The Annals of Probability
2020, Vol. 48, No. 4, 1826–1860
https://doi.org/10.1214/19-AOP1408

AVERAGING DYNAMICS DRIVEN BY FRACTIONAL BROWNIAN MOTION

BY MARTIN HAIRER* AND XUE-MEI LI**

Imperial College London, *m.hairer@imperial.ac.uk; **xue-mei.li@imperial.ac.uk

We consider slow/fast systems where the slow system is driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$. We show that unlike in the case $H = \frac{1}{2}$, convergence to the averaged solution takes place in probability and the limiting process solves the 'naïvely' averaged equation. Our proof strongly relies on the recently obtained stochastic sewing lemma.

1. Introduction. The purpose of the paper is to study a two-scale stochastic evolution on $\mathbf{R}^d$ with memory of the type

$$dx_t^\varepsilon = f(x_t^\varepsilon, y_t^\varepsilon)\, dB_t + g(x_t^\varepsilon, y_t^\varepsilon)\, dt, \tag{1.1}$$

where B is an m-dimensional fractional Brownian motion (fBm) of Hurst parameter $H > \frac{1}{2}$, $f : \mathbf{R}^d \times \mathcal{Y} \to L(\mathbf{R}^m, \mathbf{R}^d)$ and $g : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}^d$. The fast variable y_t is assumed to take values in a state space $\mathcal{Y}$ which is either an arbitrary Polish space or a compact manifold, depending on the situation. We will consider both the case in which the dynamic of y is given, independently of that of x, and the case in which the current state of x influences the dynamic of y. In the latter case, we will assume that the dynamic of y is Markovian, conditional on B.

We recall that one-dimensional fractional Brownian motion is the centred Gaussian process with $B_0 = 0$ and covariance $\mathbf{E}(B_t - B_s)^2 = |t - s|^{2H}$. An $\mathbf{R}^m$-valued fBm $(B_t^1, \dots, B_t^m)$ is obtained by taking i.i.d. copies of a one-dimensional fBm. A fBm is not a semimartingale and does not have independent increments. It does however have a version such that almost all of its sample paths $t \mapsto B_t(\omega)$ are Hölder continuous of order α for any $\alpha < H$.

Let us first consider the simple case in which the fast process y has no feedback from x and is of the form $y_t^\varepsilon = Y_{t/\varepsilon}$ for some process Y which is almost surely Hölder continuous of order α with $\alpha + H > 1$. The integral appearing in (1.1) can then be interpreted as a Young integral. For the processes $\{x_\cdot^\varepsilon, \varepsilon > 0\}$ to have a limit, we would at the very least need uniform bounds. The usual Young bound however only gives an estimate of the form

$$\left| \int_0^t f(x_s^\varepsilon, y_s^\varepsilon)\, db_s \right| \lesssim |f(x_\cdot^\varepsilon, y_\cdot^\varepsilon)|_\alpha |b|_\beta,$$

which is not very helpful since the process y^ε is in general expected to have a Hölder norm of order $\frac{1}{\varepsilon^\alpha}$. Proving these bounds present unexpected difficulties. In the case where B is a Brownian motion, the desired estimates follow quite easily from Itô's isometry and/or the Burkholder–Davis–Gundy inequality. They are of course not available in our setting, but we would nevertheless like to exploit the stochastic nature of the fractional Brownian motion. We resolve this problem by using a carefully chosen approximation to the Young integral and using a recently discovered stochastic sewing lemma by Lê [23]. Our main result in this setting is given by Theorem 3.13 below. When combining it with Lemma 3.14, this can be formulated as follows.

Received March 2019.
MSC2010 subject classifications. 60G22, 60H10, 60H05.
Key words and phrases. Fractional Brownian motion, averaging, slow/fast system, sewing lemma.

THEOREM A. *Let $\mathcal{Y}$ be a Polish space and let f, g be bounded measurable and of class BC^2 in their first argument. Let $y_t^\varepsilon = Y_{t/\varepsilon}$ for a $\mathcal{Y}$-valued stationary stochastic process Y that is independent of B and is strongly mixing with rate $t^{-\delta}$ for some $\delta > 0$ in the sense that*

$$\sup\{\mathbf{P}(A \cap \bar{A}) - \mathbf{P}(A)\mathbf{P}(\bar{A}) : A \in \sigma(Y_0), \bar{A} \in \sigma(Y_t)\} \lesssim t^{-\delta}.$$

Let $\bar{f}(x) = \int f(x, y)\mu(dy)$, where μ is the law of Y_0, and similarly for g. Then, any solutions to (1.1) converge in probability to the solution to

$$d\bar{x}_t = \bar{f}(\bar{x}_t)\, dB_t + \bar{g}(\bar{x}_t)\, dt, \tag{1.2}$$

with the same initial value.

REMARK 1.1. This is very different from the case where B is a Wiener process. In that case, one cannot expect convergence in probability and the weak limit solves a diffusion with averaged generator; see, for example, [7, 15, 19, 20, 25, 26, 31, 33], which is different in general from the diffusion with averaged diffusion coefficients appearing here. In this sense, equations driven by fBm with $H > \frac{1}{2}$ behave more like ODEs rather than SDEs.

Note however that our convergence in probability refers to convergence in probability in the full 'product' probability space on which *both* B and Y live. In particular, we do not know whether the convergence to $\bar{x}$ holds with B replaced by any given $b \in \mathcal{C}^\beta$ with $\beta < H$. In the lingo of diffusions in random environment, our convergence result is 'annealed' rather than 'quenched'.

REMARK 1.2. It would be natural to take for y the solution to a SDE driven by a fractional Brownian motion independent of B. Unfortunately, it is not clear whether the results of [2, 12, 14] concerning the ergodicity of such processes can be strengthened in order to satisfy the assumptions of Theorem A.

The other case we consider is when the state of the slow variable x feeds back into the dynamic of the fast variable y. In this case, we restrict ourselves to the case when $\mathcal{Y}$ is a compact Riemannian manifold and y is given by the solution to

$$dy_t^\varepsilon = \frac{1}{\varepsilon} V_0(x_t^\varepsilon, y_t^\varepsilon)\, dt + \frac{1}{\sqrt{\varepsilon}} V(x_t^\varepsilon, y_t^\varepsilon) \circ d\hat{W}_t, \tag{1.3}$$

where, for any fixed value $x \in \mathbf{R}^d$, $V_i(x, \cdot)$ are vector fields on a state space $\mathcal{Y}$, and where $\hat{W}$ is a $\hat{m}$-dimensional standard Wiener process that is independent of B.

Since solutions to this equation are expected to be Hölder continuous of any order $\alpha < \frac{1}{2}$, the integral with respect to B appearing in (1.1) can still be interpreted as a Young integral for any fixed $\varepsilon > 0$. Since the slow and fast variables interact with each other, however, a solution theory with mixed Young and Itô integrals must be used. Such a theory is available in the literature; see, for example, the work by Guerra–Nualart [11] extending Kubilius [21], as well as [5]. Our main theorem is the following result, which is a slight reformulation of Theorem 4.3 below.

THEOREM B. *Let f, g and the V_i satisfy Assumption 4.1 below, let B_t be a fBm of Hurst parameter $H > \frac{1}{2}$ and let $\hat{W}_t$ be an independent Brownian motion. For every $x \in \mathbf{R}^d$, let μ^x denote the (unique) invariant measure for (1.3) with x_t^ε replaced by x. As before, let $\bar{f}(x) = \int f(x, y)\mu^x(dy)$, and similarly for $\bar{g}$.*

Then, as $\varepsilon \to 0$, the process x_t^ε converges in probability in $\mathcal{C}^\alpha$ (for any $\alpha < H$) to the unique limit $\bar{x}_t$ solving (1.2) with the same initial value.

1.1. *Outline of the article.* In both cases, the proof of convergence of the slow variable is based on a deterministic residue bound, Lemma 2.2. This is a quite general statement about differential equations with Young integration, not involving any stochastic element nor needing any preparation, and is therefore given at the very beginning of the article, even though it becomes relevant only in the later stages.

Section 3 is devoted to the proof of Theorem A. In order to prepare for this, we use the Mandelbrot–Van Ness representation of B_t by a Wiener process W_t. We then make use of the observation, [12], that the filtration $\mathcal{G}_t$ generated by the increments of B up to time t is the same as that generated by W, and that B can be decomposed for $t > 0$ as $B_t = \bar{B}_t + \tilde{B}_t$, where $\bar{B}_t$ is smooth in t and $\tilde{B}_t$ is independent of $\mathcal{G}_0$. Such a split can be made with reference to any $\mathcal{G}_u$ for any time u, which allows us to define integrals of the type

$$\int_u^v F(s)\, dB(s),$$

for $\mathcal{G}_u$-measurable processes F, as the sum of a Wiener integral against $\tilde{B}_t$ and a Riemann–Stieltjes integral against the smooth function $\bar{B}_t$. The stochastic sewing lemma then allows us to extend this integration to a class of adapted integrands which are allowed to be quite singular (much more than what Young integration would allow), but such that the singular part of their behaviour is independent of B in a suitable sense. This is the content of Lemma 3.10, which is the main ingredient of the proof of Theorem A given in Section 3.4.

Section 4 is devoted to the proof of Theorem B. The main ingredient of the proof is given by Theorem 4.16 where we show that one has a bound

$$\left\| \int_s^t \big(h(x_r, y_r^\varepsilon) - \bar{h}(x_r)\big)\, dB_r \right\|_{L^p} \le C\varepsilon^\kappa \big(\|x\|_{\alpha,p}|t-s|^{\bar{\eta}} + |t-s|^\eta\big),$$

for some $\eta > \frac{1}{2}$ and $\bar{\eta} > 1$, where $\bar{h}$ is the average of h. Compared to the results in Section 3, the difficulty here is that the process y^ε does depend on x (and therefore also on B) via (1.3). The main idea is to interpret the integral appearing in this expression as the output of the stochastic sewing lemma applied to

$$A_{u,v} = \int_u^v \big(h\big(x_u, Y_r^{x_u,\varepsilon}\big) - \bar{h}(x_u)\big)\, dB_r,$$

where $Y_r^{\bar{x},\varepsilon}$ denotes the solution to (1.3), but with the process x replaced by the fixed value $\bar{x}$. In this way, the integrand is $\mathcal{F}_u$-measurable (for $\mathcal{F}$ the filtration generated by B and $\hat{W}$) and the integral can be interpreted as a mixed Wiener/Young integral as before.

The hard part is to show that δA satisfies the assumptions of the stochastic sewing lemma. For this, we use the fact that we only need bounds on $\mathbf{E}(\delta A_{s,u,t}|\mathcal{F}_s)$ and that this quantity is much better behaved than δA itself. Section 4.4 contains preliminary estimates on the Markov semigroup generated by $Y_r^{\bar{x},\varepsilon}$ as well as some form of ‘nonautonomous Markov semigroup’ generated by (1.3), while Section 4.6 then contains the uniform bounds on the conditional expectation of δA.

Notation. We gather here the most common notation:

- $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $\|\cdot\|_p$ denotes the norm in $L^p(\Omega)$.
- For $s \le t$ and x_t a one-parameter process with values in $\mathbf{R}^d$, we set $\delta x_{s,t} \stackrel{\text{def}}{=} x_t - x_s$. We also set $\|x\|_{\alpha,p} = \sup_{s,t} |t-s|^{-\alpha}\|\delta x_{s,t}\|_p$.
- For $s < u < t$ and A, a two-parameter stochastic process, we set

$$\delta A_{\text{sut}} \stackrel{\text{def}}{=} A_{s,t} - A_{s,u} - A_{u,t}.$$

We also set

$$\|A\|_{\alpha,p} \stackrel{\text{def}}{=} \sup_{s<t} \frac{\|A_{s,t}\|_p}{|t-s|^\alpha}, \qquad \|\!|A|\!\|_{\alpha,p} \stackrel{\text{def}}{=} \sup_{s<u<t} \frac{\|\mathbf{E}(\delta A_{\mathrm{sut}}|\mathcal{F}_s)\|_p}{|t-s|^\alpha}.$$

- $H_\eta^p = \{A_{s,t} \in L_p(\Omega, \mathcal{F}_t, \mathbf{P}) : \|A\|_{\eta,p} < \infty\}$.
- $\mathcal{B}_{\alpha,p} = \{x_t \in \mathcal{F}_t : \delta x_{s,t} \in H_\alpha^p\}$, $\mathcal{B}_{\alpha,p} \subset L_p(\Omega, \mathcal{C}^\gamma)$ (up to modification) for $\gamma < \alpha - \frac{1}{p}$.
- $\bar{H}_\eta^p = \{A_{s,t} : \|\!|A|\!\|_{\eta,p} < \infty\}$.
- W_t and $\hat{W}_t$ are two independent two-sided Wiener processes of dimension m and $\hat{m}$, respectively.
- $\mathcal{G}_t$ and $\hat{\mathcal{G}}_t$ are the filtrations generated by the independent Wiener processes W and $\hat{W}$, respectively, and $\mathcal{F}_t = \mathcal{G}_t \vee \hat{\mathcal{G}}_t$.
- B_t (also denoted by B_t^H) is a fractional Brownian motion of Hurst parameter H, which is related to W_t via the Mandelbrot–Van Ness representation.
- For $u < t$, $\bar{B}_t^u \stackrel{\text{def}}{=} \mathbf{E}(B_t - B_u|\mathcal{G}_u)$ and $\tilde{B}_t^u \stackrel{\text{def}}{=} B_t - B_u - \bar{B}_t^u$.
  Also $\bar{B}_t \stackrel{\text{def}}{=} \bar{B}_t^0$, $\tilde{B}_t = \tilde{B}_t^0$.
- $f \lesssim g$ means that $f \le Cg$ for a universal constant C.
- $\mathcal{C}_K^\infty$ denotes the space of smooth functions with compact support.
- BC^k is the space of bounded $\mathcal{C}^k$ functions with bounded derivatives of all orders up to k.
- For $\alpha \in (0,1)$, $|x|_\alpha = \sup_{s\neq t} \frac{|x_t - x_s|}{|t-s|^\alpha}$ is the homogeneous Hölder semi-norm.
- $|\cdot|_\infty$ and $|\cdot|_{\mathrm{Lip}}$ denote the supremum norm and minimal Lipschitz constant, respectively.
- $|f|_{\mathrm{Osc}} = \sup f - \inf f$.
- For $h \in \mathcal{C}(\mathbf{R}, \mathbf{R}^d)$, $\kappa \in (0,1)$, $|h|_{-\kappa} = \sup_{s,t\le T} |t-s|^{\kappa-1} |\int_s^t h(r)\,dr|$.
- For $f : \mathbf{R} \times \mathbf{R}^d \to \mathbf{R}$, $|f|_{-\kappa,\gamma}$ is the smallest possible choice of constant K with the property

$$\sup_x \big|f(\cdot,x)\big|_{-\kappa} \le K, \qquad \sup_{x\neq y} \frac{|f(\cdot,x) - f(\cdot,y)|_{-\kappa}}{|x-y|^\gamma} \le K.$$

- We write $\mathcal{B}(\mathcal{Y})$ for the Borel σ-algebra of a topological space $\mathcal{Y}$. For $s < u$, we set

$$\mathcal{U}_s^u = \big\{F : \Omega \times \mathcal{Y} \to \mathbf{R} : \text{ bounded } (\mathcal{F}_s \vee \mathcal{G}_u) \otimes \mathcal{B}(\mathcal{Y}) \text{ measurable}\big\}.$$

2. A deterministic residual bound. We first state a bound on the difference between two solutions to a differential equation driven by a Hölder continuous signal, given a bound on the corresponding residual. In the following, the reader may think of b_t as a realisation of the fractional Brownian motion B_t or a realisation of $(B_t, t) \in \mathbf{R}^{m+1}$, but our statement is purely deterministic.

A basic tool is the following estimate, the proof of which is elementary and follows for example easily from [9], equation (2.8).

LEMMA 2.1. *Assume that $F : \mathbf{R}^d \to \mathbf{R}$ has two bounded derivatives and let $\alpha \in (0,1)$. Then the composition operator $x \mapsto F(x) = (t \mapsto F(x_t))$ satisfies the bound*

$$\big|F(x) - F(y)\big|_\alpha \lesssim \big|F'\big|_\infty |x-y|_\alpha + \big|F''\big|_\infty |x-y|_\infty \big(|x|_\alpha + |y|_\alpha\big).$$

The announced bound goes as follows.

LEMMA 2.2. *Let $F \in \mathrm{BC}^2$, let $b \in \mathcal{C}^\beta$ for some $\beta > \frac{1}{2}$ and let $Z, \bar{Z} \in \mathcal{C}^\alpha$ for some $\alpha \in (0,\beta]$ such that $\alpha + \beta > 1$. Let $z, \bar{z}$ be the solutions to*

$$z_t = Z_t + \int_0^t F(z_s)\,db_s, \qquad \bar{z}_t = \bar{Z}_t + \int_0^t F(\bar{z}_s)\,db_s.$$

Then there exists a constant C depending only on F such that, on the time interval $[0, 1]$, *one has the bound*

$$|z-\bar z|_\alpha \le C\exp\big(C|b|_\beta^{1/\beta} + C|Z|_\alpha^{1/\alpha} + C|\bar Z|_\alpha^{1/\alpha}\big)|Z-\bar Z|_\alpha.$$

PROOF. Since, by Lemma 2.1, we have the bound

$$\big|F(z)-F(\bar z)\big|_\alpha \lesssim |z-\bar z|_\alpha\big|F'\big|_\infty + \big|F''\big|_\infty |z-\bar z|_\infty\big(|z|_\alpha+|\bar z|_\alpha\big), \tag{2.1}$$

we conclude that on $[0, T]$ with $T \le 1$ one has

$$|z-\bar z|_\infty \lesssim \big(T^\beta L|z-\bar z|_\infty + \big|F'\big|_\infty T^{\alpha+\beta}|z-\bar z|_\alpha\big)|b|_\beta + |Z-\bar Z|_\infty,$$
$$|z-\bar z|_\alpha \lesssim \big(T^{\beta-\alpha} L|z-\bar z|_\infty + \big|F'\big|_\infty T^{\beta}|z-\bar z|_\alpha\big)|b|_\beta + |Z-\bar Z|_\alpha,$$

where $L = |F|_{\mathrm{Lip}} + T^\alpha |F''|_\infty(|z|_\alpha + |\bar z|_\alpha)$. The two inequalities are proved similarly, we demonstrate with the second one. By (3.11) in Section 3, we obtain on $[0, T]$ the bound

$$\begin{aligned}|z-\bar z|_\alpha &\le \left|\int_0^\cdot \big(F(z_s)-F(\bar z_s)\big)\,db_s\right|_\alpha + |Z-\bar Z|_\alpha \\ &\lesssim \big|F(z)-F(\bar z)\big|_\alpha T^\beta|b|_\beta + \big|F(z)-F(\bar z)\big|_\infty T^{\beta-\alpha}|b|_\beta + |Z-\bar Z|_\alpha,\end{aligned}$$

and the requested bound then follows from (2.1). In a similar way, using the fact that $|F(z)|_\alpha \le |F'|_\infty |z|_\alpha$, we obtain the a priori bound

$$|z|_\alpha \lesssim |Z|_\alpha + T^{\beta-\alpha}|b|_\beta + T^\beta|b|_\beta|z|_\alpha,$$

and similarly for $\bar z$. Provided that we choose T in such a way that

$$T^\beta|b|_\beta \le c, \qquad T^\alpha|Z|_\alpha \le 1, \qquad T^\alpha|\bar Z|_\alpha \le 1, \tag{2.2}$$

for some sufficiently small constant c that only depends on F, we thus obtain the bound $|z|_\alpha \lesssim |Z|_\alpha + T^{\beta-\alpha}|b|_\beta$, and similarly for $|\bar z|_\alpha$. In particular, this shows that for T as in (2.2) one has $L \lesssim 1$.

