SLOW-FAST SYSTEMS WITH FRACTIONAL ENVIRONMENT AND DYNAMICS

BY XUE-MEI LI[a] AND JULIAN SIEBER[b]

Department of Mathematics, Imperial College London, [a]*xue-mei.li@imperial.ac.uk,* [b]*j.sieber19@imperial.ac.uk*

We prove a fractional averaging principle for interacting slow-fast systems. The mode of convergence is in Hölder norm in probability. The main technical result is a quenched ergodic theorem on the conditioned fractional dynamics. We also establish geometric ergodicity for a class of fractional-driven stochastic differential equations, improving a recent result of Panloup and Richard.

1. Introduction and main results. We study slow-fast systems driven by fractional Brownian motions (fBm):

$$(1.1) \qquad dX_t^\varepsilon = f(X_t^\varepsilon, Y_t^\varepsilon)\, dt + g(X_t^\varepsilon, Y_t^\varepsilon)\, dB_t, \quad X_0^\varepsilon = X_0,$$

$$(1.2) \qquad dY_t^\varepsilon = \frac{1}{\varepsilon} b(X_t^\varepsilon, Y_t^\varepsilon)\, dt + \frac{1}{\varepsilon^{\hat{H}}} \sigma\, d\hat{B}_t, \quad Y_0^\varepsilon = Y_0,$$

where B and $\hat{B}$ are independent fBms on an underlying complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with Hurst parameters $H \in (\frac{1}{2}, 1)$ and $\hat{H} \in (1 - H, 1)$, respectively. Here, $g : \mathbb{R}^d \times \mathbb{R}^n \to L(\mathbb{R}^m, \mathbb{R}^d)$ and $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ is nondegenerate. As the scale parameter $\varepsilon > 0$ is taken to 0, one hopes that the *slow motion* X^ε is well approximated by an *effective dynamics* $\bar{X}$. For $H = \hat{H} = \frac{1}{2}$, this convergence has been studied by myriad authors since the seminal works of Bogolyubov–Mitropol'skiĭ [7] and Hasminskii [26]; see for example, the monographs and survey articles [5, 18, 32, 33, 40, 48] and references therein for a comprehensive overview. It is still a very active research area [34, 45, 46].

For $H, \hat{H} \neq \frac{1}{2}$, the SDEs (1.1)–(1.2) provide a suitable model for economic, medical, and climate phenomena exhibiting a genuinely non-Markovian behavior in both the system and its environment. It is for example, very well known that neglecting temporal memory effects in climate modeling by resorting to a diffusion model results in prediction notoriously mismatching observational data [2, 4, 13, 28]. It thus became widely popular to use fBm in climate modeling [16, 49, 51].

While slow-fast systems with fractional noise have seen a tremendous spike of interest in the last two years [8, 9, 23, 25, 41, 42], all of these works resort to Markovian, strongly mixing fast processes by choosing $\hat{H} = \frac{1}{2}$ in (1.2). The main contribution of this article is to establish the convergence $X^\varepsilon \to \bar{X}$ even for a *non-Markovian* fast dynamics by allowing $\hat{H} \neq \frac{1}{2}$. It hardly comes as a surprise that this renders the analysis much more delicate and it is not clear at all if an averaging principle can even hold for a fractional, *nonmixing* environment. In fact, the usual assumption in the aforementioned works on Markovian averaging principles is a strong mixing condition with an algebraic rate [1, 27]. This condition is essentially never satisfied for a fractional dynamics [3].

