{B}$ ,  $p < \theta$ . Suppose that, for all  $m \geq 1$  and  $\pi \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^m)$ ,

$$(\pi W, (\pi \otimes \pi)\mathbb{W}) \sim (W_\pi, \mathbb{W}_\pi) \tag{4.5}$$

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

$$\mathbb{W}_\pi^{ij}(t) = \int_0^t W_\pi^i(s) dW_\pi^j(s) + \Gamma_\pi^{ij}t.$$

For every  $x \in \mathbb{R}^d$ , let us define  $\Sigma(x) := \Sigma_{H(x)}$  and, for  $i = 1, \dots, d$ ,

$$\Gamma^i(x) := \sum_{k=1}^d \Gamma_{H(x) \oplus DH(x)}^{k(ki)},$$

where we treat  $H(x) \oplus DH(x) \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^d \oplus (\mathbb{R}^d)^* \otimes \mathbb{R}^d)$ . Suppose further that

$$\sup_{x \in \mathbb{R}^d} \sum_{i=1}^d |\Sigma^{ii}(x)| + |\Gamma^i(x)| < \infty. \quad (4.6)$$

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  $\mathbf{W}$ . We first show that  $M : [0, 1] \rightarrow \mathbb{R}^d$  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

$$[M^i, M^j]_t = \int_0^t \Sigma^{ij}(X(s)) ds. \quad (4.7)$$

Indeed, the definition of the rough integral readily implies that  $X$  and  $M$  are adapted to  $\mathcal{F}$  (cf. [17, Lem 6.3]). Furthermore, for fixed  $0 \leq s < t \leq 1$ , we have

$$\begin{aligned} M(t) - M(s) &= \int_s^t H(X(u)) d\mathbf{W}(u) - \int_s^t \Gamma(X(u)) du \\ &= \lim_{|\mathcal{P}| \rightarrow 0} \sum_{[u,v] \in \mathcal{P}} M_{[u,v]}^{\mathcal{P}}, \end{aligned}$$

where the limit is taken over partitions  $\mathcal{P}$  of  $[s, t]$ , and

$$M_{[u,v]}^{\mathcal{P}} := H(X(u))W(u, v) + (H \otimes DH)(X(u))\mathbb{W}(u, v) - \Gamma(X(u))(v - u).$$

Note that the same argument as in [17, Lem. 6.2] implies that  $\pi W(u, v)$  and  $(\pi \otimes \pi)\mathbb{W}(u, v)$  are independent of  $\mathcal{F}_s$  for any  $\pi \in \tilde{\mathbf{L}}(\mathcal{B}, \mathbb{R}^m)$ . Taking  $\pi(x) = H(x) \oplus DH(x)$  in (4.5), it follows that

$$\mathbb{E}[M_{[u,v]}^{\mathcal{P}} | \mathcal{F}_u] = 0.$$

Furthermore, for  $i, j = 1, \dots, d$ ,

$$\mathbb{E}[H^i(X(u))W(u, v)H^j(X(u))W(u, v) | \mathcal{F}_u] = \Sigma^{ij}(X(u))(v - u)$$

and, by Itô isometry,

$$\mathbb{E}[|(H \otimes DH)(X(u))\mathbb{W}(u, v) - \Gamma(X(u))(v - u)|^2 | \mathcal{F}_u] \lesssim |v - u|^2,$$

where the proportionality constant depends only on  $\Sigma(X(u))$ . Using the bound (4.6), it follows that  $M$  is a martingale with quadratic variation (4.7) as claimed.

Let  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$  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

$$\varphi(X(t)) - \varphi(X(s)) - \int_s^t \left[ D\varphi(\bar{a} + \Gamma) + \frac{1}{2} D^2\varphi\Sigma \right](X(u)) du$$