This then suggests the introduction of the norm

$$|z|_{\alpha,T} = |z|_\infty + T^\alpha|z|_\alpha,$$

with suprema taken over $[0, T]$, for which we obtain the bound

$$|z-\bar z|_{\alpha,T} \lesssim T^\beta|z-\bar z|_{\alpha,T}|b|_\beta + |Z-\bar Z|_{\alpha,T},$$

thus yielding

$$|z-\bar z|_{\alpha,T} \le 2|Z-\bar Z|_{\alpha,T},$$

on $[0, T]$ where T is as in (2.2). Iterating this bound, we conclude that on any sub-interval $[s, s+T]$ of $[0, 1]$, one has a bound of the type

$$\big|(z-\bar z)\restriction[s, s+T]\big|_{\alpha,T} \le 2\exp\big(C(1+s/T)\big)|Z-\bar Z|_{\alpha,T},$$

whence we conclude that on $[0, 1]$, for a possibly larger constant C, one has

$$|z-\bar z|_\alpha \lesssim \exp\big(C\big(1+T^{-1}\big)\big)|Z-\bar Z|_\alpha.$$

Since (2.2) allows us to choose T such that $1/T \lesssim |b|_\beta^{1/\beta} + |Z|_\alpha^{1/\alpha} + |\bar Z|_\alpha^{1/\alpha}$, the claim follows. □

3. Averaging without feedback. In this section, we provide an interpretation of the integral against fractional Brownian motion that is more stable than the Young integral in situations in which the integrand exhibits fast oscillations. The idea is to exploit the adaptedness of the integrand in a way that allows us to apply the stochastic version of the sewing lemma [10] recently obtained in [23].

To take one step back, we recall that integration of a deterministic function with respect to a fractional Brownian motion (fBm) of Hurst parameter $H > \frac{1}{2}$ is called a Wiener integral (with respect to Gaussian processes): the integrands are smooth stochastic processes completed with the norm given by the inner product

$$\langle \phi, \psi \rangle = \mathbf{E}\Big(\int_{\mathbf{R}} \phi_s \, dB_s \int_{\mathbf{R}} \psi_s \, dB_s\Big).$$

(Limits of smooth functions with respect to this norm can be Schwartz distributions.) When the integrand is sufficiently smooth, this is just the Young integral.

Let B_t be a m-dimensional fractional Brownian motion with Hurst parameter $H \in (\frac{1}{2}, 1)$, it has an integral representation with respect to a two sided standard Wiener process W_t, which was introduced by Mandelbrot and Van Ness [29]. We consider H as being fixed throughout this article and, therefore, omit the superscript. For $r > u$, write the increment of fractional Brownian motion as a sum of two processes:

$$\begin{aligned} B_r - B_u &= \int_{-\infty}^{u} \big((r-v)^{H-\frac{1}{2}} - (u-v)^{H-\frac{1}{2}}\big) \, dW_v + \int_u^r (r-v)^{H-\frac{1}{2}} \, dW_v \\ &\overset{\text{def}}{=} \bar{B}_r^u + \tilde{B}_r^u. \end{aligned} \tag{3.1}$$

Writing $\mathcal{G}_t$ for the filtration generated by the increments of W, $\bar{B}_t^u$ is $\mathcal{G}_u$-measurable and smooth in t on (u, ∞), while $\tilde{B}_t^u$ is independent of $\mathcal{G}_u$. For the special case $u = 0$, we simply write $\bar{B}_t = \bar{B}_t^0$, $\tilde{B}_t = \tilde{B}_t^0$. Recall also that the filtration $\mathcal{G}_t$ coincides with that generated by the increments of B.

3.1. *Mixed Riemann and Wiener integrals.* If $f : \mathbf{R} \times \mathbf{R}^d \to L(\mathbf{R}^m, \mathbf{R}^d)$ is a measurable function and x_t a $\mathcal{G}_t$-adapted stochastic process, our first task is to define $\int_0^t f(r, x_r) \, dB_r$ as the limit of 'Riemann sums' of the type $\sum_i \int_{s_i}^{s_{i+1}} f(r, x_{s_i}) \, dB_r$, provided that f and x satisfy suitable assumptions. Prior to justifying its convergence, we explain how each individual integration in the sum is defined. For any $s < t$, set

$$A_{s,t} \overset{\text{def}}{=} \int_s^t f(r, x_s) \, dB_r \overset{\text{def}}{=} \int_s^t f(r, x_s) \, d\bar{B}_r^s + \int_s^t f(r, x_s) \, d\tilde{B}_r^s.$$

The first integral will be considered as a Riemann–Stieltjes integral which will exploit the fact that $\bar{B}^s$ is a smooth function with a well-behaved singularity at time s. The second term will be interpreted as a Wiener integral with respect to the Gaussian process $\tilde{B}^s$, which we can do since x_s is $\mathcal{G}_s$-measurable and, therefore, independent of it. Since $r \mapsto \bar{B}_r^u$ is smooth for $r > u$ and its derivative has an integrable singularity at $r \sim u$, the Riemann integral $\int_u^t f(r, x_u) \, d\bar{B}_r^u$ can be defined in a pathwise sense as soon as f is continuous in both of its arguments. If $x_\cdot$ has continuous sample paths, then the same is true for the Wiener integral since the map $F \mapsto \int_u^t F_r \, d\tilde{B}_r^u$, viewed as a linear map from $\mathcal{C}^\infty$ into $L^2(\Omega)$, can be extended to all $F \in \mathcal{C}^0$ (and actually even to $F \in \mathcal{C}^{-\kappa}$ for κ small enough; see Lemmas 3.2 and 3.3 below). Think now of u as being fixed and consider an arbitrary stochastic process F on $[u, t]$, but we think of the case $F_r = f(r, x_u)$.

REMARK 3.1. If F is either deterministic and Hölder continuous of order α or $F \in \mathcal{B}_{\alpha,p}$ where $p > 2$ and $\alpha + H > 1$, then the mixed integral coincides with the Young integral.

The first follows from the deterministic sewing lemma and that $\int_u^t (F_r - F_u)\, d\tilde{B}_r \lesssim |t-u|^{\alpha+H}$. The second follows from the stochastic sewing lemma; alternatively, this is a special case of Lemma 3.12 below.

In situations where f and x are sufficiently regular so that the usual Riemann sums converge, we will see in Lemmas 3.12 and 4.10 that the notion of integration used here coincides with the classical Young integral. The advantage of this set-up however is that we can exploit the stochastic cancellations of the Wiener integral through the use of the stochastic sewing lemma, which allows us to substantially expand the class of admissible integrands and is fundamental for extracting uniform estimates for SDEs with random inputs.

We begin with building up estimates for the mixed stochastic integral explained earlier. Let R denote the covariance function of $\tilde{B}$. We work componentwise, so that instead of complicating our notation with i.i.d. copies of the one-dimensional fBm's, we may assume that $\tilde{B}$ is one dimensional in the formulation below. It follows from the scaling properties of $\tilde{B}$ that

$$R(r,s) = \mathbf{E}\tilde{B}_r\tilde{B}_s = (r\wedge s)^{2H}\hat{R}\Big(\frac{|r-s|}{r\wedge s}\Big),$$
$$\hat{R}(t) = \mathbf{E}\tilde{B}_1\tilde{B}_{1+t} = \int_0^1 (1-s)^{H-\frac{1}{2}}(1+t-s)^{H-\frac{1}{2}}\, ds, \tag{3.2}$$

so that their distributional derivatives $\partial^2_{r,s}R(r,s) \stackrel{\text{def}}{=} \frac{\partial^2}{\partial r\,\partial s}R(r,s)$ satisfy

$$\partial^2_{r,s}R(r,s) = (r\wedge s)^{2H-2}G\Big(\frac{|r-s|}{r\wedge s}\Big),$$
$$G(t) = (2H-1)\hat{R}'(t) - (t+1)\hat{R}''(t). \tag{3.3}$$

Convention. We now fix a filtration $\mathcal{F}_s$ with $\mathcal{G}_s \subset \mathcal{F}_s$ and such that, for every s, $\tilde{B}^s$ is independent of $\mathcal{F}_s$. The example to have in mind which will be relevant in Section 4 is to take $\mathcal{F}_s = \mathcal{G}_s \vee \hat{\mathcal{G}}_s$, where $\hat{\mathcal{G}}$ is the filtration generated by the increments of a Wiener process independent of B.

Recall that Wiener integrals are centred Gaussian processes. In our case, F_s is random but with $F_s \in \mathcal{F}_u$ for any $s \in [u,t]$, so that $\int_u^t F_s\, d\tilde{B}^u_s$ is a centred Gaussian process, conditional on $\mathcal{F}_u$.

LEMMA 3.2. *$G(t) \approx t^{2H-2}$ for $t \ll 1$ and $G(t) \approx t^{H-\frac{3}{2}}$ for $t \gg 1$. In particular, $\partial^2_{r,s}R(r,s)$ is integrable over any bounded region and there exists $c_1 \in \mathbf{R}$ s.t.,*

$$\int_0^t\int_0^t \big|\partial^2_{r,s}R(r,s)\big|\, dr\, ds \le c_1 t^{2H},$$

for every $t \in \mathbf{R}_+$. For some fixed $u \ge 0$, let F_s be pathwise smooth in s and suppose F_s is $\mathcal{F}_u$-measurable for any $s \in [u,t]$, then the following Itô isometry holds:

$$\mathbf{E}\Big(\Big(\int_u^t F_s\, d\tilde{B}^u_s\Big)^2\Big|\mathcal{F}_u\Big)(\omega) = \int_u^t\int_u^t \partial^2_{r,s}R(r,s)F_r(\omega)F_s(\omega)\, dr\, ds. \tag{3.4}$$

PROOF. The bound on $\partial^2_{r,s}R(r,s)$ follows at once from the representation given in Lemma A.1 below, while the fact that (3.4) holds in the distributional sense is classical; see, for example, [16]. The bounds on R given in Lemma A.1 guarantee that the distributional derivative of R coincides with its weak derivative and is an integrable function, so that (3.4) also holds with the right-hand side interpreted as a Lebesgue integral. □

As usual [3, 16], the left-hand side of (3.4) can be defined in such a way that this isometry extends to all g taking values in the completion of the space of smooth functions under the norm given by the right-hand side of (3.4). This in particular contains all $g \in L^2$, as can be shown similar to [30]. It turns out however that the space of admissible integrands for this Wiener integral contains not only functions, but also distributions of order $-\kappa$ provided that $\kappa < H - \frac{1}{2}$. More precisely, we have the following result, a proof of which is postponed to the Appendix.

LEMMA 3.3. *Let h be a continuous function. Then, for all $T > 0$ and $\kappa \in [0, H - \frac{1}{2})$, one has the bound*

$$|h|_{\mathrm{RKHS}}^2 \stackrel{\mathrm{def}}{=} \left|\int_0^T \int_0^T \partial_{r,s}^2 R(r,s) h(r) h(s)\, dr\, ds\right| \lesssim T^{2H-2\kappa} |h|_{-\kappa}^2, \tag{3.5}$$

where the negative Hölder norm $|h|_{-\kappa}$ on $[0, T]$ is given by

$$|h|_{-\kappa} = \sup_{0 \le s,t \le T} |t-s|^{\kappa-1} \left|\int_s^t h(r)\, dr\right|.$$

Since $\bar{B}_t$ is smooth in t, integrals with respect to it extend to rougher integrands, as we will show now. Below we provide a bound for integration with respect to the full fBm.

LEMMA 3.4. *Let B be a fBm with $H > \frac{1}{2}$ and fix $0 \le \kappa < H - \frac{1}{2}$. Let $s \ge 0$ be fixed. Let $r \mapsto F_r$ be smooth, with each F_r for $s \le r$ measurable with respect to $\mathcal{F}_s$. Then, for $t \ge s$ with $|t - s| \le 1$ and $2 \le p < q$ one has the bound*

$$\left\|\int_s^t F_r\, dB_r\right\|_p \lesssim \big\||F|_{-\kappa}\big\|_q |t-s|^{H-\kappa},$$

where $|F|_{-\kappa}$ denotes its negative Hölder norm on $[s, t]$.

By linearity and density, this immediately allows us to extend the notion of integral against B to any integrand in $L^q((\Omega, \mathcal{F}_r), \mathcal{C}^{-\kappa})$ for any $0 \le \kappa < H - \frac{1}{2}$ (which may no longer agree with the Young integral).

PROOF. Since our set-up is translation invariant, we restrict ourselves to the case $s = 0$ without loss of generality and we write

$$\int_0^t F_r\, dB_r = \int_0^t F_r\, d\bar{B}_r + \int_0^t F_r\, d\tilde{B}_r \stackrel{\mathrm{def}}{=} I_1 + I_2.$$

To bound I_1, we note that $r \mapsto \bar{B}_r(\omega)$ is a smooth function on $(0, \infty)$ satisfying the bounds, for any $p \ge 1$,

$$\|\dot{\bar{B}}_r\|_p \lesssim r^{H-1}, \qquad \|\ddot{\bar{B}}_r\|_p \lesssim r^{H-2}. \tag{3.6}$$

We then integrate I_1 by parts, so that

$$I_1 = \dot{\bar{B}}_t \int_0^t F(r)\, dr - \int_0^t \int_0^r F_u\, du\, \ddot{\bar{B}}_r\, dr,$$

and the required bound follows from Hölder's inequality.

Concerning I_2, since the integrand is $\mathcal{F}_0$-measurable and $\tilde{B}$ is independent of $\mathcal{F}_0$, the Wiener integral I_2 is Gaussian and its L_p norm is bounded by its L_2 norm. We can proceed as if the integrand were deterministic and use Lemma 3.2, so that one has the bound

$$\mathbf{E}|I_2|^p \lesssim \mathbf{E}\left|\int_0^t \int_0^t F(r) F(s) \partial_{r,s}^2 R(r,s)\, dr\, ds\right|^{p/2}. \tag{3.7}$$

Inserting the bound from Lemma 3.3 into (3.7), we obtain the bound $\|I_2\|_p \lesssim \||F|_{-\kappa}\|_p t^{H-\kappa}$ as required. □

3.2. *Stochastic sewing lemma.* Let $A_{s,t}$ denote a two parameter stochastic process with values in $\mathbf{R}^n$, where $s \le t$. Both s and t take values in a fixed finite interval $[a,b]$. We are interested in situations where A is close to being an increment. To quantify this, for any $s < u < t$, set

$$\delta A_{\mathrm{sut}} \stackrel{\mathrm{def}}{=} A_{s,t} - A_{s,u} - A_{u,t},$$

which vanishes if and only if $A_{s,t}$ is the increment of a one-parameter function. In the cases of interest to us, the family of 'defects' δA_{sut} is typically much smaller than $A_{s,t}$ itself for $|t-s|$ small.

Let us now quantify this more precisely. Given $p \ge 2$ and an exponent $\eta > 0$, we define the space H_η^p of continuous functions $(s,t) \mapsto A_{s,t} \in L^p(\Omega, \mathcal{F}_t)$ such that

$$\|A\|_{\eta,p} \stackrel{\mathrm{def}}{=} \sup_{s<t} \frac{\|A_{s,t}\|_p}{|t-s|^\eta} < \infty, \tag{3.8}$$

where $\|\cdot\|_p$ denotes the norm in $L^p(\Omega)$. We also define the space $\bar{H}_\eta^p$ of maps $A_{s,t}$ as above such that

$$|||A|||_{\eta,p} \stackrel{\mathrm{def}}{=} \sup_{s<u<t} \frac{\|\mathbf{E}(\delta A_{\mathrm{sut}}|\mathcal{F}_s)\|_p}{|t-s|^\eta} < \infty. \tag{3.9}$$

Then H_η^p is a Banach space with norm $\|\cdot\|_{\eta,p}$, while $|||\cdot|||_{\eta,p}$ is only a semi-norm.

We will view a partition of the interval $[a,b]$ as a collection $\mathcal{P}$ of nonempty closed intervals that cover $[a,b]$ and overlap pairwise in at most one point, so we can use the notation $\sum_{[u,v]\in\mathcal{P}} A_{u,v}$ for the 'Riemann sum' associated with A on the partition $\mathcal{P}$. Given such a partition, we write $|\mathcal{P}|$ for the length of the largest interval contained in $\mathcal{P}$. The following result was proved in [23], Theorem 2.1. The version presented here is slightly weaker than the general result, but it will be sufficient for our needs. Note that a deterministic version of the sewing lemma was given in [10] and was instrumental for the reformulation of rough path theory [28] as exposed, for example, in [8]. A multidimensional analogue to the sewing lemma is given by the reconstruction theorem from the theory of regularity structures [13], Theorem 3.23.

LEMMA 3.5 (Stochastic sewing Lemma). *Suppose that, for some $p \ge 2$, one has $A \in H_\eta^p \cap \bar{H}_{\bar\eta}^p$ with $\eta > \frac{1}{2}$ and $\bar\eta > 1$. Then, for every $t > 0$, the limit in L^p*

$$I_{s,t}(A) \stackrel{\mathrm{def}}{=} \lim_{|\mathcal{P}|\to 0} \sum_{[u,v]\in\mathcal{P}} A_{u,v}, \tag{3.10}$$

with $\mathcal{P}$ taking values in partitions of $[s,t]$, exists and there exists a constant C depending only on p and $\eta, \bar\eta$ such that

$$\big\|I_{s,t}(A)\big\|_p \le C\big(|||A|||_{\bar\eta,p}|t-s|^{\bar\eta} + \|A\|_{\eta,p}|t-s|^\eta\big),$$

$$\big\|\mathbf{E}\big(I_{s,t}(A) - A_{s,t}|\mathcal{F}_s\big)\big\|_p \le C|||A|||_{\bar\eta,p}|t-s|^{\bar\eta}.$$

Furthermore, $I(A)$ satisfies the identity $I_{s,u}(A) + I_{u,t}(A) = I_{s,t}(A)$ for any $s \le u \le t$, so that there exists a stochastic process $I_t(A) = I_{0,t}(A)$ with $I_{s,t}(A) = I_t(A) - I_s(A)$. If one furthermore has the bound $\|\mathbf{E}(A_{s,t}|\mathcal{F}_s)\|_p \lesssim |t-s|^{\bar\eta}$, then $I(A) \equiv 0$.

REMARK 3.6. In the general case, the bound (3.8) is required for δA only, but we will always have this stronger bound at our disposal. Note also that [23], Theorem 2.1, requires joint continuity of $\mathbf{E}(A_{s,t}|\mathcal{F}_s)$, but this is only ever used to obtain [23], equation (2.8), which we do not need.

REMARK 3.7. A simple special case is the classical result by Young [35]: for $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^\beta$ with $\alpha + \beta > 1$, one has the bound

$$\left|\int_s^t f_r\, dg_r - f_s(g_t - g_s)\right| \lesssim |f|_\alpha |g|_\beta |t-s|^{\alpha+\beta}. \tag{3.11}$$

Strictly speaking, setting $A_{s,t} = f_s(g_t - g_s)$, this is only a special case of Lemma 3.5 for $\beta > \frac{1}{2}$, but this is only due to the fact that, as already mentioned, the formulation given here is slightly weaker than the one given in [23].