Recent work of Hairer and the first author of this article suggests the following ansatz for the effective dynamics:

$$(1.3) \qquad d\bar{X}_t = \bar{f}(\bar{X}_t)\, dt + \bar{g}(\bar{X}_t)\, dB_t, \quad \bar{X}_0 = X_0,$$

where $\bar{f}(x) \triangleq \int f(x, y)\, \pi^x(dy)$ and similar for $\bar{g}$ [23]. For $\hat{H} = \frac{1}{2}$, this work showed that the average is taken with respect to the unique invariant π^x of the fast dynamics with *frozen* slow input

$$(1.4) \qquad dY_t^x = b(x, Y_t^x)\, dt + \sigma\, d\hat{B}_t.$$

For $\hat{H} \neq \frac{1}{2}$, it is a priori not clear what π^x should be. We show that it is the one-time marginal of the unique stationary path space law $\mathbb{P}_{\pi^x} \in \mathcal{P}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^n))$; see Section 2.1 for details. [Here and in the sequel, $\mathcal{P}(\mathcal{X})$ denotes the set of Borel probability measures on a Polish space $\mathcal{X}$.]

In addition to standard regularity requirements ensuring well-posedness of the slow-fast system (see Condition 3.1 below), we shall impose a contractivity condition on the drift in (1.2).

DEFINITION 1.1. Let $\lambda, R \geq 0$ and $\kappa > 0$. We write $S(\kappa, R, \lambda)$ for the set of Lipschitz continuous functions $b : \mathbb{R}^n \to \mathbb{R}^n$ satisfying

$$(1.5) \qquad \langle b(x) - b(y), x - y \rangle \leq \begin{cases} -\kappa |x - y|^2 & |x|, |y| \geq R, \\ \lambda |x - y|^2 & \text{otherwise.} \end{cases}$$

Note that λ may be smaller than $|b|_{\text{Lip}}$, whence its prescription is not necessarily redundant. If $b = -\nabla V$ is a gradient vector field with potential V, then (1.5) is equivalent to V being at most λ-concave on $|x| < R$ and κ-convex on $|x| \geq R$. If $V \in \mathcal{C}^2(\mathbb{R}^n)$, these requirements are in turn equivalent to $\nabla^2 V \geq -\lambda$ and $\nabla^2 V \geq \kappa$ on the respective sets.

THEOREM 1.2 (Fractional averaging principle). *Consider the slow-fast system* (1.1)–(1.2). *Suppose that* $f, g \in \mathcal{C}_b^2$ *and* b *satisfies Condition* 3.1. *Let* $\alpha < H$ *and* $\kappa, R > 0$. *Then there is a number* $\lambda_0 > 0$ *such that, if* $b(x, \cdot) \in S(\kappa, R, \lambda_0)$ *for every* $x \in \mathbb{R}^d$, *all of the following hold*:

- *For every* $x \in \mathbb{R}^d$, *there exists a unique stationary path space law* $\mathbb{P}_{\pi^x} \in \mathcal{P}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^n))$ *for the frozen fast dynamics* (1.4).
- *Let* $\pi^x \in \mathcal{P}(\mathbb{R}^n)$ *be the one-time marginal of* $\mathbb{P}_{\pi^x}$. *If*

$$x \mapsto \bar{g}(x) \triangleq \int_{\mathbb{R}^n} g(x, y)\, \pi^x(dy) \in \mathcal{C}_b^2\big(\mathbb{R}^d, L(\mathbb{R}^m, \mathbb{R}^d)\big),$$

then there is a unique pathwise solution to (1.3) *and* $X^\varepsilon \to \bar{X}$ *as* $\varepsilon \to 0$ *in* $\mathcal{C}^\alpha([0, T], \mathbb{R}^d)$ *in probability for any* $T > 0$.

The regularity of $\bar{g}$ not only hinges on the regularity of g but also on the fast dynamics. First we note that the requirement on $\bar{g}$ clearly holds for a diffusion coefficient depending only on the slow motion X^ε:

$$dX_t^\varepsilon = f(X_t^\varepsilon, Y_t^\varepsilon)\, dt + g(X_t^\varepsilon)\, dB_t.$$

Another class of examples is provided by Corollary 4.25 below.