is a martingale.  $\square$

*Proof of Theorem 4.3.* The fact that  $(V_n, \mathbf{W}_n) \rightarrow_{\lambda_n} (V, \mathbf{W})$ , where  $(V, \mathbf{W})$  is a.s. continuous, follows from Proposition 4.9. By Lemma 4.8,  $\mathbf{W}$  satisfies assumption (4.5) of Lemma 4.10 with  $\Sigma_{\pi}^{ij} = \mathfrak{B}(\pi^i b, \pi^j b) + \mathfrak{B}(\pi^j b, \pi^i b)$  and  $\Gamma_{\pi}^{ij} = \mathfrak{B}_2(\pi^i b, \pi^j b)$ . In particular,  $\Gamma$  in Lemma 4.10 is given by  $\Gamma^i(x) = \sum_{k=1}^d \mathfrak{B}_2(b^k(x, \cdot), \partial_k b^i(x, \cdot))$ . Furthermore,  $\mathfrak{B} = \frac{1}{2}\mathfrak{B}_1 + \mathfrak{B}_2$  is bounded by Assumption 2.12(ii), so  $\Sigma^{ii}(x) \lesssim |b(x, \cdot)|_{C^{\kappa}}^2 \leq |b|_{C^{0,\kappa}}^2$  and  $\Gamma^i(x) \lesssim |b(x, \cdot)|_{C^{\kappa}} |\nabla b(x, \cdot)|_{C^{\kappa}} \leq |b|_{C^{0,\kappa}} |b|_{C^{1,\kappa}}$ . Hence all the assumptions of Lemma 4.10 are verified, and the conclusion follows from [26, Thm. 4.5.2] by the equivalence of weak solutions to SDEs and the martingale problem.  $\square$

*Proof of Theorem 2.17.* This follows from Corollary 4.4 and the exact same localization argument as in [17, Sec. 7].  $\square$

## 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 [17] 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 [17]), 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 2.1 and fix parameters  $q > 1$ ,  $\kappa, \bar{\kappa} \in (0, 1)$ , and  $\alpha > 2 + \frac{d}{q}$ . Let  $a_{\varepsilon} \in C^{1+\bar{\kappa},0}(\mathbb{R}^d \times M, \mathbb{R}^d)$  and  $b_{\varepsilon}, b_0 \in C^{\alpha,\kappa}(\mathbb{R}^d \times M, \mathbb{R}^d)$ , for  $\varepsilon \in (0, 1]$ , such that

$$\sup_{\varepsilon \in (0,1]} |a_{\varepsilon}|_{C^{1+\bar{\kappa},0}} + |b_{\varepsilon}|_{C^{\alpha,\kappa}} < \infty, \quad \lim_{\varepsilon \rightarrow 0} |b_{\varepsilon} - b_0|_{C^{\alpha,\kappa}} = 0.$$

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  $g_{\varepsilon} : M \rightarrow TM$  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  $\bar{a} \in C^{1+\bar{\kappa}}(\mathbb{R}^d, \mathbb{R}^d)$  such that, for all  $t \in [0, 1]$  and  $x \in \mathbb{R}^d$ ,*

$$|V_{\varepsilon}(t)(x) - t\bar{a}(x)| \rightarrow_{\lambda_{\varepsilon}} 0 \quad \text{as} \quad \varepsilon \rightarrow 0,$$

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

$$W_{v,\varepsilon}(t) = \varepsilon \int_0^{t\varepsilon^{-2}} v \circ g_{\varepsilon,s} ds, \quad \mathbb{W}_{v,w,\varepsilon}(s,t) = \int_s^t (W_{v,\varepsilon}(r) - W_{v,\varepsilon}(s)) \otimes dW_{w,\varepsilon}(r).$$

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

Recall our notational convention about subspaces  $C_\varepsilon^\kappa(M)$  of  $C^\kappa(M)$  introduced before Assumption 2.3.

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

(i) *for all  $v = (v_\varepsilon), w = (w_\varepsilon) \in \prod_{\varepsilon \in (0,1]} C_\varepsilon^\kappa(M)$  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  $\varepsilon > 0$ ,*

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

(ii) *There exists a bounded bilinear operator  $\mathfrak{B} : C_0^\kappa(M) \times C_0^\kappa(M) \rightarrow \mathbb{R}$  such that for every  $m \geq 1$  and all  $\mathbf{v} = (v_\varepsilon)_{\varepsilon \in [0,1]}$  with  $v_\varepsilon \in C_\varepsilon^\kappa(M, \mathbb{R}^m)$  and  $\lim_{\varepsilon \rightarrow 0} |v_\varepsilon - v_0|_{C^\kappa} = 0$ , it holds that  $(W_{v_\varepsilon,\varepsilon}, \mathbb{W}_{v_\varepsilon,\varepsilon}) \rightarrow_{\lambda_\varepsilon} (W_{\mathbf{v}}, \mathbb{W}_{\mathbf{v}})$  as  $\varepsilon \rightarrow 0$  in the sense of finite-dimensional distributions, where  $W_{\mathbf{v}}$  is an  $\mathbb{R}^m$ -valued Brownian motion and*