3.3. *A stochastic integral with respect to fBm.* For $\kappa, \gamma \in [0, 1]$, we introduce a space $\mathcal{C}^{-\kappa,\gamma}$ of distributions of order $-\kappa$ (in time) with values in the space of Hölder continuous functions of order γ (in space). More precisely, an element $f \in \mathcal{C}^{-\kappa,\gamma}$ is interpreted as the distributional derivative with respect to the first argument of a continuous function $\hat{f} : \mathbf{R} \times \mathbf{R}^d \to \mathbf{R}$, such that $\hat{f}(0, x) = 0$ and

$$\begin{aligned} \big|\hat{f}(t,x) - \hat{f}(s,x)\big| &\le K|t-s|^{1-\kappa}, \\ \big|\hat{f}(t,x) - \hat{f}(s,x) - \hat{f}(t,y) + \hat{f}(s,y)\big| &\le K|t-s|^{1-\kappa}|x-y|^\gamma, \end{aligned} \tag{3.12}$$

uniformly over $|s - t| \le 1$ and $x, y \in \mathbf{R}^d$. Alternatively, one has

$$\sup_x \big|f(\cdot, x)\big|_{-\kappa} \le K, \qquad \sup_{x\neq y} \frac{|f(\cdot,x) - f(\cdot,y)|_{-\kappa}}{|x-y|^\gamma} \le K.$$

We write $|f|_{-\kappa,\gamma}$ for the smallest possible choice of proportionality constant K in (3.12). In particular if f is bounded and $f(r, \cdot)$ uniformly γ-Hölder continuous (uniformly in r), then $f \in \mathcal{C}^{-\kappa,\gamma}$ for every $\kappa > 0$.

The following lemma is only used for the proof of Theorem A.

LEMMA 3.8. *For $\alpha, \kappa \in (0, 1)$, the map*

$$(f, x) \mapsto \big(t \mapsto f(t, x_t)\big),$$

extends to a continuous map from $\mathcal{C}^{-\kappa,\gamma} \times \mathcal{C}^\alpha$ into $\mathcal{C}^{-\kappa}$ provided that $\gamma\alpha > \kappa$. Furthermore, one has the bound

$$\big|t \mapsto f(t, x_t)\big|_{-\kappa} \lesssim |f|_{-\kappa,\gamma}\big(1 + |x|_\alpha^\gamma T^{\gamma\alpha}\big), \tag{3.13}$$

on any interval of length T.

PROOF. This is an immediate consequence of the deterministic sewing lemma [10]: Let $\Xi_{s,t}$ be a deterministic two parameter process with

$$|\Xi_{\mathrm{st}}| \le \hat{K}|t-s|^\eta, \qquad |\delta\Xi_{\mathrm{sut}}| \le \hat{K}_{\mathrm{Lip}}|t-s|^{\bar{\eta}},$$

for some $\eta > 0$ and $\bar{\eta} > 1$. Then, for every $s < t \le T$, the limit $I_{s,t}(\Xi) \stackrel{\text{def}}{=} \lim_{|\mathcal{P}|\to 0} \sum_{[u,v]\in\mathcal{P}} \Xi_{u,v}$ exists along partitions of $[s, t]$ and one has

$$\big|I_{s,t}(\Xi) - \Xi_{s,t}\big| \lesssim \hat{K}_{\mathrm{Lip}}|t-s|^{\bar{\eta}}. \tag{3.14}$$

To make sense of the distribution $r \mapsto f(r, x_r)$, we need to be able to make sense of its integral over any interval $[s, t]$. A good candidate for this is $I_t(\Xi)$, where

$$\Xi_{s,t} = \int_s^t f(r, x_s)\, dr = \hat{f}(t, x_s) - \hat{f}(s, x_s).$$

Writing $K = |f|_{-\kappa,\gamma}$, the bounds (3.12) imply that $|\Xi_{s,t}| \lesssim K|t-s|^{1-\kappa}$. On the other hand, we have

$$|\delta\Xi_{\mathrm{sut}}| = \big|\hat f(t,x_s) - \hat f(u,x_s) - \hat f(t,x_u) + \hat f(u,x_u)\big| \\ \lesssim |f|_{-\kappa,\gamma}|t-u|^{1-\kappa}|s-u|^{\gamma\alpha}|x|_\alpha^\gamma, \tag{3.15}$$

so that the corresponding 'integral' $I(\Xi)$ is well defined since $\gamma\alpha > \kappa$ by assumption. Furthermore, it follows from (3.14) that

$$\big|I_{s,t}(\Xi)\big| \lesssim |t-s|^{1-\kappa}|f|_{-\kappa,\gamma}\big(1 + |x|_\alpha^\gamma|t-s|^{\gamma\alpha}\big),$$

which does indeed show that the distributional derivative of $t \mapsto I_{0,t}(\Xi)$ belongs to $\mathcal{C}^{-\kappa}$. In the particular case where f is actually a β-Hölder continuous function in its first argument, we have

$$\big|\Xi_{s,t} - f(s,x_s)(t-s)\big| \lesssim |t-s|^{1+\beta},$$

so that we do have $I_{s,t}(\Xi) = \int_s^t f(r,x_r)\,dr$ and, therefore, $\frac{d}{dt}I_{0,t}(\Xi) = f(t,x_t)$ as required. □

We now introduce a space of stochastic processes that will be a natural candidate for containing our solutions. Recall that $\mathcal{F}_t$ denotes a filtration of the underlying probability space as in Section 3.1, namely it contains $\mathcal{G}_s = \sigma(\{B_u - B_v : u, v \le s\})$ and is such that $\tilde B^s_\cdot$ is independent of $\mathcal{F}_s$ for every s.

DEFINITION 3.9. For $\alpha > 0$ and $p \ge 1$, let $\mathcal{B}_{\alpha,p}$ denote the Banach space consisting of all $\mathcal{F}_t$-adapted processes x_t such that $\delta x \in H_\alpha^p$, where $\delta x_{s,t} = x_t - x_s$. We also write

$$\|x\|_{\alpha,p} = \sup_{s,t}|t-s|^{-\alpha}\|x_t - x_s\|_{L^p}. \tag{3.16}$$

(To be consistent with (3.8), we should really write $\|\delta x\|_{\alpha,p}$, but we drop the δ for the sake of conciseness.)

Our aim is to lay the foundations for a solution theory of SDEs driven by fractional Brownian motion with right-hand sides determined by functions $f : \mathbf{R}\times\mathbf{R}^d \to L(\mathbf{R}^m,\mathbf{R}^d)$ such that the following holds.

LEMMA 3.10. *Let $p \ge 2$ and $\alpha > \frac12$. Assume that $x_\cdot \in \mathcal{B}_{\alpha,p}$, let $f \in \mathcal{C}^{-\kappa,\gamma}$ (deterministic) for some $\kappa, \gamma \ge 0$ such that $\eta = H - \kappa > \frac12$ and $\bar\eta = H - \kappa + \gamma\alpha > 1$, and define the two parameter stochastic process*

$$A_{s,t} = \int_s^t f(r,x_s)\,dB_r,$$

where the integral is interpreted as a conditional Wiener integral as constructed in Lemma 3.4. Then one has $A \in H_\eta^p \cap \bar H_{\bar\eta}^p$ and we take the resulting process as our definition of the stochastic integral against B:

$$\int_s^t f(r,x_r)\,dB_r \stackrel{\mathrm{def}}{=} I_{s,t}(A). \tag{3.17}$$

This integral satisfies the bounds

$$\left\|\int_s^t f(r,x_r)\,dB_r\right\|_p \\ \lesssim |f|_{-\kappa,\gamma}\big(|t-s|^{H-\kappa} + \|x\|_{\alpha,p}^\gamma|t-s|^{\bar\eta}\big), \tag{3.18}$$

$$\left\| \mathbf{E}\left(\int_s^t (f(r, x_r) - f(r, x_s))\, dB_r \Big| \mathcal{F}_s \right) \right\|_p \tag{3.19}$$
$$\lesssim |f|_{-\kappa,\gamma} \|x\|_{\alpha,p}^{\gamma} |t - s|^{\bar{\eta}}$$

uniformly over $s, t \in [0, T]$.

PROOF. By Lemma 3.5, it suffices to show that for κ, γ as in the assumption, one has $A \in H_\eta^p \cap \bar{H}_{\bar{\eta}}^p$ with

$$\eta = H - \kappa, \qquad \|A\|_{\eta,p} \lesssim |f|_{-\kappa,\gamma}. \tag{3.20a}$$

$$\bar{\eta} = H - \kappa + \gamma\alpha, \qquad \|\!|A|\!\|_{\bar{\eta},p} \lesssim |f|_{-\kappa,\gamma} \|x\|_{\alpha,p}^{\gamma}. \tag{3.20b}$$

Since $f(\cdot, x_s)$ is $\mathcal{F}_s$-measurable and $f \in \mathcal{C}^{-\kappa,\gamma}$ with $\kappa \in [0, H - \frac{1}{2}]$, it follows from Lemma 3.4 that one has the bound

$$\|A_{\mathrm{st}}\|_p \lesssim \sup_{x \in \mathbf{R}^d} \big| f(\cdot, x) \big|_{-\kappa} |t - s|^{H-\kappa}, \tag{3.21}$$

where C_p is a universal constant, thus yielding the bound (3.20a).

We now bound δA_{sut} for u between s and t. Since

$$\delta A_{\mathrm{sut}} = \int_u^t \big(f(r, x_s) - f(r, x_u) \big)\, dB_r,$$

and since $s < u$, we are again in the setting of Lemma 3.4, which yields

$$\|\delta A_{\mathrm{sut}}\|_p \lesssim \big\| \big| f(\cdot, x_s) - f(\cdot, x_u) \big|_{-\kappa} \big\|_q |t - u|^{H-\kappa}. \tag{3.22}$$

We then note that

$$\big| f(\cdot, x_s) - f(\cdot, x_u) \big|_{-\kappa} \le |f|_{-\kappa,\gamma} |x_s - x_u|^{\gamma}.$$

Choosing $q = p/\gamma$ in (3.22), we thus obtain

$$\|\delta A_{\mathrm{sut}}\|_p \lesssim |f|_{-\kappa,\gamma} \|x\|_{\alpha,p}^{\gamma} |u - s|^{\alpha\gamma} |t - u|^{H-\kappa},$$

so that (3.20b) follows. Since $H - \kappa > \frac{1}{2}$ and $H + \alpha\gamma - \kappa > 1$, we can now apply Lemma 3.5 and immediately deduce $I_{s,t}(A) = \lim_{|\mathcal{P} \to 0} \sum_{[u,v] \in \mathcal{P}} A_{u,v}$ and the required bounds (3.18)–(3.19). □

REMARK 3.11. Since $f(t, x)$ is not assumed to be Hölder continuous in t, the integral defined in the theorem by sewing up the mixed integrals (Riemann–Stieltjes integral with respect to the smooth $\bar{B}_t$ and the Wiener integral with respect to $\tilde{B}_t$ with essentially 'nonrandom' integrand) cannot necessarily be interpreted as a Young integral.

Finally, we note that in 'nice' situations where the integral against B also makes sense as a Young integral, the two integrals coincide. The precise statement is as follows.

LEMMA 3.12. *Under the assumptions of Lemma* 3.10 *and assuming that* f *is such that, for some* δ *with* $\delta + H > 1$, *one has*

$$\sup_x \sup_{|t-s| \le 1} |t - s|^{-\delta} \big| f(t, x) - f(s, x) \big| < \infty,$$

the integral given by (3.17) *coincides with the usual Young integral*.

PROOF. By Lemma 3.4 with $\kappa = 0$, and $q > p \geq 2$,

$$\left\| \int_u^v \big(f(r, x_u) - f(u, x_u)\big)\, dB_r \right\|_p \leq \left\| \sup_{r\in[u,v]} \big|f(r, x_u) - f(u, x_u)\big| \right\|_q |v-u|^H \lesssim |v-u|^{H+\delta}.$$

The final part of Lemma 3.5 then leads to the desired conclusion, namely that

$$\lim_{|\mathcal{P}|\to 0} \sum_{[u,v]\in\mathcal{P}} \left(\int_u^v f(r, x_u)\, dB_r - f(u, x_u)(B_v - B_u) \right) = 0,$$

in probability. □

3.4. *A semideterministic averaging result.* In order to state the main theorem of this section, which is then going to lead us to the proof of Theorem A, we introduce the space $\mathcal{C}^{\alpha,2}$ of functions that are α-Hölder continuous in time, with values into the space BC^2. With this notation, we then have the following.

THEOREM 3.13. *For $H > \frac{1}{2}$, let $\alpha, \kappa, \gamma > 0$ satisfy the assumptions of Lemma 3.10 and $\alpha < H - \kappa$. Let furthermore $\zeta \in (\alpha, 1]$ and let $f_n, \bar{f} : \mathbf{R}_+ \times \mathbf{R}^d \to L(\mathbf{R}^m, \mathbf{R}^d)$ be in $\mathcal{C}^{\zeta,2}$ such that*

$$\lim_{n\to\infty} |f_n - \bar{f}|_{-\kappa,\gamma} = 0.$$

Let x^n and x be the $\mathcal{C}^\alpha$ solutions to the equations

$$dx_t^n = f_n(t, x_t^n)\, dB_t, \qquad dx_t = \bar{f}(x_t)\, dB_t, \tag{3.23}$$

with $x_0^n = x_0$ and the integrals interpreted pathwise in Young's sense. Then $x^n \to x$ in probability in $\mathcal{C}^\alpha$. The same holds if the equations include a term with dB replaced by dt.

PROOF. The fact that (3.23) admits unique solutions in $\mathcal{C}^\alpha$ for every realisation of $B \in \mathcal{C}^\beta$ with $\beta > \alpha$ and $\alpha + \beta > 1$ is standard. (Combine Lemma 2.1 with [35] to show that the Picard iteration is contracting in $\mathcal{C}^\alpha$ with fixed initial value.)

Let us first obtain bounds on x^n that are uniform in n. For any $\frac{1}{2} < \alpha < H - \kappa$ satisfying $\alpha\gamma > 1 - H + \kappa$, we can apply bound (3.18) of Lemma 3.10 so that, over any interval $[0, T]$, we obtain the bound

$$\|x^n\|_{\alpha,p} \lesssim T^{H-\alpha-\kappa}\big(1 + T^{\gamma\alpha} \|x^n\|_{\alpha,p}^\gamma\big),$$

which immediately implies that $\|x^n\|_{\alpha,p} \leq 1$, uniformly over n, provided that we choose a sufficiently short time interval. This bound can be iterated and, therefore, yields an order one a priori bound on $\|x^n\|_{\alpha,p}$ over any fixed time interval.

We then note that we can write

$$x_t = Z_t + \int_0^t \bar{f}(x_s)\, dB_s, \qquad x_t^n = Z_t^n + \int_0^t \bar{f}(x_s^n)\, dB_s,$$

with

$$Z_t = x_0, \qquad Z_t^n = x_0 + \int_0^t \big(f_n(s, x_s^n) - \bar{f}(x_s^n)\big)\, dB_s.$$

It now follows again from (3.18) in Lemma 3.10 that, over any fixed time interval, one has the bound

$$\|Z^n - Z\|_{H-\kappa,p} \lesssim |f_n - \bar{f}|_{-\kappa,\gamma}\big(1 + \|x^n\|_{\alpha,p}^\gamma\big).$$

Note now that by Kolmogorov's continuity theorem, we have for any $\delta, \zeta > 0$ the inclusions

$$L^p(\Omega, \mathcal{C}^\zeta) \subset \mathcal{B}_{\zeta,p} \subset L^p(\Omega, \mathcal{C}^{\zeta-\frac{1}{p}-\delta}) \tag{3.24}$$

so that, choosing p large enough, we conclude that $|Z^n - Z|_\alpha \to 0$ in L_p, for any $p \geq 2$, as $n \to \infty$. The claim now follows from Lemma 2.2. □

The type of application of this theorem that we have in mind is that when f_n is, for example, given by

$$f_n(t, x) = F(x, y_{nt}), \tag{3.25}$$

for some smooth function F and for a stationary stochastic process y_t that is independent of the driving noise B. (This is so that f_n can be considered deterministic.) To give a more concrete setting, given any two random variables X and Y, we can measure their degree of independence $\alpha(X, Y)$ (also called the 'strong mixing coefficient') by

$$\alpha(X, Y) = \sup\{\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B) : A \in \sigma(X), B \in \sigma(Y)\}.$$

Note that if F and G are two bounded centred functions, then

$$\left|\mathbf{E}F(X)G(Y)\right| \leq 4\alpha(X, Y)|F|_\infty|G|_\infty, \tag{3.26}$$

see [17]. The following proposition is then crucial.

LEMMA 3.14. *Let $\mathcal{Y}$ be a Polish space and let $(y_t)_{t\in\mathbf{R}}$ be a stationary $\mathcal{Y}$-valued stochastic process such that $\alpha(y_0, y_t) \lesssim t^{-\delta}$ for some $\delta > 0$. Let $F : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}$ be a measurable function; $\mathcal{C}^2$ in the first variable, such that*

$$\left|F(x, y)\right| \leq K, \qquad \left|F(x, y) - F(z, y)\right| \leq K|x - z|,$$

uniformly over $y \in \mathcal{Y}$ and $x, z \in \mathbf{R}^d$. Assume for simplicity that outside of a compact set F is periodic in its first argument.

Then, for every $\kappa > 0$, every $\gamma < 1$ and every $p \geq 1$, the sequence f_n defined as in (3.25) is such that $|f_n - \bar{f}|_{-\kappa,\gamma} \to 0$ in L_p as $n \to \infty$, with $\bar{f}(x) = \int F(x, y)\mu(dy)$, where μ denotes the law of y_t for any fixed t.