The technical core of the proof of Theorem 1.2 is a quantitative *quenched ergodic theorem* on the conditional evolution of the process (1.4). We prove this by means of a control argument, which is of independent interest. In fact, it allows us to improve recent work of

Panloup and Richard [39] by establishing geometric ergodicity for a class of SDEs driven by additive fractional noise. To our best knowledge, this is the first result achieving an exponential convergence rate for a fractional dynamics (excluding the trivial instance of an everywhere contractive drift).

Let $\|\mu\|_{\mathrm{TV}} \triangleq \sup_A |\mu(A)|$ denote the total variation norm of μ, $\mathcal{W}^p$ the p-Wasserstein distance, and $\mathbb{W}^p$ the Wasserstein-like metric for generalized initial conditions introduced in Definition 2.5.

THEOREM 1.3 (Geometric ergodic theorem). *Let $(Y_t)_{t\geq 0}$ be the solution to the SDE*

$$dY_t = b(Y_t)\, dt + \sigma\, dB_t \tag{1.6}$$

started in the generalized initial condition μ, where $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ is nondegenerate and B is an fBm with Hurst parameter $H \in (0,1)$. Then, for any $p \geq 1$ and any $\kappa, R > 0$, there exists a $\Lambda = \Lambda(\kappa, R, p) > 0$ such that, whenever $b \in \mathrm{S}(\kappa, R, \Lambda)$, there is a unique invariant measure $\mathcal{I}_\pi$ for (1.6) in the sense of Definition 2.1. Moreover,

$$\mathcal{W}^p(\mathcal{L}(Y_t), \pi) \leq C e^{-ct} \mathbb{W}^p(\mu, \mathcal{I}_\pi) \quad \forall t \geq 0 \tag{1.7}$$

and

$$\|\mathcal{L}(Y_{\cdot+t}) - \mathbb{P}_\pi\|_{\mathrm{TV}} \leq C e^{-ct} \mathbb{W}^1(\mu, \mathcal{I}_\pi) \quad \forall t \geq 0, \tag{1.8}$$

where $c, C > 0$ are numerical constants independent of $t \geq 0$ and μ.

The work [21] already contains a result on the rate of convergence. There, the author assumed an off-diagonal contraction condition (see Condition 2.6 below) and obtained an algebraic rate in (1.8). Very recently Panloup and Richard [39] studied $b \in \mathrm{S}(\kappa, R, 0)$ for which they found a rate of order e^{-Dt^γ} for some $\gamma < \frac{2}{3}$ in both (1.7) and (1.8). Albeit these works did not require a global Lipschitz condition on the drift for Hurst parameters $H < \frac{1}{2}$, we emphasize that they do impose this assumption for $H > \frac{1}{2}$ to obtain (1.8). This is due to the lack of regularity of a certain fractional integral operator. Theorem 1.3 thus provides a genuine ramification of the results of [39] in the latter case. We note that similarly to the work of Panloup and Richard, the Wasserstein decay (1.7) also holds for more general Gaussian driving noises with stationary increments. We shall briefly comment on this in Section 3.5.

With the spiking interest in numerical methods based on the generalized Langevin equation with memory kernel [10, 30], Theorem 1.3 and the quenched quantitative ergodic theorem underpinning it can give a better theoretical understanding. A first step would be to derive quantitative estimates on the constants c, C, and Λ; a possible pathway is outlined in Remark 3.25 below. It is an interesting open question if there is indeed a finite threshold value of Λ beyond which the exponential rates (1.7)–(1.8) no longer hold. As established by Eberle, such a transition from exponential to subexponential rates does not happen in case $H = \frac{1}{2}$ [15].

EXAMPLE 1.4. Let us give an example of a drift not covered by the subexponential convergence theorems of [39]. Consider the double-well potential

$$V(x) = \alpha |x|^4 - \beta |x|^2$$

for $\alpha, \beta > 0$. We modify V outside of a compact such that its Hessian is bounded. Set $b = -\nabla V$. It is clear that $b \notin \bigcup_{\kappa, R > 0} \mathrm{S}(\kappa, R, 0)$ as soon as $\beta > 0$. However, for $\frac{\beta}{\alpha}$ sufficiently small, Theorem 1.3 furnishes an exponential rate of convergence.