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

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

$$|W_{v_\varepsilon,\varepsilon}(t)|_{L^{2q}(\lambda_\varepsilon)} \leq Kt^{1/2} \quad \text{and} \quad |\mathbb{W}_{v_\varepsilon,w_\varepsilon,\varepsilon}(0,t)|_{L^q(\lambda_\varepsilon)} \leq Kt \quad \text{for all } \varepsilon, t \in (0, 1]$$

*imply Assumption 5.2(i).*

**Remark 5.4.** *As in Proposition 2.15, one can show that Assumption 5.2 implies that the covariance of  $W_{\mathbf{v}}$  is given by*

$$\mathbb{E}[W_{\mathbf{v}}^i(1)W_{\mathbf{v}}^j(1)] = \mathfrak{B}(v_0^i, v_0^j) + \mathfrak{B}(v_0^j, v_0^i).$$

*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  $\mathfrak{B}$  is bounded in Assumption 5.2 since this follows automatically (see [17, Prop. 2.8]).*

Consider the quadratic form

$$\Sigma^{ij}(x) = \mathfrak{B}(b_0^i(x, \cdot), b_0^j(x, \cdot)) + \mathfrak{B}(b_0^j(x, \cdot), b_0^i(x, \cdot)), \quad i, j = 1, \dots, d. \quad (5.1)$$

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  $\tilde{a} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ .

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  $\xi_\varepsilon \rightarrow \xi \in \mathbb{R}^d$ . Then  $x_\varepsilon \rightarrow_{\lambda_\varepsilon} X$  in the uniform topology as  $\varepsilon \rightarrow 0$ , where  $X$  is the unique weak solution of the SDE*

$$dX = \tilde{a}(X) dt + \sigma(X) dB, \quad X(0) = \xi. \quad (5.2)$$

Here,  $B$  is a standard Brownian motion in  $\mathbb{R}^d$ ,  $\sigma$  is the unique positive semi-definite square root of  $\Sigma$  given by (5.1), and  $\tilde{a}$  is the Lipschitz function given by

$$\tilde{a}^i(x) = \bar{a}^i(x) + \sum_{k=1}^d \mathfrak{B}(b_0^k(x, \cdot), \partial_k b_0^i(x, \cdot)), \quad i = 1, \dots, d.$$

## A Rough path Besov-variation embedding

We adapt Friz–Victoir [11, 12] 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  $\mathcal{B}^{\otimes 2}, \dots, \mathcal{B}^{\otimes N}$  with a system of admissible tensor norms. For a continuous multiplicative function  $\mathbf{W} = (1, \mathbf{W}^1, \dots, \mathbf{W}^N) : [0, T]^2 \rightarrow \oplus_{k=0}^N \mathcal{B}^{\otimes k}$  define the homogeneous Besov norm

$$\|\mathbf{W}\|_{W^{\alpha, q}; [s, t]}^q := \sum_{k=1}^N \iint_{[s, t]^2} \frac{|\mathbf{W}_{v, u}^k|_{\mathcal{B}^{\otimes k}}^{q/k}}{|u - v|^{q\alpha+1}} du dv.$$

**Proposition A.1.** *Suppose  $q > 1$  and  $\alpha \in (\frac{1}{q}, 1)$ . There exists a constant  $C = C(\alpha, q, N)$  such that*

$$\sum_{k=1}^N |\mathbf{W}_{s, t}^k|^{q/k} \leq C |t - s|^{q\alpha-1} \|\mathbf{W}\|_{W^{\alpha, q}; [s, t]}^q.$$

*Proof.* We follow a similar strategy to [12, Proposition A.9]. We proceed by induction on  $N$ . The case  $N = 1$  follows directly from the GRR lemma [12, Corollary A.2]. Suppose the result is true for  $N - 1$ . Since both sides scale homogeneously with dilations, we may suppose that  $\|\mathbf{W}\|_{W^{\alpha, q}; [s, t]}^q \leq 1$ . 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  $\Upsilon_{s, t} = \sup_{u, v \in [s, t]} \frac{|\mathbf{W}_{u, v}|}{|v - u|^{N/p}}$ , and observe that it suffices to show  $\Upsilon_{s, t} \leq C$ . We have