PROOF. Since F is bounded measurable and $f_n(t, x) = F(x, y_{nt}) \in \mathcal{C}^{0,1}$, we note that $\bar{f}$ is bounded Lipschitz continuous. Replacing f_n by $(f_n - \bar{f})/K$, we can assume without loss of generality that $K = 1$ and $\bar{f} = 0$. Making use of (3.26), we then have the bound

$$\begin{aligned}
\mathbf{E}\left(\int_s^t f_n(r, x)\, dr\right)^2 &= n^{-2}\mathbf{E}\int_{ns}^{nt}\int_{ns}^{nt} F(x, y_r)F(x, y_{\bar{r}})\, dr\, d\bar{r} \\
&\leq |F(x, \cdot)|_\infty^2 4n^{-2}\int_{ns}^{nt}\int_{ns}^{nt} \alpha(y_r, y_{\bar{r}})\, dr\, d\bar{r} \\
&\lesssim |F(x, \cdot)|_\infty^2 n^{-2}\int_{ns}^{nt}\int_{ns}^{nt} |r - \bar{r}|^{-\delta}\, dr\, d\bar{r} \\
&\lesssim |F(x, \cdot)|_\infty^2 n^{-2}|nt - ns|^{2-\delta} \lesssim n^{-\delta}|t - s|^{2-\delta}.
\end{aligned}$$

On the other hand, we have the trivial bound $|\int_s^t f_n(r, x)\, dr| \leq |t - s|$, so that for any $p \geq 2$,

$$\left\|\int_s^t f_n(r, x)\, dr\right\|_p \lesssim n^{-\frac{\delta}{p}}|t - s|^{1-\frac{\delta}{p}}.$$

Replacing $f_n(r,x)$ by $f_n(r,x) - f_n(r,z)$, we similarly obtain

$$\left\| \int_s^t \big(f_n(r,x) - f_n(r,z)\big)\,dr \right\|_p \lesssim n^{-\frac{\delta}{p}} |x-z| |t-s|^{1-\frac{\delta}{p}}.$$

Applying Kolmogorov's continuity criterion, we obtain over any finite time interval the bound

$$\left\| \left| \int_0^\cdot \big(f_n(r,x) - f_n(r,z)\big)\,dr \right|_{1-\kappa} \right\|_p \lesssim n^{-\frac{\delta}{p}} |x-z|,$$

provided that we choose p large enough so that $\frac{\delta}{p} + \frac{1}{p} < \kappa$. In other words, we have the bound

$$\big\| |f_n(\cdot,x) - f_n(\cdot,z)|_{-\kappa} \big\|_p \lesssim n^{-\frac{\delta}{p}} |x-z|.$$

This allows us to apply Kolmogorov's criterion a second time; this time in the spatial variable, showing that for any compact set $\mathcal{K}$, we have

$$\left\| \sup_{\substack{x \neq z \\ x,z \in \mathcal{K}}} \frac{|f_n(\cdot,x) - f_n(\cdot,z)|_{-\kappa}}{|x-z|^\gamma} \right\|_p \lesssim n^{-\frac{\delta}{p}},$$

provided that p is such that $\frac{d}{p} < 1-\gamma$, where the proportionality constant in front of $n^{-\frac{\delta}{p}}$ depends on the compact set $\mathcal{K}$. Since F is furthermore assumed to be periodic outside of a compact set, this bound extends to the whole space, proving the claim that $|f_n - \bar f|_{\kappa,\gamma} \to 0$ in L_p. □

PROOF OF THEOREM A. As above, we define

$$f_n(t,x) = f(x, y_{nt}),$$

for f as in the statement of the theorem. Since $\bar f$ is Lipschitz continuous and $\mathcal{C}^2$, the equation $\dot x_t = \bar f(x_t)\,dB_t$ has a unique solution. Assume now that x^n is a $\mathcal{C}^\alpha$ solution to $dx_t^n = f_n(t, x_t^n)\,dB_t$. Note that since $f_n \in \mathcal{C}^{0,1}$, $t \mapsto f_n(t, x_t^n) \in \mathcal{C}^{-\kappa}$ for any $\kappa < \alpha$ (by Lemma 3.8) and the integral $\int_0^t f_n(s, x_s^n)\,dB_s$ makes sense by Lemma 3.10, so the notion of what constitutes a solution is unambiguous.

We now modify f outside of the ball B_R of radius R centred at 0 so that the resulting function f_R is periodic in its first argument and satisfies the conditions of Lemma 3.14. Write $f_n^R(t,x) \stackrel{\text{def}}{=} f_R(x, y_{tn})$ and denote by $x_t^{n,R}$ and x_t^R the respective solutions to $dx_t = f_R(x_t, y_{nt})\,dB_t$ and $dx_t = \bar f_R(x_t)\,dB_t$. By Theorem 3.13, as $n \to \infty$, $x_t^{n,R}$ converges to x_t^R in probability. The reason why we are able to apply this result is that, even though the functions f_n^R are not deterministic, they are independent of B. Furthermore, one has $x_t^{n,R} = x_t^n$ and $x_t^R = x_t$ before they exit B_R so that, sending $R \to \infty$, we conclude that $x_t^n \to x_t$ in probability, as required. □

To conclude this section, we give a deterministic example of averaging. Fix a function

$$F : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}_+ \to L(\mathbf{R}^m, \mathbf{R}^d),$$
$$(t, x, \tau) \mapsto F_\tau(t,x)$$

with the property that, for any fixed $\tau \in \mathbf{R}_+$, the function F_τ is of class BC^2 and is periodic with period τ in its first argument. We furthermore assume that, for some positive Radon measure μ on $\mathbf{R}_+$ and some $\kappa > 0$, one has

$$\int_0^\infty |F_\tau|_{\mathrm{BC}^2} (1 + \tau^\kappa) \mu(d\tau) < \infty. \tag{3.27}$$

We then set

$$f_n(t,x) = \int_0^\infty F_\tau(nt,x)\mu(d\tau), \qquad \bar f(x) = \int_0^\infty \frac{1}{\tau}\int_0^\tau F_\tau(t,x)\,dt\,\mu(d\tau).$$

REMARK 3.15. A special case is when μ is atomic, which corresponds to the case when f_n is a sum of periodic functions.

PROPOSITION 3.16. *In the above setting, if we set*

$$dx_t^{(n)} = f_n\big(t, x_t^{(n)}\big)\,dB_t, \qquad d\bar x_t = \bar f(\bar x_t)\,dB_t,$$

then $x^{(n)}$ converges in probability to $\bar x$.

PROOF. We have $f_n, \bar f \in \mathrm{BC}^2$ as an immediate consequence of their definitions and (3.27). Note also that if g_ε is periodic with period ε and averages to 0, then one has the bound

$$\left|\int_s^t g_\varepsilon(r)\,dr\right| \le |g_\varepsilon|_\infty\big(|t-s|\wedge\varepsilon\big).$$

As a consequence, one has

$$\begin{aligned}
&\left|\int_s^t \big(f_n(r,x)-\bar f(x)\big)\,dr\right| \\
&\quad\lesssim \int_0^\infty \big|F_\tau(\cdot,x)\big|_\infty\big(|t-s|\wedge n^{-1}\tau\big)\mu(d\tau) \\
&\quad\lesssim |t-s|^{1-\kappa}\int_0^\infty \big|F_\tau(\cdot,x)\big|_\infty n^{-\kappa}\tau^\kappa\mu(d\tau) \lesssim n^{-\kappa}|t-s|^{1-\kappa}.
\end{aligned}$$

Since one similarly has the bound

$$\left|\int_s^t \big(f_n(r,x)-f_n(r,y)-\bar f(x)+\bar f(y)\big)\,dr\right| \lesssim n^{-\kappa}|x-y||t-s|^{1-\kappa},$$

it follows that $|f_n-\bar f|_{-\kappa,1} \lesssim n^{-\kappa}$, and we conclude by Theorem 3.13. □

4. Averaging with feedback. We now turn to the main result of this article, where we allow for feedback from the slow dynamic into the fast dynamic. The trade-off is that our averaging result is not as general as Theorem 3.13, as we require that the fast dynamic is Markovian.

4.1. *A class of slow/fast processes.* Fix a smooth compact manifold $\mathcal{Y}$ for the fast variable and consider the slow/fast system

$$dx_t^\varepsilon = f\big(x_t^\varepsilon, y_t^\varepsilon\big)\,dB_t + g\big(x_t^\varepsilon, y_t^\varepsilon\big)\,dt, \tag{4.1a}$$

$$dy_t^\varepsilon = \frac{1}{\varepsilon}V_0\big(x_t^\varepsilon, y_t^\varepsilon\big)\,dt + \frac{1}{\sqrt{\varepsilon}}V\big(x_t^\varepsilon, y_t^\varepsilon\big)\circ d\hat W_t, \tag{4.1b}$$

$$x_0^\varepsilon = x_0 \in \mathbf{R}^d, \qquad y_0^\varepsilon = y_0 \in \mathcal{Y},$$

where B is a m-dimensional fractional Brownian motion with Hurst parameter $H > \frac12$ and $\hat W$ a $\hat m$-dimensional standard Wiener process independent of B. Also, $f : \mathbf{R}^d \times \mathcal{Y} \to L(\mathbf{R}^m, \mathbf{R}^d)$ and $g : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}^d$. We use the shorthand

$$V\big(x_t^\varepsilon, y_t^\varepsilon\big)\circ d\hat W_t \stackrel{\text{def}}{=} \sum_{k=1}^{\hat m} V_k\big(x_t^\varepsilon, y_t^\varepsilon\big)\circ d\hat W_t^k$$

for vector fields $V_i(x,\cdot)$ on $\mathcal{Y}$. Similarly, $f(x_t^\varepsilon, y_t^\varepsilon)\,dB_t \stackrel{\text{def}}{=} \sum_{k=1}^m f_k(x_t^\varepsilon, y_t^\varepsilon)\,dB_t^k$. We fix a Riemannian metric on $\mathcal{Y}$ and, furthermore, assume that the following holds.

ASSUMPTION 4.1. The drift vector field g is uniformly bounded and globally Lipschitz continuous. Also $f, V_0 \in \mathrm{BC}^2$ and $V_k \in \mathrm{BC}^3$ for $k > 0$. Furthermore, there exists $\lambda > 0$ such that, for all $x \in \mathbf{R}^d$, $y \in \mathcal{Y}$ and $v \in T_y\mathcal{Y}$, one has $\sum_{j>0}\langle v, V_j(x,y)\rangle^2 \geq \lambda |v|^2$.

Solutions will be interpreted as follows. We fix a realisation of the fractional Brownian motion in $\mathcal{C}^\beta$ for some $\beta > \frac{1}{2}$ and we will look for solutions that are Hölder continuous of order α for some $\alpha < \frac{1}{2}$ with $\alpha + \beta > 1$, so that integration with respect to the realisation of fractional Brownian motion can (and will) be interpreted as a Young integral. We will see in Theorem 4.6 below that (4.1) is well-posed and admits solutions in $\mathcal{B}_{\alpha,p}$ for any $\alpha < \frac{1}{2}$. Indeed, by the consideration in Remark 4.5, it is sufficient to consider the case where $\mathcal{Y}$ is a Euclidean space. A posteriori, one can easily show that the slow variables actually satisfy $x_\cdot^\varepsilon \in \mathcal{B}_{\beta,p}$ for any $\beta < H$, but this will not be needed.

For any fixed $(x,y) \in \mathbf{R}^d \times \mathcal{Y}$ and any time $s \in \mathbf{R}$, consider the SDE

$$dY_{s,t} = \frac{1}{\varepsilon} V_0(x, Y_{s,t})\,dt + \frac{1}{\sqrt{\varepsilon}} V(x, Y_{s,t}) \circ d\hat{W}_t, \quad Y_{s,s} = y, \tag{4.2}$$

with $t \geq s$. We write $Y_{s,t} = \bar{\Phi}_{s,t}^x(y)$ for the solution flow associated to this equation, which exists for all time and is unique under our assumptions. The superscript denotes the frozen variable x and we have refrained from adding also ε to the notation.

Write now $\mathcal{P}_t^x$ for the Markov transition semigroup on $\mathcal{Y}$ with generator

$$\mathcal{L}^x = \frac{1}{\varepsilon} V_0(x,\cdot) + \frac{1}{2\varepsilon}\sum_{i=1}^{\hat{m}} \big(V_i(x,\cdot)\big)^2,$$

(vector fields are identified with first-order differential operators in the usual way), so for any points $x \in \mathbf{R}^d$, $y \in \mathcal{Y}$, and bounded measurable $F : \mathcal{Y} \to \mathbf{R}$, $\mathcal{P}_t^x F(y) = \mathbf{E} F(\bar{\Phi}_{s,t}^x(y))$. Note that since $\mathcal{Y}$ is compact and the diffusion for y is uniformly elliptic for any $x \in \mathbf{R}^d$, the solution $\bar{\Phi}_{s,t}^x(y)$ with generator $\mathcal{P}_t^x$ admits a unique invariant probability measure μ^x on $\mathcal{Y}$.

REMARK 4.2. The semigroup $\mathcal{P}_t^x$ depends on ε, but in a trivial way, that is, only though a time change. In particular, the family μ^x of invariant measures does not depend on ε.

Writing $\bar{f}(x) = \int_{\mathcal{Y}} f(x,y)\mu^x(dy)$, and similarly for $\bar{g}$, the following is our main result. The proof of which will be given in Section 4.8.

THEOREM 4.3. *Assume Assumption 4.1. Let B_t be a fBm with Hurst parameter $H > \frac{1}{2}$ and let $\hat{W}_t$ be an independent Brownian motion. Then over any finite time interval $[0,T]$ and for any $\beta < H$, the solution x_t^ε to (4.1) converges, in probability in $\mathcal{C}^\beta$ as $\varepsilon \to 0$, to the unique limit $\bar{x}_t$ solving*

$$d\bar{x}_t = \bar{f}(\bar{x}_t)\,dB_t + \bar{g}(\bar{x}_t)\,dt, \quad \bar{x}_0 = x_0.$$

Furthermore, there exists an exponent $\kappa > 0$ (depending on β) such that

$$\lim_{\varepsilon \to 0} \mathbf{P}\big(|x_t^\varepsilon - \bar{x}_t|_\beta > \varepsilon^\kappa\big) = 0.$$

REMARK 4.4. The lengthy part of the proof is to obtain a priori estimates on the slow variables $\{x^\varepsilon_\cdot\}$ that are uniform in ε. The Young bound is of course useless since the Hölder norm of $r \mapsto y^\varepsilon_r$ diverges as $\varepsilon \to 0$. It is more advantageous to use the sewing technique. Since y^ε_r contains noise not independent of the increments of B_r, (3.17) cannot be directly applied to $\int_s^t f(x^\varepsilon_s, y^\varepsilon_r)\, dB_r$. Instead we break the interaction between the slow and the fast variables by replacing y^ε_r by $Y_{s,r} \stackrel{\text{def}}{=} \bar{\Phi}^{x_s}_{s,r}(y_s)$ and exploit the fact that, since x^ε_s is left frozen at its value at time s and $Y_{s,t}$ is driven by $\hat{W}$, the integrand is independent of $\tilde{B}^s_r$. In order to use these for obtaining estimates, we first prove that for any fixed $\varepsilon > 0$, our notion of integral, used for the purpose of estimation, does still coincide with usual Young integration; see Section 4.3.

We also consider equation (4.1b) separately, with x_t a given $\mathcal{F}_t$ stochastic process not necessarily the solution to (4.1a). More precisely, we consider

$$dy^\varepsilon_t = \frac{1}{\varepsilon} V_0\big(x_t, y^\varepsilon_t\big)\, dt + \frac{1}{\sqrt{\varepsilon}} V\big(x_t, y^\varepsilon_t\big) \circ d\hat{W}_t, \tag{4.3}$$

for x any given sufficiently regular $\mathcal{F}_t$-adapted stochastic process. Its solution flow will be denoted by $\Phi^x_{s,t}$, namely $\Phi^x_{s,t}(y, \omega)$ is the solution to the equation at time $t \geq s$ with $\Phi^x_{s,s}(y, \omega) = y$. For its existence, see Remark 4.5. The chance element ω in the flow is often omitted for simplicity and the superscript denotes the dependence on the auxiliary process x_t.

Given an adapted $\mathbf{R}^d$-valued stochastic processes x_t and an initial condition y_0, we will henceforth use the symbol $Y_{s,t}$ in order to denote the process

$$Y_{s,t} \stackrel{\text{def}}{=} \bar{\Phi}^{x_s}_{s,t}\big(\Phi^x_{0,s}(y_0)\big), \tag{4.4}$$

namely $Y_{s,\cdot}$ is the solution to (4.2) with frozen parameter $x = x_s$ and with initial condition $y = y_s$, with y_s itself given by the solution to (4.3).

REMARK 4.5. First, we assume that (4.3) is defined on the Euclidean space $\mathbf{R}^{d'}$. Suppose $V_0 \in C^1$ and $V_1, \ldots, V_m$ are C^2. Set $\tilde{V}(t, y, \omega) = V(x_t(\omega), y)$, the randomness in x_t is independent of that in $\hat{W}_t - \hat{W}_s$, so there exists a unique global solution to the SDE $dy_t = \tilde{V}(t, y_t, \omega) \circ d\hat{W}_t + \tilde{V}_0(t, y_t, \omega)\, dt$, which follows from the fixed-point argument and the condition $|V(x, y) - V(x, y')| \leq K\, d(y, y')$. On a compact manifold the global existence is trivial. Let (U_i, ϕ_i) be an atlas of charts in M with the property that for every i, $\phi_i(U_i)$ contains the centred ball $B(3r)$ of radius $3r$ and the pre-image of $V_i \stackrel{\text{def}}{=} \phi_i(B(r))$ covers the manifold. Consider the SDE $dy_t = (\phi_i)_* \tilde{V}(t, y_t) \circ d\hat{W}_t + (\phi_i)_* \tilde{V}_0(t, y_t)\, dt$. Since the the vector fields $(\phi_i)_*(V_i)$ have uniform bounds on $B(3r)$, there are uniform estimates, in i, on the exit time of y_t from V_i to U_i.

4.2. *Estimates for SDEs with mixed Young and Itô integration.* In this section, we show that SDEs driven by a fractional Brownian motion with $H > \frac{1}{2}$ and a Wiener process do admit solutions in $\mathcal{B}_{\alpha,p}$ for arbitrary p. This is similar to the results obtained in [11, 21], but since our spaces are slightly different, our result does not appear to follow immediately from theirs. The first estimate below is a semipath by path result: we fixed a realisation of the fBm, then consider the Itô integral.

THEOREM 4.6. *Let $b \in \mathcal{C}^\beta$ where $\beta > \frac{1}{2}$ and consider the equation in $\mathbf{R}^d$*,

$$dz_t = F(z_t)\, db_t + \sigma(z_t)\, d\hat{W}_t + G(z_t)\, dt, \tag{4.5}$$

for some $F \in \mathrm{BC}^2$ and $\sigma, G \in \mathrm{BC}^1$. Here, the first term on the right-hand side is a Young integral while the second term is an Itô integral. Given a time interval $[0, T]$ and numbers $\alpha > 0$ and $p \geq 1$ with $1 - \beta < \alpha < \frac{1}{2}$ (allowing $1 - \beta < \alpha < \beta$ if σ vanishes identically), there exists a unique solution in $\mathcal{B}_{\alpha,p}$. Furthermore,

$$\|z\|_{\alpha,p} \lesssim |b|_\beta^{\frac{1}{\beta}} + 1. \tag{4.6}$$

PROOF. Note that the conclusion is more restrictive for larger values of p so we can choose p sufficiently large so that $\alpha > \frac{1}{p}$ and $\alpha + \beta > 1 + \frac{1}{p}$. Let $z \in \mathcal{B}_{\alpha,p}([0, \delta])$ and recall that, by Kolmogorov's continuity theorem, we have $\mathcal{B}_{\alpha,p} \subset L_p(\Omega, \mathcal{C}^\gamma)$ for any $\gamma < \alpha - \frac{1}{p}$ and, for any fixed $\kappa > 0$, one has

$$\big\| |z|_\gamma \big\|_p \lesssim \delta^{\alpha - \frac{1}{p} - \kappa - \gamma} \|z\|_{\alpha,p}. \tag{4.7}$$

(This is because, on an interval of length δ, $|z|_\gamma \leq \delta^{\bar\gamma - \gamma} |z|_{\bar\gamma}$ for any $0 < \gamma \leq \bar\gamma \leq 1$.) Let us define, for $z \in \mathcal{B}_{\alpha,p}([0, \delta])$,

$$\begin{aligned}\Psi(z) &= z_0 + \int_0^\cdot F(z_r)\, db_r + \int_0^\cdot \sigma(z_r)\, d\hat{W}_r + \int_0^\cdot G(z_r)\, dr \\ &= z_0 + \Psi_1(z) + \Psi_2(z) + \Psi_3(z),\end{aligned}$$

and show that for a sufficiently small value of δ, Ψ maps the ball of radius 1 in $\mathcal{B}_{\alpha,p}([0, \delta])$ centred around z_0 to itself.