Outline of the article. The next section features a brief overview of preliminary material. In Section 3, we prove the quantitative quenched ergodic theorem and deduce Theorem 1.3. The proof of Theorem 1.2 is concluded in Section 4.

2. Preliminaries. Recall that one-dimensional fractional Brownian motion with Hurst parameter $H \in (0, 1)$ is the centered Gaussian process $(B_t)_{t \geq 0}$ with

$$\mathbb{E}[(B_t - B_s)^2] = |t - s|^{2H}, \quad s, t \geq 0.$$

To construct d-dimensional fBm one lets the coordinates evolve as independent one-dimensional fBms with the same Hurst parameter. We will make frequent use of the following classical representation of one-dimensional fBm as a fractional integral of a two-sided Wiener process $(W_t)_{t \in \mathbb{R}}$, which is due to Mandelbrot and van Ness [37]:

$$(2.1) \quad B_t = \alpha_H \int_{-\infty}^0 (t - u)^{H - \frac{1}{2}} - (-u)^{H - \frac{1}{2}} \, dW_u + \alpha_H \int_0^t (t - u)^{H - \frac{1}{2}} \, dW_u, \quad t \geq 0.$$

Here, $\alpha_H > 0$ is some explicitly known normalization constant and we also declare $B_t = \bar{B}_t + \tilde{B}_t$.

2.1. *Invariant measures of fractional SDEs.* Albeit being non-Markovian on its own, the solution to (1.6) can actually be cast as the marginal of an infinite-dimensional Feller process $Z_t \triangleq (Y_t, (W_s)_{s \leq t})$ with values in $\mathbb{R}^n \times \mathcal{H}_H$. Here, W is the two-sided Wiener process driving the equation through (2.1) and $\mathcal{H}_H$ is a Hölder-type space of paths $\mathbb{R}_- \to \mathbb{R}^n$ supporting the Wiener measure W. More concretely, $\mathcal{H}_H$ is the closure of the space $\{f \in \mathcal{C}_c^\infty(\mathbb{R}_-, \mathbb{R}^n) : f(0) = 0\}$ in the norm

$$\|f\|_{\mathcal{H}_H} \triangleq \sup_{s,t \leq 0} \frac{|f(t) - f(s)|}{|t - s|^{\frac{1-H}{2}} \sqrt{1 + |t| + |s|}}.$$

To ensure that this construction actually furnishes a solution to (1.6), we of course have to assume that the law of the second marginal of Z coincides with W for each time $t \geq 0$. This motivates the following definition.

DEFINITION 2.1 ([21]). A measure $\mu \in \mathcal{P}(\mathbb{R}^n \times \mathcal{H}_H)$ with $\Pi_{\mathcal{H}_H}^* \mu = \mathsf{W}$ is called a *generalized initial condition*. A generalized initial condition $\mathcal{I}_\pi$, which is invariant for the Feller process Z is called an *invariant measure* for the SDE (1.6). We write $\pi \triangleq \Pi_{\mathbb{R}^n}^* \mathcal{I}_\pi$ for the first marginal and $\mathbb{P}_\pi \in \mathcal{P}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^n))$ for the law of the first coordinate of Z when started in $\mathcal{I}_\pi$.

By only adding the past of the driving noise to the auxiliary process Z, Hairer's framework rules out the existence of "unphysical" invariant measures, which frequently occur in the theory of random dynamical systems; see [22] for details.

There are only a few examples for which the invariant measure can be written down explicitly.