$$\mathbf{W}_{s, v}^N - \mathbf{W}_{s, u}^N = \mathbf{W}_{u, v}^N + \sum_{j=1}^{N-1} \mathbf{W}_{s, u}^{N-j} \otimes \mathbf{W}_{u, v}^j,$$

and thus

$$\left( \iint \frac{|\mathbf{W}_{s, u}^N - \mathbf{W}_{s, v}^N|^q}{|v - u|^{q\alpha+1}} du dv \right)^{1/q} \leq \Delta_1 + \Delta_2,$$

where

$$\begin{aligned} \Delta_1 &= \sum_{j=1}^{N-1} \left( \iint |\mathbf{W}_{s, u}^{N-j}|^q \frac{|\mathbf{W}_{u, v}^j|^q}{|u - v|^{q\alpha+1}} du dv \right)^{1/q}, \\ \Delta_2 &= \left( \iint \frac{|\mathbf{W}_{u, v}^N|^q}{|u - v|^{q\alpha+1}} du dv \right)^{1/q}. \end{aligned}$$

For  $\Delta_1$ , by the inductive hypothesis, we have  $|\mathbf{W}_{s, u}^{(N-j)}|^q \leq |t - s|^{q(N-j)/p}$ , so that

$$\Delta_1 \leq \sum_{j=1}^{N-1} |t - s|^{(N-j)/p} \left( \iint \frac{|\mathbf{W}_{u, v}^j|^q}{|u - v|^{q\alpha+1}} du dv \right).$$



Again by the inductive hypothesis, we have

$$|\mathbf{W}_{u,v}^j|^{q(1-1/j)} \leq |t-s|^{q(j-1)/p},$$

so that

$$\Delta_1 \leq \sum_{j=1}^N |t-s|^{(N-1)/p} \left( \iint \frac{|\mathbf{W}_{u,v}^j|^{q/j}}{|u-v|^{q\alpha+1}} du dv \right) \leq \sum_{j=1}^N |t-s|^{(N-1)/p}.$$

For  $\Delta_2$ , we have

$$\begin{aligned} \Delta_2 &\leq \left( \iint \Upsilon_{s,t}^{q(1-1/N)} |t-s|^{q(N-1)/p} \frac{|\mathbf{W}_{u,v}^N|^{q/N}}{|v-u|^{q\alpha+1}} du dv \right)^{1/q} \\ &\leq \Upsilon_{s,t}^{1-1/N} |t-s|^{(N-1)/p}. \end{aligned}$$

Combining the above two estimates, we have

$$\left( \iint \frac{|\mathbf{W}_{s,u}^N - \mathbf{W}_{s,v}^N|^q}{|v-u|^{q\alpha+1}} du dv \right)^{1/q} \leq C |t-s|^{(N-1)/p} (1 + \Upsilon_{s,t}^{1-1/N}).$$

Applying the GRR lemma to the continuous path  $\mathbf{W}_{s,\cdot}^N : [s, t] \rightarrow \mathcal{B}^{\otimes N}$  we have

$$\begin{aligned} |\mathbf{W}_{s,t}^N| &\leq C |t-s|^{1/p} |t-s|^{(N-1)/p} (1 + \Upsilon_{s,t}^{1-1/N}) \\ &\leq C |t-s|^{N/p} (1 + \Upsilon_{s,t}^{1-1/N}). \end{aligned}$$

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  $\gamma \in (0, 1]$

$$\|\mathbf{W}\|_{\gamma\text{-Höl};[s,t]} := \sum_{k=1}^N \sup_{u,t \in [s,t]} \frac{|\mathbf{W}_{v,u}^k|^{1/k}}{|u-v|^\gamma}.$$

**Corollary A.2.** *Let  $q > 1$  and  $\alpha \in (\frac{1}{q}, 1)$ . There exists a constant  $C = C(\alpha, q, N)$  such that*

$$\|\mathbf{W}\|_{(\alpha-1/q)\text{-Höl};[s,t]} \leq C \|\mathbf{W}\|_{W^{\alpha,q};[s,t]}.$$

*Proof.* Immediate from Proposition A.1.  $\square$

Recall the homogeneous  $p$ -variation “norm” for  $p \geq 1$

$$\|\mathbf{W}\|_{p\text{-var};[s,t]}^p := \sup_{\mathcal{P}} \sum_{[u,v] \in \mathcal{P}} \sum_{k=1}^N |\mathbf{W}_{u,v}^k|^{p/k},$$

where the supremum runs over all partitions  $\mathcal{P}$  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*

$$\|\mathbf{W}\|_{1/\alpha\text{-var};[s,t]} \leq C |t-s|^{\alpha-1/q} \|\mathbf{W}\|_{W^{\alpha,q};[s,t]}.$$