Choosing γ such that $\gamma + \beta > 1$ (which is always possible by taking p large enough and κ small enough) and using Young's bound (3.11), we obtain the estimate

$$\begin{aligned}\big|\Psi_1(z)\big|_\alpha &\lesssim \big|F(z)\big|_\gamma |b|_\beta \delta^{\beta + \gamma - \alpha} + |F|_\infty |b|_\beta \delta^{\beta - \alpha} \\ &\leq |F|_{\mathrm{Lip}} |z|_\gamma |b|_\beta \delta^{\beta + \gamma - \alpha} + |F|_\infty |b|_\beta \delta^{\beta - \alpha},\end{aligned} \tag{4.8}$$

where the Hölder seminorm $|z|_\gamma$ is really the Hölder seminorm of $z \restriction [0, \delta]$, while the Hölder seminorm $|b|_\beta$ is considered on the full interval $[0, T]$. Combining this with (4.7) and using the fact that $\| \cdot \|_{\alpha,p} \leq \| | \cdot |_\alpha \|_p$, we obtain the a priori bound

$$\big\|\Psi_1(z)\big\|_{\alpha,p} \lesssim \|z\|_{\alpha,p} |b|_\beta \delta^{\beta - \frac{1}{p} - \kappa} + |b|_\beta \delta^{\beta - \alpha},$$

where we used that $\alpha + \beta - \frac{1}{p} > 1$. As a consequence of the Burkholder–Davis–Gundy inequality, we immediately obtain the bound

$$\big\|\Psi_2(z)\big\|_{\alpha,p} \lesssim |\sigma|_\infty \delta^{\frac{1}{2} - \alpha},$$

provided that $\alpha < \frac{1}{2}$, p is large enough, and κ is small enough. This is the only place where we require that $\alpha < \frac{1}{2}$, which is due to the regularity of the Wiener process. If the Wiener process is absent, we can choose any $\alpha \in (0, \beta)$. Finally, it is trivial that

$$\big\|\Psi_3(z)\big\|_{\alpha,p} \lesssim |G|_\infty \delta^{1 - \alpha}.$$

This shows that, assuming that Ψ does admit a fixed point in $\mathcal{B}_{\alpha,p}$, this fixed point necessarily satisfies the bound

$$\|z\|_{\alpha,p} \lesssim |b|_\beta \delta^{\beta - \alpha} + \delta^{\frac{1}{2} - \alpha} + \delta^{1 - \alpha},$$

provided that $\delta^\beta |b|_\beta \leq c$ for a sufficiently small constant $c > 0$ (depending on F). Since this bound is independent of the initial condition, it can be iterated and necessarily holds on any interval of size δ. It is then straightforward to verify from the definitions that, on the interval $[0, T]$, one has

$$\|z\|_{\alpha,p} \lesssim \delta^{\alpha-1} \sup_{s\in[0,T-\delta]} \|z\restriction[s, s+\delta]\|_{\alpha,p},$$

which implies that, fixing δ with $\delta^\beta |b|_\beta = \frac{1}{2}$,

$$\|z\|_{\alpha,p} \lesssim |b|_\beta \delta^{\beta-1} + \delta^{-\frac{1}{2}} + 1 \lesssim 1 + |b|_\beta^{\frac{1}{\beta}},$$

as claimed.

To show that such a solution exists and is unique, we note that by [11], Theorem 2.2, there exists a unique adapted process $z \in L^2(\Omega, \mathcal{C}^\alpha)$ solving (4.5). To show that it furthermore satisfies the stronger bound (4.6), write τ for the stopping time given by

$$\tau_M = T \wedge \inf\{t \in (0, T] : |z\restriction[0, t]|_\alpha \geq M\}.$$

The process z^M obtained by stopping z at time τ_M then belongs to $\mathcal{B}_{\alpha,p}$ and the same calculation as above shows that it satisfies the bound (4.6). Since one has $\lim_{M\to\infty} \mathbf{P}(z^M = z) = 1$, the claim follows at once. □

Since the fractional Brownian motion has moments of all order, we may take average over all fractional Brownian paths and obtain the corollary below.

CORROLARY 4.7. *Suppose that B and $\hat{W}$ are independent and let F, G and σ be as in Theorem 4.6. Let $\alpha \in (\frac{1}{p}, \frac{1}{2})$ be such that $\alpha + H > 1$. Then for any initial value and any interval $[0, T]$, there exists a unique solution in $\mathcal{B}_{\alpha,p}$ to*

$$dz_t = F(z_t)\, dB_t + \sigma(z_t)\, d\hat{W}_t + G(z_t)\, dt.$$

Furthermore, $\|z\|_{\alpha,p} \lesssim 1 + \||B|_\beta^{\frac{1}{\beta}}\|_p$. (If σ vanishes, we may take $\alpha \in (\frac{1}{p}, H)$.)

4.3. *Stochastic equals Young.* We will make use of the following fact.

LEMMA 4.8. *Let $\alpha \in (0, 1)$ and let F and G be two positive functions such that*

$$F(t) \leq \int_0^t F^{1-\alpha}(s) G(s)\, ds.$$

Then one has the bound $F^\alpha(t) \leq \alpha \int_0^t G(s)\, ds$.

PROOF. Let $\hat{F}(t) = \int_0^t F^{1-\alpha}(s) G(s)\, ds$, so that $\hat{F}$ is a continuous increasing function and we have the bound

$$\frac{d}{dt} \hat{F}^\alpha(t) = \alpha \hat{F}^{\alpha-1}(t) F^{1-\alpha}(t) G(t) \leq \alpha G(t),$$

since $F^{1-\alpha} \leq \hat{F}^{1-\alpha}$. The claim then follows since $F^\alpha \leq \hat{F}^\alpha$. □

Recalling $Y_{s,t}$ given by (4.4) (with x_t arbitrary, not necessarily solution to (4.1a)), standard methods for estimating its deviation from y_t on the time scale of $[s, t]$ blow up exponentially fast as $\varepsilon \to 0$. (For longer times, we will use the smoothing properties of the semigroup.) Recall that $y_u = \Phi^x_{0,u}(y_0) = \Phi^x_{s,u}(y_s)$ and $Y_{s,u} = \bar{\Phi}^{x_s}_{s,u}(y_s)$, and denote by ρ the Riemannian distance on $\mathcal{Y}$.

LEMMA 4.9. *Suppose $x_\cdot \in \mathcal{B}_{\alpha,p}$ where $\alpha < \frac{1}{2}$ and $\alpha + H > 1$. For $p \geq 2$,*

$$\Big\| \sup_{s \leq u \leq t} \rho(y_u, Y_{s,u}) \Big\|_p \lesssim \|x\|_{\alpha,p} \cdot \varepsilon^{-\frac{1}{2}} |t-s|^{\frac{1}{2}+\alpha}, \tag{4.9}$$

provided that $|t-s| \leq \min(\frac{1}{4}\delta, c)\varepsilon$ where δ is the injectivity radius of $\mathcal{Y}$ and c a constant depending on the bounds on V, V_0.

PROOF. Since $\mathcal{Y}$ is compact, we can find a function d which agrees with the Riemannian distance in a neighbourhood of the diagonal and such that d^2 is globally smooth. (Take $d = g \circ \rho$ for $g : \mathbf{R}_+ \to \mathbf{R}$ a smooth concave function with $g(r) = r$ when $r < \delta/4$ and $g(r) = \delta/2$ when $r \geq 3\delta/4$, where δ denotes the injectivity radius of $\mathcal{Y}$.)

We now claim that, by applying Itô's formula to $d^{2p}(y_u, Y_{s,u})$ and then using the Burkholder–Davis–Gundy inequality, one obtains for $p \geq 1$ the bound

$$\begin{aligned} &\mathbf{E} \sup_{s \leq u \leq t} d^{2p}(y_u, Y_{s,u}) \\ &\quad \lesssim \frac{1}{\varepsilon} \int_s^t \mathbf{E} d^{2p}(y_r, Y_{s,r})\, dr + \frac{1}{\varepsilon} \int_s^t \mathbf{E}\big(d^{2p-2}(y_r, Y_{s,r}) |x_u - x_r|^2\big)\, dr. \end{aligned} \tag{4.10}$$

We proceed with the proof based on this, and return to give more explanation at the end of the proof. Let $\sigma_t = \mathbf{E} \sup_{u \leq t} d^{2p}(y_u, Y_{s,u})$. Using Hölder's inequality on the last term, we obtain the bound

$$\sigma_t \lesssim \frac{1}{\varepsilon} \int_s^t \sigma_r^{\frac{2p-2}{2p}} \big(\|x_u - x_r\|_{2p}^2 + \sigma_r^{\frac{1}{p}} \big)\, dr,$$

so that Lemma 4.8 yields

$$\sigma_t^{\frac{1}{p}} \lesssim \frac{1}{\varepsilon} \int_s^t \big(\|x_u - x_r\|_{2p}^2 + \sigma_r^{\frac{1}{p}} \big)\, dr.$$

This allows us to apply Gronwall's inequality, yielding

$$\sigma_t^{\frac{1}{p}} \lesssim \frac{e^{c\frac{t-s}{\varepsilon}}}{\varepsilon} \int_s^t \|x_r - x_u\|_{2p}^2\, dr \lesssim \frac{|t-s|}{\varepsilon} \|x\|_{\alpha,2p}^2 |t-s|^{2\alpha},$$

where we used the bound $|t-s| \leq \varepsilon$ to make sure that the exponential factor doesn't cause an explosion. This is precisely the required bound since we considered d^{2p} rather than ρ^p.

The bound (4.10) is straightforward in Euclidean space using the $x \leftrightarrow y$ symmetry of the distance $|x-y|^2$. If $\mathcal{Y}$ is a compact manifold, $\langle \nabla_x \rho(x,y), v \rangle = -\langle \nabla_y \rho(x,y), \tilde{v} \rangle$, where $v, \tilde{v}$ are tangent vectors at $T_x M$ and $T_y M$, respectively, and are obtained by parallel translations along the geodesic from one to the other. This holds because we only consider x, y such that their Riemannian distance $\rho(x,y)$ is smaller than $1/2$ of the injectivity radius.

This means the stochastic differential dy_t and $dY_{s,t}$ can then be compared using the Lipschitz continuity assumption on the vector fields, modulo the Stratonovich correction term which can be dealt with by the Lipschitz continuity assumption on $\sum_{i=1}^{\hat{m}} \nabla_{V_i} V_i$. The same consideration holds for our modified distance function which is of the form $g \circ \rho$ where g is a smooth real-valued function. □

We consider the two parameter family of stochastic processes:

$$A_{s,t}^{\varepsilon} = \int_s^t h(x_s, Y_{s,r})\, dB_r. \tag{4.11}$$

This is for fixed process $x_\cdot \in \mathcal{B}_{\alpha,p}$ and for $Y_{s,r} = \bar\Phi^{x_s}_{s,t}(\Phi^x_{0,s}(y_0))$, as in (4.4). We will also write y_t for the process obtained by solving (4.3), that is, $y_t = \Phi^x_{0,t}(y_0, \omega)$, and we recall that both y_t and $Y_{s,t}$ depend of course on ε.

We have the following conclusion, which holds for any $x_\cdot \in \mathcal{B}_{\alpha,p}$ (we emphasise that this does in particular include the solution to (4.1a) but we do not restrict ourselves to that case).

LEMMA 4.10. *Assume that $h \in \mathrm{BC}^1$ and that $x \in \mathcal{B}_{\alpha,p}$ with $\alpha + H > 1 + \frac{1}{p}$. Then for each ε fixed, the process $I_t(A^\varepsilon)$ given by Lemma 3.5 coincides with the Young integral $\int_0^t h(x_s, y_s)\, dB_s$.*

PROOF. Note first that under our assumptions, the process $s \mapsto h(x_s, y_s)$ belongs almost surely to $\mathcal{C}^\beta$ for every $\beta < (\alpha - \frac{1}{p}) \wedge \frac{1}{2}$, so that the Young integral is well defined for every $B \in \mathcal{C}^{H-\kappa}$ for κ sufficiently small.

Let $\tilde A^\varepsilon_{s,t} = h(x_s, y_s)(B_t - B_s)$, so that $I_t(\tilde A^\varepsilon)$ coincides with that Young integral by (3.10). As a consequence of the last statement of Lemma 3.5, it is sufficient to show that

$$\|A^\varepsilon_{s,t} - \tilde A^\varepsilon_{s,t}\|_p \lesssim |t-s|^{\bar\eta},$$

for some $\bar\eta > 1$. Note that $\varepsilon > 0$ is fixed, so we are allowed to obtain bounds which diverge as $\varepsilon \to 0$. Apart from x_s, the evolution of $Y_{s,r}$ has no further dependence on the fractional Brownian motion, so that $h(x_s, y_s) - h(x_s, Y_{s,r})$ is $\mathcal{F}_s$-measurable for $\mathcal{F}_s = \mathcal{G}_s \vee \sigma(\hat W)$. We then use Lemma 3.4 with $\kappa = 0$ and $p' < p$, so that

$$\begin{aligned} &\left\| \int_s^t \big(h(x_s, y_s) - h(x_s, Y_{s,r})\big)\, dB_r \right\|_{p'} \\ &\quad \le \big\| |h(x_s, y_s) - h(x_s, Y_{s,\cdot})|_\infty \big\|_p |s-t|^H. \end{aligned} \tag{4.12}$$

By the Lipschitz continuity of h, we have the bound

$$\big\| |h(x_s, y_s) - h(x_s, Y_{s,\cdot})|_\infty \big\|_p \lesssim \Big\| \sup_{r\in[s,t]} \rho(y_s, Y_{s,r}) \Big\|_p.$$

Since the distance ρ is bounded, it follows from Lemma 4.9 that, for every $\kappa > 0$, one has the bound

$$\Big\| \sup_{r\in[s,t]} \rho(y_s, Y_{s,r}) \Big\|_p \lesssim 1 \wedge \big(\varepsilon^{-\frac{1}{2}} |t-s|^{\frac{1}{2}+\alpha}\big) \le \big(\varepsilon^{-1/2}|t-s|^{\frac{1}{2}+\alpha}\big)^{1-\kappa}.$$

Combining this with (4.12), the claim follows. □

4.4. *Semigroups with a parameter: Ergodicity and continuity.* Denote by $|F|_{\mathrm{Lip}}$ the best Lipschitz constant of F and set $|F|_{\mathrm{Osc}} = \sup F - \inf F$. On a space with bounded radius, $|F|_{\mathrm{Osc}} \lesssim |F|_{\mathrm{Lip}}$. Recall that $\mathcal{P}^x_t$ denotes the semigroup associated to (4.2). By differentiating the solution flow, a brutal bound on the derivative flow which is the solution to the equation $dv_t = \varepsilon^{-1/2} \nabla_{v_t} V \circ d\hat W_t + \varepsilon^{-1} \nabla_{v_t} V_0\, dt$ (the estimates depend on the covariant derivatives up to order 2), yields a bound of the type $|\mathcal{P}^x_t F|_{\mathrm{Lip}} \le C e^{Ct/\varepsilon} |F|_{\mathrm{Lip}}$. This can be improved using ergodicity; for large time, we will also make use of the smoothing properties of the Markov semigroups $\mathcal{P}^x$.

LEMMA 4.11. *Under Assumption 4.1, the following holds:*

(1) *There exist constants c, C such that for any $x \in \mathbf{R}^d$,*

$$|\mathcal{P}^x_t F|_{\mathrm{Osc}} \le C e^{-ct/\varepsilon} |F|_{\mathrm{Osc}}, \tag{4.13}$$

$$|\mathcal{P}^x_t F|_{\mathrm{Lip}} \le C e^{-ct/\varepsilon} |F|_{\mathrm{Lip}}, \tag{4.14}$$

$$|\mathcal{P}^x_t F|_{\mathrm{BC}^2} \le C \varepsilon t^{-1} e^{-ct/\varepsilon} |F|_\infty, \tag{4.15}$$

uniformly over $t \in \mathbf{R}_+$.

(2) *For any* $x, \bar{x} \in \mathbf{R}^d$, *the bound*

$$|\mathcal{P}_t^x F - \mathcal{P}_t^{\bar{x}} F|_\infty \le C|x - \bar{x}||F|_{\mathrm{Lip}}, \tag{4.16}$$

holds for all $t \ge 0$.

(3) *If* $h : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}$ *is a Lipschitz continuous bounded function with* $\int_{\mathcal{Y}} h(x, y)\mu^x(dy) = 0$ *for every* x, *then for any* $\kappa \in (0, 1)$,

$$|\mathcal{P}_t^x h(x, \cdot) - \mathcal{P}_t^{\bar{x}} h(\bar{x}, \cdot)|_\infty \lesssim |h|_\infty^\kappa |h|_{\mathrm{Lip}}^{1-\kappa} |x - \bar{x}|^{1-\kappa} e^{-\kappa ct/\varepsilon}. \tag{4.17}$$

PROOF. It follows from standard estimates (see, e.g., [6]) that for any $t \in (0, \varepsilon)$, one has the bounds

$$\begin{aligned} |\mathcal{P}_t^x F|_{\mathrm{Lip}} &\le C\big(1 \vee (t/\varepsilon)^{-1/2}\big)|F|_{\mathrm{Osc}}, \\ |\mathcal{P}_t^x F|_{\mathrm{BC}^2} &\le C\big(1 \vee (t/\varepsilon)^{-1}\big)|F|_\infty, \end{aligned} \tag{4.18}$$

where the constant C only depends on derivatives of V_0 up to order 2 and of the remaining V_k up to order 3. The reason why one obtains the oscillation norm on the right-hand side of the first bound is that $|F|_{\mathrm{Lip}}$ does not change under constant shifts and $|F|_{\mathrm{Osc}} = \inf_{c \in \mathbf{R}} 2|F - c|_\infty$. In fact, (4.18) holds for all t, but we do not need it.

It furthermore follows from the uniform positive lower bounds on the heat kernel (see [1] for the case of $\mathbf{R}^n$ and, e.g., [4, 32] for versions that apply to manifolds; see also [34]) that $|\mathcal{P}_\varepsilon^x F(y_1) - \mathcal{P}_\varepsilon^x F(y_2)| \le (1 - \lambda)|F|_{\mathrm{Osc}}$ for some constant $\lambda > 0$, so that (4.13) follows by iterating this bound (Doeblin's condition).

Using this last inequality for time $t - \varepsilon$ and (4.18) for the remaining time ε, we obtain for $t \ge \varepsilon$ the bound

$$|\mathcal{P}_t^x F|_{\mathrm{Lip}} = |\mathcal{P}_\varepsilon^x \mathcal{P}_{(t-\varepsilon)}^x F|_{\mathrm{Lip}} \le C e^{-ct/\varepsilon} |F|_{\mathrm{Lip}}.$$

For $t \le \varepsilon$ on the other hand, this bound follows from L^p bounds on the Jacobian, so that (4.14) holds. The bound (4.15) follows in the same way.

It is also rather straightforward to verify that

$$|\mathcal{P}_t^x F - \mathcal{P}_t^{\bar{x}} F|_\infty \le C e^{Ct/\varepsilon} |x - \bar{x}||F|_{\mathrm{Lip}}. \tag{4.19}$$

While this bound is good for $t \le \varepsilon$, it can be significantly improved for $t > \varepsilon$. Indeed, for $t \ge \varepsilon$ and any partition Δ of $[0, t]$ into subintervals of size at most ε and at least $\varepsilon/2$, we have

$$\begin{aligned} |\mathcal{P}_t^x F - \mathcal{P}_t^{\bar{x}} F|_\infty &\le \sum_{[s,u] \in \Delta} |\mathcal{P}_{t-u}^x (\mathcal{P}_{u-s}^x - \mathcal{P}_{u-s}^{\bar{x}}) \mathcal{P}_s^{\bar{x}} F|_\infty \\ &\lesssim \sum_{[s,u] \in \Delta} e^{-cs/\varepsilon} |x - \bar{x}||F|_{\mathrm{Lip}} \le C|x - \bar{x}||F|_{\mathrm{Lip}}. \end{aligned}$$

Here, we used (4.14), the small time bound (4.19), and the fact that $\mathcal{P}_{t-u}^x$ is a contraction in L^∞.