EXAMPLE 2.2. Let Y be the fractional Ornstein–Uhlenbeck process [11], that is,

$$dY_t = -Y_t \, dt + dB_t.$$

Then it is well known that its invariant measure is given by

$$\mathcal{I}_\pi(dy, dw) = \delta_{F(w)}(dy)\mathsf{W}(dw), \qquad F(w) \triangleq -\int_{-\infty}^0 e^s D_H w(s) \, ds,$$

where $D_H : \mathcal{H}_H \to \mathcal{H}_{1-H}$ is a continuous linear operator switching between Wiener and fBm paths' see [21], equation (3.6), for the precise definition. The first marginal of $\mathcal{I}_\pi$ and the stationary path space law are given by

$$\pi = \mathcal{L}\left(\int_{-\infty}^0 e^s \, dB_s\right) \quad \text{and} \quad \mathbb{P}_\pi = \mathcal{L}\left(\int_{-\infty}^t e^s \, dB_s\right)_{t \geq 0}.$$

REMARK 2.3. The invariant measure of (1.6) is in general not of product form.

Since $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ is nondegenerate, one can show that there is an isomorphism between the strictly stationary solutions to (1.6) and the set of invariant measures (provided one quotients the latter by the equivalence relation identifying generalized initial conditions which generate the same evolution in the first marginal). It is also not hard to prove the following.

PROPOSITION 2.4 ([21]). *If $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ and $b \in S(\kappa, R, \lambda)$ for some $\kappa > 0$, $R, \lambda \geq 0$, then there exists an invariant measure for* (1.6) *in the sense of Definition* 2.1. *Moreover, $\mathcal{I}_\pi$ has moments of all orders.*

The conclusion of Proposition 2.4 actually holds for a merely locally Lipschitz off-diagonal large scale contractive drift (see Condition 2.6 below). See also [14, 24] for versions for multiplicative noise. Finally, we introduce a Wasserstein-type distance for generalized initial conditions:

DEFINITION 2.5. Let μ and ν be generalized initial conditions. Let $\mathcal{C}_\Delta(\mu, \nu)$ denote the set of couplings of μ and ν concentrated on the diagonal $\Delta_{\mathcal{H}_H} \triangleq \{(w, w') \in \mathcal{H}_H^2 : w = w'\}$. For $p \geq 1$, we set

$$\mathbb{W}^p(\mu, \nu) \triangleq \inf_{\rho \in \mathcal{C}_\Delta(\mu, \nu)} \left(\int_{(\mathbb{R}^n \times \mathcal{H}_H)^2} |x - y|^p \, \rho(dx, dw, dy, dw')\right)^{\frac{1}{p}}.$$

Note that clearly $\mathcal{W}^p(\Pi_{\mathbb{R}^n}^* \mu, \Pi_{\mathbb{R}^n}^* \nu) \leq \mathbb{W}^p(\mu, \nu)$ and the inequality is strict in general.

2.2. *Large scale contractions.* Known ergodic theorems on (1.6) require either a Lyapunov-type stability or a large scale contractivity condition on the drift b. The former indicates that once far out, the solutions have the tendency to come back to a neighborhood of the origin. Under this condition, it is conceivable that two distinct solutions can come back from diverging routes, thus allowing to couple them. The Lyapunov stability condition was used in [14, 17] for multiplicative noise.

A large scale contraction on the other hand will force two solutions to come closer once they have left a ball B_R of sufficiently large radius $R > 0$. The following two conditions appeared in previous works.

CONDITION 2.6 (Off-diagonal large scale contraction, [21]). There exist numbers $\tilde{\kappa} > 0$ and $D, \lambda \geq 0$ such that

$$(2.2) \qquad \langle b(x) - b(y), x - y \rangle \leq \left(D - \tilde{\kappa}|x - y|^2\right) \wedge \left(\lambda|x - y|^2\right) \quad \forall x, y \in \mathbb{R}^n.$$

CONDITION 2.7 (Large scale contraction, [39]). There exist numbers $R \geq 0$ and $\kappa > 0$ such that

$$(2.3) \qquad \langle b(x) - b(y), x - y \rangle \leq -\kappa|x - y|^2 \quad \forall x, y \in \mathbb{R}^n \setminus B_R.$$

EXAMPLE 2.8.   The function $b(x) = x - x^3$ is a large scale contraction.