*Proof.* By Proposition A.1 we have for all  $u, v \in [s, t]$  and  $k = 1, \dots, N$

$$\begin{aligned} |\mathbf{W}_{u,v}^k|^{\frac{1}{\alpha k}} &= \left( |\mathbf{W}_{u,v}^k|^{q/k} \right)^{\frac{1}{\alpha q}} \\ &\leq C (|u - v|^{q\alpha - 1})^{\frac{1}{q\alpha}} \left( \|\mathbf{W}\|_{W^{\alpha,q};[u,v]}^q \right)^{\frac{1}{q\alpha}}. \end{aligned}$$

Note however that  $\omega_1(u, v) = |u - v|$  and  $\omega_2(u, v) := \|\mathbf{W}\|_{W^{\alpha,q};[u,v]}^q$  are controls, and thus so is  $\omega := \omega_1^{1 - \frac{1}{q\alpha}} \omega_2^{\frac{1}{q\alpha}}$ . Hence

$$\|\mathbf{W}\|_{1/\alpha\text{-var};[s,t]}^{1/\alpha} \leq \omega(s, t),$$

from which the conclusion follows.  $\square$

**Remark A.4.** *Besov (rough path) regularity effectively interpolates between the well-known Hölder- and  $p$ -variation cases, see [9] for a discussion.*

## References

- [1] I. Bailleul and R. Catellier. Rough flows and homogenization in stochastic turbulence. *J. Differential Equations* **263** (2017) 4894–4928.
- [2] P. Billingsley. *Convergence of probability measures*. Second edition. Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., New York, 1999.
- [3] V. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, 2007.
- [4] I. Chevyrev. Random walks and Lévy processes as rough paths. *Probab. Theory Related Fields* **170** (2018) 891–932.
- [5] I. Chevyrev and P. K. Friz. Canonical RDEs and general semimartingales as rough paths. *Ann. Probab.* **47** (2019) 420–463.
- [6] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang. Multiscale systems, homogenization, and rough paths. *Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016* (P. Friz et al., ed.). Springer Proceedings in Mathematics & Statistics **283**, Springer, 2019, pp. 17–48.
- [7] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 2005.
- [8] P. K. Friz and M. Hairer. *A course on rough paths*. Universitext, Springer, Cham, 2014.
- [9] P. K. Friz and D. Prömel. Rough path metrics on a Besov–Nikolskii-type scale. *Trans. Amer. Math. Soc.* **370** (2018) 8521–8550.
- [10] P. K. Friz and A. Shekhar. General rough integration, Lévy rough paths and a Lévy–Kintchine-type formula. *Ann. Probab.* **45** (2017) 2707–2765.
- [11] P. K. Friz and N. Victoir. A variation embedding theorem and applications. *J. Funct. Anal.* **239** (2006) 631–637.
- [12] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*. Cambridge Stud. Adv. Math. **120**, 2010.

- [13] P. K. Friz and H. Zhang. Differential equations driven by rough paths with jumps. *J. Differential Equations* **264** (2018) 6226–6301.
- [14] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. London A* **469** (2013) 20130201.
- [15] D. Kelly. Rough path recursions and diffusion approximations. *Ann. Appl. Probab.* **26** (2016) 425–461.
- [16] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520.
- [17] D. Kelly and I. Melbourne. Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise. *J. Funct. Anal.* **272** (2017) 4063–4102.
- [18] A. Korepanov, Z. Kosloff and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **35** (2018) 859–885.
- [19] A. Korepanov, Z. Kosloff and I. Melbourne. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1. Preprint, 2020. To appear in *Ann Inst. H. Poincaré (B) Probab. Statist.*
- [20] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999) 671–685.
- [21] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** (1998) 215–310.
- [22] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity* **24** (2011) 1361–1367.
- [23] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197.
- [24] R. Ryan. *Introduction to tensor products of Banach spaces*. Springer Monographs in Mathematics, Springer-Verlag London, 2002.
- [25] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967) 747–817.
- [26] D. Stroock and S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag, Berlin, 2006.
- [27] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.
- [28] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188.