For the last bound, we make use of the fact that, since the integral of $h(x)$ against the invariant measure for $\mathcal{P}_t^x$ vanishes, its supremum norm is controlled by its oscillation and vice versa, so that (4.13) yields the bound

$$|\mathcal{P}_t^x h(x, \cdot)|_\infty \le C e^{-ct/\varepsilon} |h|_\infty.$$

Combining this with (4.16), we see that

$$\begin{aligned} |\mathcal{P}_t^x h(x, \cdot) - \mathcal{P}_t^{\bar{x}} h(\bar{x}, \cdot)|_\infty &\lesssim \inf\{|h|_\infty e^{-ct/\varepsilon}, |x - \bar{x}||h|_{\mathrm{Lip}}\} \\ &\lesssim |h|_\infty^\kappa |h|_{\mathrm{Lip}}^{1-\kappa} |x - \bar{x}|^{1-\kappa} e^{-\kappa ct/\varepsilon}, \end{aligned}$$

completing the proof. □

LEMMA 4.12. *Fix $\bar{x} \in \mathbf{R}^d$ and $y \in \mathcal{Y}$ and write $y_t = \bar{\Phi}^{\bar{x}}_{0,t}(y)$. Fix $p \geq 2$ and $F : \mathcal{Y} \to \mathbf{R}$ bounded measurable, and write $\bar{F}(\bar{x}) = \int F(y)\mu^{\bar{x}}(dy)$. Then one has the bound*

$$\left\| \int_0^t \big(F(y_r) - \bar{F}(x)\big)\, dr \right\|_p \lesssim \varepsilon^{\frac{1}{p}} t^{1-\frac{1}{p}} |F|_{\mathrm{Osc}}, \tag{4.20}$$

uniformly over $\bar{x} \in \mathbf{R}^d$, $y \in \mathcal{Y}$ and $t \geq 0$.

PROOF. Since $\inf F \leq \bar{F}(x) \leq \sup F$, one has the almost sure bound

$$\left| \int_0^t \big(F(y_r) - \bar{F}(x)\big)\, dr \right| \leq 2t|F|_{\mathrm{Osc}},$$

so that the general case of (4.20) follows from the case $p = 2$ by interpolation.

For $p = 2$, we write $\tilde{F}(y) = F(y) - \bar{F}(\bar{x})$. With this notation, we have the identity

$$\mathbf{E}\left(\int_0^t \tilde{F}(y_s)\, ds\right)^2 = 2\int_0^t \int_0^r \big(\mathcal{P}^{\bar{x}}_s\big(\tilde{F} \cdot \mathcal{P}^{\bar{x}}_{r-s}\tilde{F}\big)\big)(y)\, ds\, dr.$$

By Lemma 4.11, we can bound the supremum of $\mathcal{P}^{\bar{x}}_{r-s}\tilde{F}$ by $Ce^{-c|r-s|/\varepsilon}|\tilde{F}|_{\mathrm{Osc}}$, giving the bound

$$\mathbf{E}\left(\int_0^t \tilde{F}(y_s)\, ds\right)^2 \leq C|\tilde{F}|^2_{\mathrm{Osc}} \int_0^t \int_0^r e^{-\frac{c|s-r|}{\varepsilon}}\, ds\, dr.$$

Since this integral is bounded by a multiple of εt and since $|\tilde{F}|_{\mathrm{Osc}} = |F|_{\mathrm{Osc}}$, the claim follows. □

CORROLARY 4.13. *We fix an adapted stochastic process x_t with values in $\mathbf{R}^d$ and y_t with values in $\mathcal{Y}$. Let $h : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}$ be bounded measurable, set $Y_{s,t} = \bar{\Phi}^{x_s}_{s,t}(y_s)$, and set $\bar{h}(x) = \int h(x, y)\mu^x(dy)$. For $p \geq 2$ and $s \leq u \leq t$, one has the bound*

$$\left\| \int_u^t \big(h(x_s, Y_{s,r}) - \bar{h}(x_s)\big)\, dr \right\|_p \lesssim |h|_{\mathrm{Osc}} \varepsilon^{\frac{1}{p}} |t-u|^{1-\frac{1}{p}},$$

uniformly over $\varepsilon \in (0, 1]$, and over the processes x_t and y_t.

PROOF. Applying Lemma 4.12 with $\bar{x} = x_s$, $y = Y_{s,u} \equiv \bar{\Phi}^{x_s}_{s,u}(y_s)$ and $F = h(x_s, \cdot)$, we then obtain

$$\mathbf{E}\left(\left| \int_u^t \big(h\big(x_s, \bar{\Phi}^{x_s}_{u,r}(Y_{s,u})\big) - \bar{h}(x_s)\big)\, dr \right|^p \middle| \mathcal{F}_u\right) \lesssim \big|h(x_s, \cdot)\big|_{\mathrm{Osc}} \varepsilon^{\frac{1}{p}} |t-u|^{1-\frac{1}{p}}.$$

Since this bound is uniform over y and since $|h(\bar{x}, \cdot)|_{\mathrm{Osc}} \leq |h|_{\mathrm{Osc}}$ for every $\bar{x}$, the claim follows. □

4.5. *Regularity of limiting drift.* In this section, we show that Assumption 4.1 guarantees that the limiting functions $\bar{f}$ and $\bar{g}$ do again belong to BC^2. Throughout Section 4.5 (and only here), we write $\mathcal{P}^x_t$ for the semigroup generated by (4.2), but with $\varepsilon = 1$, that is, we work on the fast timescale. The reason why we can do this is that we are only interested in showing a result about the invariant measures, and these do not depend on ε. We also write $\bar{\Phi}^x_t$ for the corresponding flow map, so that $(\mathcal{P}^x_t F)(y) = \mathbf{E}F(\bar{\Phi}^x_t(y))$. The main ingredient for the proof is the following claim.

LEMMA 4.14. *Under Assumption 4.1, for any fixed $\tau > 0$, the map $x \mapsto \mathcal{P}^x_\tau$ is differentiable, uniformly in x, as a map $\mathbf{R}^d \to L(\mathrm{BC}^2, \mathrm{BC}^1)$ and as a map $\mathbf{R}^d \to L(\mathrm{BC}^1, \mathrm{BC}^0)$. It is also twice differentiable, uniformly in x, as a map $\mathbf{R}^d \to L(\mathrm{BC}^2, \mathrm{BC}^0)$.*

PROOF. Consider the semigroup $\bar{\mathcal{P}}_t$ on $\mathbf{R}^d \times \mathcal{Y}$ given by

$$(\bar{\mathcal{P}}_t F)(x, y) = \mathbf{E} F\big(x, \bar{\Phi}_t^x(y)\big).$$

By [6], Proposition A8, and [24] (see also [22] for related results), $\bar{\mathcal{P}}_t$ maps BC^k into itself for $k \in \{0, 1, 2\}$, and the claim follows. □

LEMMA 4.15. *Under Assumption* 4.1, *the map* $\bar{f}(x) = \int_{\mathcal{Y}} f(x, y)\mu^x(dy)$ *belongs to* BC^2.

PROOF. For any given x and t, we view $\mathcal{P}_t^x$ as a bounded linear operator on the spaces BC^k of k times continuously differentiable functions on $\mathcal{Y}$, $k = 0, 1, 2$. Writing Π_x for the projection operator given by $\Pi_x F = \langle \mu^x, F\rangle \mathbf{1}$, where $F : \mathcal{Y} \to \mathbf{R}$ and $\mathbf{1}$ denotes the constant function, we note that Π_x commutes with $\mathcal{P}_t^x$ and that, by part (1) of Lemma 4.11, we can choose t sufficiently large so that $|(1 - \Pi_x)\mathcal{P}_t^x F|_{\mathrm{BC}^k} \le \frac{1}{2}|F|_{\mathrm{BC}^k}$ for $k \in \{0, 1, 2\}$, uniformly over x. We used the fact that $F - \Pi_x F$ is centred.

We fix such a value of t from now on. Writing $R^x(\lambda)$ for the resolvent of $\mathcal{P}_t^x$, it follows that that for each $k \in \{0, 1, 2\}$, the operator norm of $R^x(\lambda)$ in BC^k is bounded by 4, uniformly in x and uniformly over λ belonging to the circle γ of radius $\frac{1}{4}$ centred at 1. Indeed, for $B = \mathrm{BC}^k$, we write $B = \langle \mathbf{1}\rangle \oplus B^\perp$ where $B^\perp = \{F : \langle \mu^x, F\rangle = 0\}$, and view Π_x as the projection onto $\langle \mathbf{1}\rangle$ for this decomposition.

Since Π_x commutes with $\mathcal{P}_t^x$, that operator splits with respect to this decomposition as $\mathcal{P}_t^x = \mathrm{id} \oplus (1 - \Pi_x)\mathcal{P}_t^x$ and its resolvent is given by

$$R^x(\lambda) = \big(\lambda - \mathcal{P}_t^x\big)^{-1} = (\lambda - \mathrm{id})^{-1} \oplus \big(\lambda - (1 - \Pi_x)\mathcal{P}_t^x\big)^{-1}.$$

The first term is obviously bounded by 4, while the second term is given by the convergent Neumann series

$$\lambda^{-1}\bigg(1 + \frac{1}{\lambda}(1 - \Pi_x)\mathcal{P}_t^x + \frac{1}{\lambda^2}\big((1 - \Pi_x)\mathcal{P}_t^x\big)^2 + \cdots\bigg),$$

which is also bounded by 4 in operator norm since $|\lambda| \ge \frac{3}{4}$ and $\|(1 - \Pi_x)\mathcal{P}_t^x\| \le \frac{1}{2}$. We claim that, uniformly over $\lambda \in \gamma$ and $x \in \mathbf{R}^d$, the map $x \mapsto R^x(\lambda)$ is $\mathcal{C}^2$ as a map into $L(\mathrm{BC}^2, \mathrm{BC}^0)$ and $\mathcal{C}^1$ as a map into $L(\mathrm{BC}^2, \mathrm{BC}^1)$ and into $L(\mathrm{BC}^1, \mathrm{BC}^0)$. Indeed, this is an immediate consequence of the identities

$$\begin{aligned} D_x R^x &= R^x\big(D_x \mathcal{P}_t^x\big), R^x, \\ D_x^2 R^x &= 2R^x\big(D_x \mathcal{P}_t^x\big)R^x\big(D_x \mathcal{P}_t^x\big)R^x + R^x\big(D_x^2 \mathcal{P}_t^x\big)R^x, \end{aligned} \tag{4.21}$$

combined with Lemma 4.14.

We now recall that one has [18], Theorem III.6.17,

$$\Pi_x = \frac{1}{2i\pi}\oint_\gamma R^x(\lambda)\, d\lambda.$$

In particular, for any fixed probability measure μ on $\mathcal{Y}$, one has the identity

$$\bar{f}(x) = \frac{1}{2i\pi}\oint_\gamma \big\langle \mu, R^x(\lambda) f(x, \cdot)\big\rangle\, d\lambda. \tag{4.22}$$

It now suffices to note that, by our assumptions, one has $D_x^k f(x, \cdot) \in \mathrm{BC}^{2-k}$ for $k \in \{0, 1, 2\}$ uniformly over x and, therefore, the claim follows from (4.21) combined with Lemma 4.14. □

4.6. *Uniform estimates on the slow variable.* The main theorem in this section is a uniform estimate for the slow variables. The divergence of the Hölder norm of $r \mapsto y_r^\varepsilon$ means that the bounds given on A^ε and $\tilde{A}^\varepsilon$ in the proof of Lemma 4.10 diverge as $\varepsilon \to 0$, and are thus inadequate to show any kind of tightness result. This is where our precise choice of A^ε comes in. We will show that A^ε belongs to the Banach space $H_\eta^p \cap \bar{H}_{\bar{\eta}}^p$ with uniform (in ε) upper bounds on its norms.

The estimates obtained in Lemma 4.11 on $\mathcal{P}_t^x$ are not quite sufficient for our use, we will also introduce a second family of random semigroups $\mathcal{Q}_{s,t}^x$ generated by the flow $\Phi_{s,t}^x$; see Definition 4.18 below. Given the Wiener process $\hat{W}_t$ and the fractional Brownian motion B_t as before, we define the filtration $\mathcal{F}_t = \mathcal{G}_t \vee \hat{\mathcal{G}}_t$, where

$$\mathcal{G}_t = \sigma\{B_u - B_r : r \le u \le t\}, \qquad \hat{\mathcal{G}}_t = \sigma\{\hat{W}_u - \hat{W}_r : r \le u \le t\}.$$

Observe that $\mathcal{G}_t = \sigma\{W_u - W_r : r \le u \le t\}$. We will also make use of the 'noise'

$$\hat{\mathcal{G}}_t^s = \sigma\{\hat{W}_u - \hat{W}_r : s \le r \le u \le t\}, \tag{4.23}$$

and also $\mathcal{G}_t^s$ defined using W_t. In the definition for $\mathcal{B}_{\alpha,p}$, we use the filtration $\mathcal{F}_t$ given above.

The following theorem is the main technical tool in the proof of Theorem 4.3 as it yields uniform bounds in ε on the fixed-point map defining x. Its proof relies on three further lemmas: Lemmas 4.23, 4.22 and 4.24, given after the proof of the theorem.

THEOREM 4.16. *Let $T > 0$, let $p \ge 2$ and $\alpha \in (0, \frac{1}{2})$ be such that $1 + \frac{1}{p} < \alpha + H$, let $x_\cdot$ be a $\mathbf{R}^d$-valued process in $\mathcal{B}_{\alpha,p}$ and let $y_\cdot^\varepsilon$ be the solution to (4.3) where V is assumed to satisfy Assumption 4.1. Then there are exponents $\eta > \frac{1}{2}$ and $\bar{\eta} > 1$ such that, for $h : \mathbf{R}^d \times \mathcal{Y} \to \mathbf{R}$ a bounded uniformly Lipschitz continuous function, one has the following*:

1. *Suppose in addition that $\bar{h}(x) \stackrel{\text{def}}{=} \int_{\mathcal{Y}} h(x, y)\mu^x(dy) = 0$ for all x. Then for any $\beta < H$, there exists a constant $\kappa > 0$ such that*

$$\left\| \int_s^t h(x_r, y_r^\varepsilon)\, dB_r \right\|_p \lesssim \varepsilon^\kappa |h|_{\mathrm{BC}^1} \big(\|x\|_{\alpha,p} |t-s|^{\bar{\eta}} + |t-s|^\eta \big), \tag{4.24}$$

holds uniformly over ε and over x_t.

2. *For general h, one has the bound*

$$\left\| \int_s^t h(x_r, y_r^\varepsilon)\, dB_r \right\|_p \lesssim |h|_{\mathrm{BC}^1} \big(\|x\|_{\alpha,p} |t-s|^{\bar{\eta}} + |t-s|^\eta \big). \tag{4.25}$$

PROOF. Fix an arbitrary $\mathbf{R}^d$-valued stochastic process x_t in $\mathcal{B}_{\alpha,p}$ (which is not necessarily a solution to our equation). We first note that the bound (4.25) for general h (i.e., with $\bar{h} \neq 0$) follows from the bound (4.24), combined with the estimate (3.18) for $\int_0^t \bar{h}(x_r)\, dB_r$. We therefore only focus on the proof of (4.24) and assume that $\bar{h} = 0$ from now on.

During the rest of the section, we will also suppress the superscript ε whenever possible. Note that standard estimates for the Young integral $\int_0^t h(x_r, y_r^\varepsilon)\, dB_r$ obtained by taking limits in the Riemann sum $\sum_{[v,u]\in\mathcal{P}} h(x_s, y_s)(B_u - B_v)$ would require uniform (in ε) bounds on the Hölder norms on y^ε, which we do not have.

In order to obtain estimates that are uniform in ε, it will be useful to use Lemma 4.10 and write the integral as $\lim_{|\mathcal{P}|\to 0} \sum_{[u,v]\in\mathcal{P}} A_{u,v}$, where A is given by

$$A_{s,t} = \int_s^t h(x_s, Y_{s,r})\, dB_r.$$

In order to obtain the bound (4.24) from Lemma 3.5, it therefore remains to obtain a bound on $\|A\|_{\eta,p}$ and $|||A|||_{\bar{\eta},p}$ for some $p \ge 2$, some $\eta > \frac{1}{2}$ and some $\bar{\eta} > 1$. The bound on $\|A\|_{\eta,p}$ is contained in the following lemma, the proof of which is relatively straightforward.

LEMMA 4.17. *Assume that* $\bar h = 0$. *For every* $p \geq 2$ *and* $\kappa > 0$, *one has* $\{A_{s,t}\} \in H^p_\eta$ *with* $\eta = H - \kappa$ *and*

$$\|A_{s,t}\|_p \lesssim \varepsilon^{\bar\kappa} |t-s|^{H-\kappa},$$

uniformly over $s, t \in [0, T]$, *provided that* $\bar\kappa < (\kappa/2) \wedge (1/p)$.