We will later use the following standard result, a slightly weaker version of which was proven in [39], Lemma 5.1.

LEMMA 2.9.   *If b is locally Lipschitz continuous and satisfies the large scale contraction condition* (2.3), *then for any $\bar{\kappa} \in (0, \kappa)$, there is an $\bar{R} > 0$ such that*

$$\langle b(x) - b(y), x - y \rangle \leq -\bar{\kappa}|x - y|^2 \quad \forall y \in \mathbb{R}^n, |x| > \bar{R}.$$

PROOF.   Since $\langle b(x) - b(y), x - y \rangle \leq -\kappa|x - y|^2$ for x and y outside of the ball B_R, we only need to show that the required contraction holds for any $|y| \leq R$ and $|x| > \bar{R}$. Fix such x and y.

Without loss of generality, we may also assume that $\bar{R} \geq R + 1$. Then there is a $\beta \in (0, 1)$ such that $z_\beta \triangleq (1 - \beta)x + \beta y$ has norm $|z_\beta| = R + 1$. Since $x - y = \frac{1}{\beta}(x - z_\beta)$ and, since x, z_β are outside of B_R,

$$\langle b(x) - b(z_\beta), x - y \rangle \leq -\frac{1}{\beta}\kappa|x - z_\beta|^2 = -\kappa\beta|x - y|^2.$$

Let $K \triangleq |b|_{\text{Lip}; B_{R+1}}$ denote the Lipschitz constant of b on B_{R+1}. Since $|z_\beta - y| = (1 - \beta)|x - y|$, it holds that

$$\langle b(x) - b(y), x - y \rangle = \langle b(x) - b(z_\beta), x - y \rangle + \langle b(z_\beta) - b(y), x - y \rangle$$

$$\leq -\kappa\beta|x - y|^2 + K(1 - \beta)|x - y|^2.$$

Since β is the length of the proportion of the line segment outside of B_{R+1}, we can choose it as close to 1 as we like by choosing $\bar{R}$ sufficiently large ($\beta = \frac{|x - z_\beta|}{|x - y|} \geq \frac{|x| - R - 1}{|x| + R} \geq \frac{\bar{R} - R - 1}{\bar{R} + R}$).   □

REMARK 2.10.   We make the following observations.

(i) Let $b : \mathbb{R}^n \to \mathbb{R}^n$ be a globally Lipschitz continuous function. Then the large scale contraction condition (2.3) is equivalent to $b \in \bigcup_{\lambda > 0} S(\kappa, R, \lambda)$. In view of Lemma 2.9, condition (1.5) also holds for a merely locally Lipschitz continuous b at the cost of a smaller contractive rate and a bigger contractive range. In fact, choose $\bar{\kappa} \in (0, \kappa)$ and let $\bar{R} > R$ be the corresponding radius furnished by Lemma 2.9. This gives (1.5) with $\kappa \rightsquigarrow \bar{\kappa}$, $R \rightsquigarrow \bar{R}$, and $\lambda \rightsquigarrow |b|_{\text{Lip}; B_{\bar{R}}}$.

(ii) The off-diagonal large scale contraction condition is weaker than the large scale contraction condition. With the former, there may be no $\kappa > 0$ such that (2.3) holds in the region $\{|x - y| \leq \frac{D}{2\bar{\kappa}}\} \cap \{|x| \geq R, |y| \geq R\}$. On the other hand, if (2.3) holds and b is locally Lipschitz continuous, we can choose any $\tilde{\kappa} < \kappa$. In fact, denoting the radius from Lemma 2.9 by $\bar{R} > 0$, one only needs to show (2.2) when both x and y are in $B_{\bar{R}}$. To this end, we pick $\lambda = |b|_{\text{Lip}; B_{\bar{R}}}$ and $D \geq \sup_{x, y \in B_{\bar{R}}}(\tilde{\kappa} + \lambda)|x - y|^2$.