PROOF. To prove that $\|A\|_{\eta,p} = \sup_{s<t} \frac{\|A_{s,t}\|_p}{|t-s|^\eta}$ is finite, we first write

$$F_s(r) = h(x_s, Y_{s,r}),$$

so that

$$A_{s,t} = \int_s^t F_s(r)\, dB_r.$$

Furthermore, conditional on $\mathcal{F}_s$, F_s and B are independent, so that we can apply Lemma 3.4, yielding for $q > p$ and $\kappa \in [0, 1)$ the bound

$$\|A_{s,t}\|_p \lesssim \big\| |F_s|_{-\kappa} \big\|_q |t-s|^{H-\kappa}.$$

On the other hand, it follows from Corollary 4.13 exploiting ergodicity, that for $s \leq u \leq v \leq t$, one has the bound

$$\left\| \int_u^v F_s(r)\, dr \right\|_q \lesssim \varepsilon^{\frac{1}{q}} |v-u|^{1-\frac{1}{q}},$$

so that Kolmogorov's continuity theorem implies the bound

$$\big\| |F_s|_{-\kappa} \big\|_q = \left\| \sup_{u \neq v} |u-v|^{1-\kappa} \int_u^v F_s(r)\, dr \right\|_q \lesssim \varepsilon^{\frac{1}{q}},$$

provided that $\kappa > \frac{2}{q}$. Choosing $q = 1/\bar\kappa$ completes the proof. □

It now remains to show that $A_{s,t} \in \bar H^p_{\bar\eta}$ for some $\bar\eta > 1$ and to obtain a suitable bound for small values of ε. We have the identity

$$\delta A_{\text{sut}} = \int_u^t \big(h(x_s, Y_{s,r}) - h(x_u, Y_{u,r})\big)\, dB_r.$$

Recall that this integral is defined as the sum of a Wiener integral against $\tilde B^u_r$ and a Riemann–Stieltjes integral against $\bar B^u_r$. Since $\tilde B^u_r$ is independent of $\mathcal{F}_u \vee \hat{\mathcal{G}}_t$, while $h(x_s, Y_{s,r}) - h(x_u, Y_{u,r})$ is measurable with respect to it, the Wiener integral has vanishing conditional expectation against $\mathcal{F}_u$, so that

$$\begin{aligned}\mathbf{E}(\delta A_{\text{sut}} | \mathcal{F}_u) &= \int_u^t \mathbf{E}\big(h(x_s, Y_{s,r}) - h(x_u, Y_{u,r}) | \mathcal{F}_u\big) \dot{\bar B}^u_r\, dr \\ &= \int_u^t \big(\mathcal{P}^{x_s}_{r-u} h(x_s, \cdot)(Y_{s,u}) - \mathcal{P}^{x_u}_{r-u} h(x_u, \cdot)(y_u)\big) \dot{\bar B}^u_r\, dr.\end{aligned}$$

In the last line, we have used

$$Y_{s,r} = \bar\Phi^{x_s}_{s,r}(y_s) = \bar\Phi^{x_s}_{u,r} \bar\Phi^{x_s}_{s,u}(y_s) = \bar\Phi^{x_s}_{u,r}(Y_{s,u}), \qquad Y_{u,r} = \bar\Phi^{x_u}_{u,r}(y_u).$$

It seems to be difficult to get a good enough bound on this expression, so we exploit the fact that we really only need to bound the conditional expectation of δA_{sut} with respect to $\mathcal{F}_s$ rather than $\mathcal{F}_u$. Conditioning on $\mathcal{F}_s$ however has the unfortunate side effect that it no longer keeps this term separate from the term $\dot{\bar{B}}_r^u$. Instead, we are going to condition on $\mathcal{F}_s \vee \mathcal{G}_u$. We have

$$\begin{aligned}
&\mathbf{E}(\delta A_{\mathrm{sut}}|\mathcal{F}_s \vee \mathcal{G}_u) \\
&= \int_u^t \mathbf{E}\big(\mathcal{P}_{r-u}^{x_s} h(x_s, \cdot)(Y_{s,u}) - \mathcal{P}_{r-u}^{x_u} h(x_u, \cdot)(y_u)|\mathcal{F}_s \vee \mathcal{G}_u\big)\dot{\bar{B}}_r^u \, dr \qquad (4.26)\\
&= \int_u^t \big(\mathcal{P}_{y-s}^{x_s}\mathcal{P}_{r-u}^{x_s} h(x_s, \cdot)(y_s) - \mathbf{E}\big(\mathcal{P}_{r-u}^{x_u} h(x_u, \cdot)(y_u)|\mathcal{F}_s \vee \mathcal{G}_u\big)\big)\dot{\bar{B}}_r^u \, dr.
\end{aligned}$$

Given a 'final time' u, we write

$$\begin{aligned}
\mathcal{U}_s^u = \big\{&F : \Omega \times \mathcal{Y} \to \mathbf{R} : \\
&F \text{ is bounded and } (\mathcal{F}_s \vee \mathcal{G}_u) \otimes \mathcal{B}(\mathcal{Y})\text{-measurable}\big\}.
\end{aligned} \qquad (4.27)$$

Given an element $F \in \mathcal{U}_s^u$, we will use various norms for its $\mathcal{Y}$-dependency, but will always keep the ω-dependency fixed, so that these norms are interpreted as $\mathbf{R}_+$-valued random variables. For example, we set

$$|F|_\infty(\omega) = \sup_{y\in\mathcal{Y}} \big|F(\omega, y)\big|, \qquad |F|_{\mathrm{Lip}}(\omega) = \sup_{\bar{y}\neq y\in\mathcal{Y}} \rho(y, \bar{y})^{-1}\big|F(\omega, y) - F(\omega, \bar{y})\big|,$$

and similarly for $|F|_{\mathrm{Osc}}(\omega)$, but we will always denote them simply by $|F|_\infty$, $|F|_{\mathrm{Lip}}$, etc.

For any stochastic process x (not necessarily a solution to our equation) adapted to the full filtration $\mathcal{F}$, we then define a collection of bounded linear operators $\mathcal{Q}_{r,v}^x : \mathcal{U}_v^u \to \mathcal{U}_r^u$ in the following way.

DEFINITION 4.18. Given a fixed value u and a process x adapted to $\mathcal{F}$, we set for $r \le v \le u$ and $F \in \mathcal{U}_v^u$,

$$(\mathcal{Q}_{r,v}^x F)(\omega, y) \stackrel{\text{def}}{=} \mathbf{E}\big(F\big(\cdot, \Phi_{r,v}^x(y, \cdot)\big)|\mathcal{F}_r \vee \mathcal{G}_u\big)(\omega). \qquad (4.28)$$

REMARK 4.19. The fact that we have $\mathcal{G}_u$ and not $\mathcal{G}_v$ in the right-hand side of (4.28) is not a typo. We always condition on the whole trajectory of the fractional Brownian motion B up to the 'final' time u. Observe that $\mathcal{Q}_{r,v}^x F$ is a three-parameter family of stochastic processes, and could be denoted by $\mathcal{Q}_{r,v}^{x,u} F$.

Since $\bar{\Phi}_{r,v}^{\bar{x}}(y)$ is independent of $\mathcal{G}_u$, for $u \ge v \ge r$, we can also build from $\mathcal{P}_{v-r}^{\bar{x}}$ an operator $\hat{\mathcal{P}}_{r,v}^{\bar{x}} : \mathcal{U}_v^u \to \mathcal{U}_r^u$ by setting

$$\big(\hat{\mathcal{P}}_{r,v}^{\bar{x}} F\big)(\omega, y) \stackrel{\text{def}}{=} \mathbf{E}\big(\big(\mathcal{P}_{v-r}^{\bar{x}} F\big)(\cdot, y)|\mathcal{F}_r \vee \mathcal{G}_u\big)(\omega).$$

By $\mathcal{P}_{v-r}^{\bar{x}} F(\omega, y)$, we mean applying the semigroup to each $F(\omega, \cdot)$, so if F happens to be $\mathcal{F}_r \vee \mathcal{G}_u$-measurable, then $\hat{\mathcal{P}}_{r,v}^{\bar{x}}$ coincides with $\mathcal{P}_{v-r}^{\bar{x}}$, applied ω-wise.

Using these notation, we can rewrite (4.26) as

$$\begin{aligned}
\mathbf{E}(\delta A_{\mathrm{sut}}|\mathcal{F}_s \vee \mathcal{G}_u) = \int_u^t &\big(\hat{\mathcal{P}}_{s,u}^{x_s}\mathcal{P}_{r-u}^{x_s} h(x_s, \cdot)(y_s) \\
&- \mathcal{Q}_{s,u}^x \mathcal{P}_{r-u}^{x_u} h(x_u, \cdot)(y_s)\big)\dot{\bar{B}}_r^u \, dr.
\end{aligned} \qquad (4.29)$$

The expression in (4.29) then naturally splits into two parts. The first part is given by

$$I_1 \stackrel{\text{def}}{=} \int_u^t \hat{\mathcal{P}}_{s,u}^{x_s}\big(\mathcal{P}_{r-u}^{x_s} h(x_s, \cdot) - \mathcal{P}_{r-u}^{x_u} h(x_u, \cdot)\big)(y_s)\dot{\bar{B}}_r^u \, dr.$$

We then apply the estimate in (4.17), namely

$$\big|\mathcal{P}_t^x h(x,\cdot) - \mathcal{P}_t^{\bar x} h(\bar x,\cdot)\big|_\infty \lesssim |h|_\infty^\kappa |h|_{\mathrm{Lip}}^{1-\kappa} |x-\bar x|^{1-\kappa} e^{-\kappa c t/\varepsilon}.$$

This term is then bounded by

$$|I_1| \lesssim |h|_{\mathrm{Lip}}^{1-\kappa} |h|_\infty^\kappa \mathbf{E}\big(|x_s - x_u|^{1-\kappa} | \mathcal{F}_s \vee \mathcal{G}_u\big) \int_u^t e^{-\kappa c(r-u)/\varepsilon} \big|\dot{\bar B}_r^u\big|\,dr. \tag{4.30}$$

Consequently, for $\frac{1}{p'} + \frac{1}{q'} = 1$,

$$\|I_1\|_p \lesssim |h|_{\mathrm{Lip}}^{1-\kappa} |h|_\infty^\kappa \big\|\mathbf{E}\big(|x_s - x_u|^{1-\kappa} | \mathcal{F}_s \vee \mathcal{G}_u\big)\big\|_{pp'} \bigg\|\int_u^t e^{-\frac{\kappa c(r-u)}{\varepsilon}} \big|\dot{\bar B}_r^u\big|\,dr\bigg\|_{pq'}.$$

We choose $p' = (1-\kappa)^{-1}$ and $q' = \kappa^{-1}$, which yields the bound

$$\big\|\mathbf{E}\big(|x_s - x_u|^{1-\kappa} | \mathcal{F}_s \vee \mathcal{G}_u\big)\big\|_{pp'} \le \|x\|_{\alpha,p} |s-u|^{(1-\kappa)\alpha}$$

recall that $\|\dot{\bar B}_r^u\|_q \lesssim |r-u|^{H-1}$ for every $q \ge 1$ so that

$$\bigg\|\int_u^t e^{-\frac{\kappa c(r-u)}{\varepsilon}} \big|\dot{\bar B}_r^u\big|\,dr\bigg\|_{pq'} \lesssim \varepsilon^H \wedge |t-u|^H.$$

Combining these bounds, we conclude that for every $\bar\eta < H + \alpha$ there exist $\kappa, \bar\kappa > 0$ such that

$$\|I_1\|_p \lesssim \varepsilon^{\bar\kappa} \|x\|_{\alpha,p} |t-u|^{\bar\eta} |h|_{\mathrm{Lip}}^{1-\kappa} |h|_\infty^\kappa.$$

The remaining term is given by

$$I_2 \stackrel{\mathrm{def}}{=} \int_u^t \big(\big(\hat{\mathcal{P}}_{s,u}^{x_s} - \mathcal{Q}_{s,u}^x\big) \mathcal{P}_{r-u}^{x_u} h(x_u,\cdot)\big)(y_s)\big) \dot{\bar B}_r^u\,dr.$$

We then apply the following estimate from Lemma 4.24, to be found after this proof,

$$\big|\hat{\mathcal{P}}_{s,u}^{x_s} F - \mathcal{Q}_{s,u}^x F\big|_\infty \lesssim \sqrt{\mathbf{E}\big(|x|_{\bar\alpha}^2 | \mathcal{F}_s \vee \mathcal{G}_u\big)} |u-s|^{\bar\alpha} |F|_{\mathrm{Lip}},$$

where the Hölder norm $|x|_{\bar\alpha}$ is taken on $[s,u]$ and $\bar\alpha < H$ is a number to be chosen, to deduce that

$$\begin{aligned} |I_2| &= \bigg|\int_u^t \big(\hat{\mathcal{P}}_{s,u}^{x_s} \mathcal{P}_{r-u}^{x_s} h(x_s,\cdot) - \mathcal{Q}_{s,u}^x \mathcal{P}_{r-u}^{x_s} h(x_s,\cdot)\big)(y_s) \dot{\bar B}_r^u\,dr\bigg| \\ &\lesssim \int_u^t \sqrt{\mathbf{E}\big(|x|_{\bar\alpha}^2 | \mathcal{F}_s \vee \mathcal{G}_u\big)} |u-s|^{\bar\alpha} \big|\mathcal{P}_{r-u}^{x_s} h(x_s,\cdot)\big|_{\mathrm{Lip}} \big|\dot{\bar B}_r^u\big|\,dr. \end{aligned}$$

We apply to this the following estimate obtained in Lemma 4.11:

$$\big|\mathcal{P}_{r-u}^{x_s} h(x_s,\cdot)\big|_{\mathrm{Lip}} \le C e^{-c(r-u)/\varepsilon} \big|h(x_s,\cdot)\big|_{\mathrm{Lip}} \le C e^{-c(r-u)/\varepsilon} |h|_{\mathrm{Lip}}.$$

Then, provided that we choose $\bar\alpha$ and p in such a way that $\alpha < \bar\alpha - \frac{1}{p}$, we can apply Kolmogorov's continuity theorem yielding

$$\|I_2\|_p \lesssim \|x\|_{\alpha,p} \varepsilon^{\bar\kappa} |h|_{\mathrm{Lip}} (t-s)^{\bar\alpha + H - \bar\kappa}.$$

Combining these estimates, we have shown that, provided that we choose p sufficiently large and $\bar\kappa$ sufficiently small, there exists $\bar\eta > 1$ and a constant $C(h)$ such that

$$\big\|\mathbf{E}(\delta A_{\mathrm{sut}} | \mathcal{F}_s \vee \mathcal{G}_u)\big\|_p \le C(h) \|x\|_{\alpha,p} \varepsilon^{\bar\kappa} (t-s)^{\bar\eta}.$$

When combining this with Lemma 4.17, we have proved that A belongs to $H^p_\eta \cap \bar H^p_{\bar\eta}$ with $\eta > \frac{1}{2}$ and $\bar\eta > 1$, and we have obtained bounds for it that are of order $\varepsilon^{\bar\kappa}$ for sufficiently small $\bar\kappa > 0$. We also know from Lemma 4.10 that $I_t(A^\varepsilon)$ equals the Young integral $\int_0^t h(x_s, y_s^\varepsilon)\, dB_s$. Applying Lemma 3.5, this leads to the bound

$$\begin{aligned}\left\| \int_s^t h(x_r^\varepsilon, y_r^\varepsilon)\, dB_r \right\|_p &\lesssim \left(\vert\!\vert\!\vert A \vert\!\vert\!\vert_{\bar\eta,p} |t-s|^{\bar\eta} + \|A\|_{\eta,p} |t-s|^{\eta} \right) \\ &\le C(h) \|x\|_{\alpha,p} \varepsilon^{\bar\kappa} |t-s|^{\bar\eta} + \varepsilon^{\bar\kappa} |t-s|^{\eta},\end{aligned}$$

uniformly over $\varepsilon \in (0,1]$, thus completing the proof of (4.24). □

CORROLARY 4.20. *Suppose that Assumption 4.1 holds. The solutions x^ε to (4.1) are uniformly bounded in $\mathcal{B}_{\alpha,p}$ for any $\alpha < H$ and $p \ge 1$.*

PROOF. The assumptions on our data guarantee that, for each $\varepsilon > 0$, there exists a unique solution to (4.1) that belong to $\mathcal{B}_{\alpha,p}$, so we only need to obtain the uniform bound. By Theorem 4.16, we obtain for the time interval $[0,T]$ the bound

$$\|x\|_{\alpha,p} \lesssim C\left(|f|_\infty, |f|_{\mathrm{Lip}}\right) T^\kappa \left(1 + \|x\|_{\alpha,p}\right) + T|g|_\infty,$$

where $\kappa > 0$, which implies the required bound on a sufficiently short time interval. One concludes by iterating the bound. □

REMARK 4.21. It is clear from the proof of Corollary 4.20 that instead of assuming that g is bounded, it suffices to guarantee that it satisfies a bound of the form $\| \int_0^T g(x_r, y_r^\varepsilon)\, dr \|_p \lesssim T^\kappa (1 + \|x\|_{\alpha,p})$ for some $\kappa > 0$. (Here, y_r^ε solves (4.3) driven by x as usual.)

4.7. *Bounds on the random semigroup.* In the rest of this section, we fix a $\mathcal{F}_t$-adapted stochastic process x_t and as usual $\Phi^x_{s,t}$ denotes the solution flow to (4.3). For any $\bar x \in \mathbf{R}^d$, fixed, $\Phi^{\bar x}_{s,t}$ denotes the solution to (4.2) with frozen variable $\bar x$. We first bound the difference between the evolutions of Φ^x and $\bar\Phi^{\bar x}$ over a short time period $[r, r']$.

LEMMA 4.22. *Suppose that $x \in \mathcal{B}_{\alpha,p}$. Let $F : \Omega \times \mathcal{Y} \to \mathbf{R}$ be bounded and $(\mathcal{F}_s \vee \mathcal{G}_u) \otimes \mathcal{B}(\mathcal{Y})$ measurable. Then, for $s \le r < r' \le u$ with $|r' - r| \le \varepsilon$ and for $\bar x \in \mathbf{R}^d$, one has the almost sure bound*

$$\left| \hat{\mathcal{P}}^{\bar x}_{r,r'} F - \mathcal{Q}^x_{r,r'} F \right|_\infty \lesssim \sqrt{\sup_{v \in [r,r']} \mathbf{E}\left(|x_v - \bar x|^2 | \mathcal{F}_r \vee \mathcal{G}_u\right)}\, |F|_{\mathrm{Lip}}. \tag{4.31}$$

PROOF. Since $\bar\Phi^{\bar x}_{r,r'}(y_s)$ depends on the filtration of B only through the value of y_s, it is measurable with respect to $\mathcal{F}_r \vee \hat{\mathcal{G}}^r_{r'}$; cf. (4.23). Since furthermore $\hat{\mathcal{G}}^r_{r'}$ is independent of $\mathcal{G}_u$, it follows that

$$\left(\hat{\mathcal{P}}^{\bar x}_{r,r'} F(\omega, \cdot)\right)(y) = \mathbf{E}\left(F\left(\omega, \bar\Phi^{\bar x}_{r,r'}(y)\right) | \mathcal{F}_r \vee \mathcal{G}_u\right).$$

We now have

$$\begin{aligned}\left| \left(\mathcal{P}^{\bar x}_{r,r'} F - \mathcal{Q}^x_{r,r'} F\right)(\omega, y) \right| &= \left| \mathbf{E}\left(F\left(\omega, \bar\Phi^{\bar x}_{r,r'}(y)\right) - F\left(\omega, \Phi^x_{r,r'}(y)\right) | \mathcal{F}_r \vee \mathcal{G}_u\right) \right| \\ &\le |F|_{\mathrm{Lip}} \mathbf{E}\left(\rho\left(\bar\Phi^{\bar x}_{r,r'}(y), \Phi^x_{r,r'}(y)\right) | \mathcal{F}_r \vee \mathcal{G}_u\right).\end{aligned}$$

We then apply Itô's formula to $d(\bar\Phi^{\bar x}_{r,r'}(y), \Phi^x_{r,r'}(y))$, where d is a modification of ρ such that d^2 is smooth. Since the increments of $\hat W$ on $[r, r']$ are independent of $\mathcal{F}_r \vee \mathcal{G}_u$, its martingale term vanishes after taking conditional expectation with respect to $\mathcal{F}_r \vee \mathcal{G}_u$. The rest of the estimate for the distance is routine (see Lemma 4.9) and the required bound follows. □

We fix a 'final time' u and recall that $\mathcal{U}_s^u$ is the space of bounded real valued functions from $\Omega \times \mathcal{Y}$ that are measurable with respect to $(\mathcal{F}_s \vee \mathcal{G}_u) \otimes \mathcal{B}(\mathcal{Y})$; cf. (4.27).

LEMMA 4.23. *Let $s < r < v$. Let $F \in \mathcal{U}_s^u$ be a function that is continuous in the second variable for almost every ω. The operators $\mathcal{Q}_{s,r}^x : \mathcal{U}_r^u \to \mathcal{U}_s^u$ defined by (4.28) satisfy the composition rule*

$$\mathcal{Q}_{s,r}^x \circ \mathcal{Q}_{r,v}^x F = \mathcal{Q}_{s,v}^x F.$$

PROOF. Since $\Phi_{r,v}^x(\Phi_{s,r}^x(y,\omega),\omega) = \Phi_{s,v}^x(y,\omega)$ and $\omega \mapsto \Phi_{s,r}^x(y,\omega)$ is $(\mathcal{F}_r \vee \mathcal{G}_u)$-measurable, for $F \in \mathcal{U}_v^u$,

$$\begin{aligned}(\mathcal{Q}_{s,r}^x \mathcal{Q}_{r,v}^x F)(y) &= \mathbf{E}\big((\mathcal{Q}_{r,v}^x F)(\cdot, \Phi_{s,r}^x(y)) | \mathcal{F}_s \vee \mathcal{G}_u\big) \\ &= \mathbf{E}\big(\mathbf{E}\big(F\big(\cdot, \Phi_{r,v}^x(\Phi_{s,r}^x(y))\big) | \mathcal{F}_r \vee \mathcal{G}_u\big) | \mathcal{F}_s \vee \mathcal{G}_u\big) \\ &= \mathbf{E}\big(F\big(\cdot, \Phi_{s,v}^x(y)\big) | \mathcal{F}_s \vee \mathcal{G}_u\big) = (\mathcal{Q}_{s,v}^x F)(y),\end{aligned}$$

as required. Here, the fact that $\omega \mapsto \Phi_{s,r}^x(y,\omega)$ is $(\mathcal{F}_r \vee \mathcal{G}_u)$-measurable was used in order to go from the first to the second line. This is a particular instance of the fact that if $x \mapsto F(x,\omega)$ is continuous in x for almost every ω and $Y : \Omega \to \mathcal{X}$ is a $\mathcal{G}$-measurable random variable for some sub-σ-algebra $\mathcal{G}$, then the identity

$$\mathbf{E}\big(F(x,\cdot)|\mathcal{G}\big)|_{x=Y} = \mathbf{E}\big(F\big(Y(\cdot),\cdot\big)|\mathcal{G}\big)$$

holds almost surely. We have also used the fact that $\Phi_{s,\cdot}^x$ has a version which is continuous in time and in the initial value. □

LEMMA 4.24. *The following estimate holds uniformly over all $s < v \le u$ and all $F \in \mathcal{U}_v^u$:*

$$|\hat{\mathcal{P}}_{s,v}^{x_s} F - \mathcal{Q}_{s,v}^x F|_\infty \lesssim \sqrt{\mathbf{E}(|x|_\alpha^2 | \mathcal{F}_s \vee \mathcal{G}_u)}|v-s|^\alpha |F|_{\mathrm{Lip}}. \tag{4.32}$$

PROOF. We know that (4.32) holds for $|v-s| \le \varepsilon$ from Lemma 4.22 with $\bar{x} = x_s$ since one has the bound

$$\mathbf{E}(|x_{r'} - x_s|^2 | \mathcal{F}_r \vee \mathcal{G}_u) \lesssim \mathbf{E}(|x|_\alpha^2 | \mathcal{F}_s \vee \mathcal{G}_u)|u-s|^{2\alpha}, \tag{4.33}$$

uniformly over $r \in [s,u]$ and $r' \in [r,u]$. We also know from Lemma 4.11 that $|\mathcal{P}_{v-s}^{x_s} F|_{\mathrm{Lip}}$ is bounded by a constant multiple of $|F|_{\mathrm{Lip}}$, uniformly in time.

We then consider a partition Δ of $[s,v]$ into subintervals of size between $\varepsilon/2$ and ε, and we write the difference of the two semigroups as a telescopic sum, then apply consecutively the following: the triangle inequalities, the contraction property of $\mathcal{Q}_{s,r}^x$, estimate (4.31) combined with (4.33) and Lemma 4.11. We also use the quasi semigroup property of $\mathcal{Q}^x$. This yields

$$\begin{aligned}|\hat{\mathcal{P}}_{s,v}^{x_s} F - \mathcal{Q}_{s,v}^x F|_\infty &\le \sum_{[r,r'] \in \Delta} |\mathcal{Q}_{s,r}^x (\hat{\mathcal{P}}_{r,r'}^{x_s} - \mathcal{Q}_{r,r'}^x) \mathcal{P}_{v-r'}^{x_s} F|_\infty \\ &\le \sum_{[r,r'] \in \Delta} |(\hat{\mathcal{P}}_{r,r'}^{x_s} - \mathcal{Q}_{r,r'}^x) \mathcal{P}_{v-r'}^{x_s} F|_\infty \\ &\lesssim \sum_{[r,r'] \in \Delta} \sqrt{\mathbf{E}(|x|_\alpha^2 | \mathcal{F}_s \vee \mathcal{G}_u)}|s-r'|^\alpha |\mathcal{P}_{v-r'}^{x_s} F|_{\mathrm{Lip}}\end{aligned}$$

$$\lesssim \sum_{[r,r']\in\Delta} \sqrt{\mathbf{E}(|x|_\alpha^2|\mathcal{F}_s\vee\mathcal{G}_u)}|s-v|^\alpha e^{-c|v-r'|/\varepsilon}|F|_{\mathrm{Lip}}$$
$$\lesssim \sqrt{\mathbf{E}(|x|_\alpha^2|\mathcal{F}_s\vee\mathcal{G}_u)}|v-s|^\alpha|F|_{\mathrm{Lip}},$$

as required. Note that there exists C such that for any ε, $\sum_{[r,r']\in\Delta} e^{-c|v-r'|/\varepsilon}\le C$ provided the size of the partition is of order ε. □

4.8. *Proof of the main result.* We now have all the ingredients in place for the proof of Theorem 4.3.

OF THEOREM 4.3. By Lemma 4.15, we know that $\bar f$ and $\bar g$ belong to BC^2. This implies that there exists a unique solution $\bar x_t$ to

$$d\bar x_t = \bar f(\bar x_t)\,dB_t + \bar g(\bar x_t)\,dt,\quad \bar x_0 = x_0,$$

where the integral against B is interpreted pathwise as a Young integral; see for example [27]. We apply Theorem 4.16 with $h = f - \bar f$, yielding the bound

$$\left\| \int_0^\cdot \big(f(x_r,y_r^\varepsilon) - \bar f(x_r)\big)\,dB_r\right\|_{\beta,p} \lesssim \varepsilon^\kappa\big(1+\|x\|_{\alpha,p}\big),$$

uniformly over x and over $\varepsilon\in(0,1]$, where y^ε is obtained from x by solving (4.1b). Here, $\kappa>0$ is small enough and $\alpha<\frac12$ and p are such that $\alpha+H>1+\frac1p$. Since $\sup_\varepsilon \|x_s^\varepsilon\|_{\alpha,p}<\infty$ by Corollary 4.20, we conclude that

$$\left\| \int_0^\cdot \big(f(x_s^\varepsilon,y_s^\varepsilon) - \bar f(x_s^\varepsilon)\big)\,dB_s\right\|_{\beta,p} \lesssim \varepsilon^\kappa,$$

and a similar bound holds for $\|\int_0^t (g(x_s^\varepsilon,y_s^\varepsilon)-\bar g(x_s^\varepsilon))\,ds\|_{\beta,p}$. Setting

$$\bar x_t^\varepsilon \stackrel{\mathrm{def}}{=} x_0 + \int_0^t \bar f(x_s^\varepsilon)\,dB_s + \int_0^t \bar g(x_s^\varepsilon)\,ds,$$

we have just shown that the processes x_t^ε and $\bar x_t^\varepsilon$ are close in $\mathcal{B}_{\beta,p}$:

$$\|\bar x^\varepsilon - x^\varepsilon\|_{\beta,p}\lesssim \varepsilon^\kappa. \tag{4.34}$$

It remains to show that $\bar x_t^\varepsilon$ and $\bar x_t$ are close in the β-Hölder norm, for which we begin by obtaining a pathwise estimates on their β-Hölder norm. Writing

$$x_t^\varepsilon = x_t^\varepsilon - \bar x_t^\varepsilon + x_0 + \int_0^t \bar f(x_s^\varepsilon)\,dB_s + \int_0^t \bar g(x_s^\varepsilon)\,ds,$$

we may apply Lemma 2.2 to compare x_t^ε and $\bar x_t$, where we take $F=(\bar f,\bar g)$ and $b_t = (B_t(\omega),t)$, $Z_0=x_0$ and $\bar Z_0 = x_0 + x_t^\varepsilon - \bar x_t^\varepsilon$. We have the pathwise estimate:

$$|x_t^\varepsilon - \bar x_t|_\beta \lesssim \exp\big(C|B|_\beta^{1/\beta} + C + C|\bar x^\varepsilon - x^\varepsilon|_\beta^{1/\beta}\big)|\bar x^\varepsilon - x^\varepsilon|_\beta.$$

(Note that the β-Hölder seminorm of the constant x_0 vanishes.) Since (modulo changing β slightly), we already know from (4.34) that $|\bar x^\varepsilon - x^\varepsilon|_\beta\to 0$ in probability at rate ε^κ. This concludes the proof. □

APPENDIX: ESTIMATES ON CONDITIONED FBM

The purpose of this Appendix is to provide a proof of Lemma 3.3, as well as to provide an explicit representation of R used in Lemma 3.2. For this, we first derive a suitable representation for the mixed derivative of the covariance function R of $\tilde{B}$.

LEMMA A.1. *Let R be as above, $c_1 = (H - \frac{1}{2}), c_3 = (H - \frac{1}{2})(H - \frac{3}{2})$ and $c_2 = -c_3 \int_0^\infty u^{H-\frac{1}{2}}(1+u)^{H-\frac{5}{2}}\,du$. Then for $r < s$, one has the identity*

$$\partial^2_{r,s} R(r,s) = c_1 r^{H-\frac{1}{2}} s^{H-\frac{3}{2}} + c_2 (s-r)^{2H-2} + c_3 \int_r^\infty v^{H-\frac{1}{2}}(s-r+v)^{H-\frac{5}{2}}\,dv. \tag{A.1}$$

PROOF. Recall that one has the identity

$$G(t) = (2H-1)\hat{R}'(t) - (t+1)\hat{R}''(t), \tag{A.2}$$

with

$$\hat{R}(t) = \int_0^1 (1-s)^{H-\frac{1}{2}}(1+t-s)^{H-\frac{1}{2}}\,ds.$$

We have

$$\begin{aligned} F'(t) &= \left(H - \frac{1}{2}\right) \int_0^1 (1-s)^{H-\frac{1}{2}}(1+t-s)^{H-\frac{3}{2}}\,ds \\ &= \left(H - \frac{1}{2}\right) t^{2H-1} \int_0^{1/t} u^{H-\frac{1}{2}}(1+u)^{H-\frac{3}{2}}\,du, \end{aligned} \tag{A.3}$$

where we used the change of variables $s = 1 + tu$. Differentiating the second line of (A.3) immediately gives

$$t\hat{R}''(t) = (2H-1)\hat{R}'(t) - c_1(1+t)^{H-\frac{3}{2}}. \tag{A.4}$$

On the other hand, differentiating the first line of (A.3), we obtain

$$\begin{aligned} \hat{R}''(t) &= \left(H - \frac{1}{2}\right)\left(H - \frac{3}{2}\right) \int_0^1 (1-s)^{H-\frac{1}{2}}(1+t-s)^{H-\frac{5}{2}}\,ds \\ &= c_3 \int_0^1 u^{H-\frac{1}{2}}(t+u)^{H-\frac{5}{2}}\,du \\ &= c_3 \int_0^\infty u^{H-\frac{1}{2}}(t+u)^{H-\frac{5}{2}}\,du - c_3 \int_1^\infty u^{H-\frac{1}{2}}(t+u)^{H-\frac{5}{2}}\,du \\ &= c_3 t^{2H-2} \int_0^\infty u^{H-\frac{1}{2}}(1+u)^{H-\frac{5}{2}}\,du - c_3 \int_1^\infty u^{H-\frac{1}{2}}(t+u)^{H-\frac{5}{2}}\,du. \end{aligned}$$

Substituting (A.4) into (A.2), we can then rewrite G as $G(t) = c_1(1+t)^{H-\frac{3}{2}} - \hat{R}''(t)$, so that for $c_2 = -c_3 \int_0^\infty u^{H-\frac{1}{2}}(1+u)^{H-\frac{5}{2}}\,du$,

$$G(t) = c_1(1+t)^{H-\frac{3}{2}} + c_2 t^{2H-2} + c_3 \int_1^\infty u^{H-\frac{1}{2}}(t+u)^{H-\frac{5}{2}}\,du,$$

and the claim follows by substituting this into (3.3). □

OF LEMMA 3.3. It follows from the fact that a conditional variance is always smaller than the full variance that

$$|h|^2_{\mathrm{RKHS}} \le C\left|\int_0^T \int_0^T |r-s|^{2H-2} h(r)h(s)\,dr\,ds\right| =: 2C|I|.$$

By homogeneity, it suffices to consider the case $T = 1$ and, by symmetry, we can restrict the domain of integration to the region where $r \le s$. To bound I, we then note that we can find a constant c such that

$$\begin{aligned} I &= \int_0^1 \int_0^s (s-r)^{2H-2} h(r)\,dr\,h(s)\,ds \\ &= c\int_0^1 \int_0^s \int_r^s (s-u)^{H-\frac{3}{2}} (u-r)^{H-\frac{3}{2}}\,du\,h(r)\,dr\,h(s)\,ds \\ &= c\int_0^1 \int_u^1 (s-u)^{H-\frac{3}{2}} h(s)\,ds \int_0^u (u-r)^{H-\frac{3}{2}} h(r)\,dr\,du. \end{aligned}$$

Performing one integration by parts, we note that

$$\begin{aligned} \left|\int_0^u (u-r)^{H-\frac{3}{2}} h(r)\,dr\right| &= \left|\left(H-\frac{3}{2}\right)\int_0^u (u-r)^{H-\frac{5}{2}} \big(\hat h(u) - \hat h(r)\big)\,dr\right| \\ &\lesssim |h|_{-\kappa} u^{H-\frac{1}{2}-\kappa}, \end{aligned}$$

where we used the fact that $\kappa < H - \frac{1}{2}$ to guarantee integrability at $r = u$ and $\hat h$ denotes a primitive of h. The other factor is bounded in the same way, so that

$$|I| \lesssim |h|^2_{-\kappa} \int_0^1 (1-u)^{H-\frac{1}{2}-\kappa} u^{H-\frac{1}{2}-\kappa}\,du \lesssim |h|^2_{-\kappa},$$

as required. □

Acknowledgments. We would like to thank Massimiliano Gubinelli for an inspiring conversation.

The first author was supported by the Leverhulme trust via a Leadership Award and the ERC via the consolidator grant 615897:CRITICAL.

REFERENCES

[1] ARONSON, D. G. (1967). Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.* **73** 890–896. MR0217444 https://doi.org/10.1090/S0002-9904-1967-11830-5

[2] BAUDOIN, F. and HAIRER, M. (2007). A version of Hörmander's theorem for the fractional Brownian motion. *Probab. Theory Related Fields* **139** 373–395. MR2322701 https://doi.org/10.1007/s00440-006-0035-0

[3] BOGACHEV, V. I. (1998). *Gaussian Measures. Mathematical Surveys and Monographs* **62**. Amer. Math. Soc., Providence, RI. MR1642391 https://doi.org/10.1090/surv/062

[4] CHEEGER, J. and YAU, S. T. (1981). A lower bound for the heat kernel. *Comm. Pure Appl. Math.* **34** 465–480.

[5] DA SILVA, J. L., ERRAOUI, M. and ESSAKY, E. H. (2018). Mixed stochastic differential equations: Existence and uniqueness result. *J. Theoret. Probab.* **31**. 1119–1141.

[6] ELWORTHY, K. D. and LI, X.-M. (1994). Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** 252–286. MR1297021 https://doi.org/10.1006/jfan.1994.1124

[7] FREIDLIN, M. I. and WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* **260**. Springer, New York. Translated from the 1979 Russian original by Joseph Szücs. MR1652127 https://doi.org/10.1007/978-1-4612-0611-8

[8] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths: With an Introduction to Regularity Structures. Universitext*. Springer, Cham. MR3289027 https://doi.org/10.1007/978-3-319-08332-2

[9] GOEBEL, M. and SACHWEH, F. (1999). On the autonomous Nemytskij operator in Hölder spaces. *Z. Anal. Anwend.* **18** 205–229. MR1701350 https://doi.org/10.4171/ZAA/878

[10] GUBINELLI, M. (2004). Controlling rough paths. *J. Funct. Anal.* **216** 86–140. https://doi.org/10.1016/j.jfa.2004.01.002

[11] GUERRA, J. and NUALART, D. (2008). Stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *Stoch. Anal. Appl.* **26** 1053–1075. MR2440915 https://doi.org/10.1080/07362990802286483

[12] HAIRER, M. (2005). Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.* **33** 703–758. https://doi.org/10.1214/009117904000000892

[13] HAIRER, M. (2014). A theory of regularity structures. *Invent. Math.* **198** 269–504.

[14] HAIRER, M. and PILLAI, N. S. (2011). Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 601–628. MR2814425 https://doi.org/10.1214/10-AIHP377

[15] HAS'MINSKII, R. Z. (1968). On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)* **4** 260–279. MR0260052

[16] HUANG, S. T. and CAMBANIS, S. (1978). Stochastic and multiple Wiener integrals for Gaussian processes. *Ann. Probab.* **6** 585–614. MR0496408

[17] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Teor. Veroyatn. Primen.* **7** 361–392. MR0148125

[18] KATO, T. (1976). *Perturbation Theory for Linear Operators*, 2nd ed. Springer, Berlin. MR0407617

[19] KIFER, Y. (1988). *Random Perturbations of Dynamical Systems. Progress in Probability and Statistics* **16**. Birkhäuser, Inc., Boston, MA. MR1015933 https://doi.org/10.1007/978-1-4615-8181-9

[20] KOMOROWSKI, T., LANDIM, C. and OLLA, S. (2012). *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation. Grundlehren der Mathematischen Wissenschaften* **345**. Springer, Heidelberg. MR2952852 https://doi.org/10.1007/978-3-642-29880-6

[21] KUBILIUS, K. (2002). The existence and uniqueness of the solution of an integral equation driven by a p-semimartingale of special type. *Stochastic Process. Appl.* **98** 289–315. MR1887537 https://doi.org/10.1016/S0304-4149(01)00145-4

[22] KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics* **24**. Cambridge Univ. Press, Cambridge. MR1070361

[23] LÊ, K. (2018). A stochastic sewing lemma and applications. Available at arXiv:1810.10500.

[24] LI, X.-M. (1994). Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds. *Probab. Theory Related Fields* **100** 485–511. MR1305784 https://doi.org/10.1007/BF01268991

[25] LI, X.-M. (2008). An averaging principle for a completely integrable stochastic Hamiltonian system. *Nonlinearity* **21** 803–822. MR2399826 https://doi.org/10.1088/0951-7715/21/4/008

[26] LI, X.-M. (2018). Perturbation of conservation laws and averaging on manifolds. In *Computation and Combinatorics in Dynamics, Stochastics and Control, the Abel Symposium, Rosendal, Norway, August* 2016, Norway.

[27] LYONS, T. (1994). Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.* **1** 451–464. MR1302388 https://doi.org/10.4310/MRL.1994.v1.n4.a5

[28] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. MR1654527 https://doi.org/10.4171/RMI/240

[29] MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR0242239 https://doi.org/10.1137/1010093

[30] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Springer, Berlin. MR2200233

[31] PAVLIOTIS, G. A. and STUART, A. M. (2008). *Multiscale Methods: Averaging and Homogenization. Texts in Applied Mathematics* **53**. Springer, New York. MR2382139

[32] QIAN, Z. and ZHENG, W. (2004). A representation formula for transition probability densities of diffusions and applications. *Stochastic Process. Appl.* **111** 57–76.

[33] SKOROKHOD, A. V., HOPPENSTEADT, F. C. and SALEHI, H. (2002). *Random Perturbation Methods with Applications in Science and Engineering. Applied Mathematical Sciences* **150**. Springer, New York. MR1912425 https://doi.org/10.1007/b98905

[34] WANG, F.-Y. (1997). Sharp explicit lower bounds of heat kernels. *Ann. Probab.* **25** 1995–2006.

[35] YOUNG, L. C. (1936). An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67** 251–282.