3. The conditional evolution of fractional dynamics.   To derive strong L^p-bounds on the Hölder norm of the slow motion in Section 4 below, we need to study the conditional distribution of the evolution (1.4). Unlike the Markovian case, the conditioning changes the dynamics and the resulting evolution may *no longer* solve the original equation. We will show that, in the limit $t \to \infty$, the law of the conditioned dynamics still converges to π^x, the first marginal of the invariant measure for the fast dynamics with frozen slow input (1.4). The rate of convergence is however slower (only algebraic rather than exponential).

Let us first state the regularity assumption imposed in Theorem 1.2. For this we introduce a convenient notation, which we shall frequently use in the sequel. We write $a \lesssim b$ if there is a constant $C > 0$ such that $a \leq Cb$. The constant C is independent of any ambient parameters on which a and b may depend.

CONDITION 3.1. The drift $b : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies the following conditions:

- *Linear growth*:
$$|b(x, y)| \lesssim 1 + |x| + |y|, \quad \forall x \in \mathbb{R}^d, y \in \mathbb{R}^n.$$

- *Uniformly locally Lipschitz in the first argument*: For each $R > 0$, there is an $L_R > 0$ such that
$$\sup_{y \in \mathbb{R}^n} |b(x_1, y) - b(x_2, y)| \leq L_R |x_1 - x_2|, \quad \forall |x_1|, |x_2| \leq R.$$

- *Uniformly Lipschitz in the second argument*: There is an $L > 0$ such that
$$\sup_{x \in \mathbb{R}^d} |b(x, y_1) - b(x, y_2)| \leq L|y_1 - y_2|, \quad \forall y_1, y_2 \in \mathbb{R}^n.$$

Let $(\mathcal{F}_t)_{t \geq 0}$ be a complete filtration to which $\hat{B}$ is adapted. For any continuous, $(\mathcal{F}_t)_{t \geq 0}$-adapted, $\mathbb{R}^d$-valued process X with continuous sample paths, and any $\varepsilon > 0$, the equation

$$(3.1) \qquad d\Phi_t^X = \frac{1}{\varepsilon} b(X_t, \Phi_t^X) \, dt + \frac{1}{\varepsilon^{\hat{H}}} \sigma \, d\hat{B}_t, \quad \Phi_t^X = y,$$

has a unique global pathwise solution under Condition 3.1; see Lemma 4.2 below. The flow $\Phi_{s,t}^X(y)$ associated with (3.1) is therefore well defined. An important special case of (3.1) is when the extrinsic process is given by a fixed point $x \in \mathbb{R}^d$. For this we reserve the notation $\bar{\Phi}^x$:

$$(3.2) \qquad d\bar{\Phi}_t^x = \frac{1}{\varepsilon} b(x, \bar{\Phi}_t^x) \, dt + \frac{1}{\varepsilon^{\hat{H}}} \sigma \, d\hat{B}_t, \quad \bar{\Phi}_0^x = y.$$

We would like the reader to observe that the dependency of flows on the scale parameter $\varepsilon > 0$ is suppressed in our notation. Note that, by self-similarity, sending $\varepsilon \to 0$ in (3.2) is equivalent to keeping $\varepsilon = 1$ fixed and taking $t \to \infty$. As the ε-dependence of the flows (3.1)–(3.2) will play a key role in Section 4, we choose to introduce a new notation in case $\varepsilon = 1$, which is used throughout the rest of this section.

DEFINITION 3.2. Let $\mathfrak{h} \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n) \triangleq \{f \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n) : f(0) = 0\}$ and $x \in \mathbb{R}^d$. We denote the flow of the ordinary differential equation

$$(3.3) \qquad dy_t = b(x, y_t) \, dt + d\mathfrak{h}_t$$

by $\Psi_{s,t}^x(y, \mathfrak{h})$, where $y \in \mathbb{R}^n$ and $0 \leq s \leq t$. It is given by the solution to the integral equation

$$\Psi_{s,t}^x(y, \mathfrak{h}) = y + \int_s^t b(x, \Psi_{s,r}^x(y, \mathfrak{h})) \, dr + \mathfrak{h}_t - \mathfrak{h}_s.$$

We also use the abbreviation $\Psi_t^x \triangleq \Psi_{0,t}^x$.

Under Condition 3.1, (3.3) is well-posed and it follows that $\Psi_{s,t}^x(y, \mathfrak{h}) = \Psi_{t-s}^x(y, \theta_s \mathfrak{h})$ for each $0 \leq s \leq t$ and $y \in \mathbb{R}^n$, where $\theta_s f = f(\cdot + s) - f(\cdot)$ is the Wiener shift operator on the path space. If $x \in \mathbb{R}^d$, $y \in \mathbb{R}^n$, or $\mathfrak{h} \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$ are random, we understand Definition 3.2 pathwise for each fixed sample $\omega \in \Omega$. The solutions to (3.1) and (3.2) are also understood in this sense.

3.1. *Processes with a locally independent increment decomposition.* The derivation of the conditioned evolution relies on the following simple fact: For $t, h \geq 0$, we have

$$(3.4) \qquad (\theta_t \hat{B})_h = \hat{B}_{t+h} - \hat{B}_t = \bar{\hat{B}}_h^t + \tilde{\hat{B}}_h^t,$$

where, in a slight abuse of notation (the integrand has to be multiplied by the identity matrix),

$$\bar{\hat{B}}_h^t \triangleq \alpha_{\hat{H}} \int_{-\infty}^t \left((t+h-u)^{\hat{H}-\frac{1}{2}} - (t-u)^{\hat{H}-\frac{1}{2}} \right) d\hat{W}_u,$$

$$\tilde{\hat{B}}_h^t \triangleq \alpha_{\hat{H}} \int_t^{t+h} (t+h-u)^{\hat{H}-\frac{1}{2}} \, d\hat{W}_u.$$

This decomposition is easily obtained by rearranging (2.1). For any $t \geq 0$, the two components $\bar{\hat{B}}^t$ and $\tilde{\hat{B}}^t$ are independent. We call $\bar{\hat{B}}^t$ the *smooth* part of the increment, whereas $\tilde{\hat{B}}^t$ is referred to as the *rough* part. This terminology is based on the fact that, away from the origin, the process $\bar{\hat{B}}^t$ has continuously differentiable sample paths and therefore the "roughness" of $\hat{B}$ essentially comes from $\tilde{\hat{B}}^t$. Indeed, it is not hard to check that $\tilde{\hat{B}}^t$ is of precisely the same Hölder regularity as $\hat{B}$. We also observe that $\tilde{\hat{B}}^t \overset{d}{=} \tilde{\hat{B}}^0 \triangleq \tilde{B}$ for all $t > 0$.

The process $\tilde{\hat{B}}$ is—up to a prefactor—known as Riemann–Liouville process (or type-II fractional Brownian motion) and was initially studied by Lévy [31]. Its use in modelling was famously discouraged in [37] due to its overemphasis of the origin and the "regularized" process (2.1) was proposed instead. In fact as we shall see below, the lack of stationarity of the increments of $\tilde{\hat{B}}$ complicates the analysis of the conditioned evolution.

DEFINITION 3.3. Let $(\mathcal{F}_t)_{t\geq 0}$ be a complete filtration. An $(\mathcal{F}_t)_{t\geq 0}$-adapted stochastic process Z is said to have a *locally independent decomposition of its increments* with respect to $(\mathcal{F}_t)_{t\geq 0}$ if for any $t \geq 0$, there exists an increment decomposition of the form

$$(\theta_t Z)_h = \tilde{Z}_h^t + \bar{Z}_h^t, \quad h \geq 0,$$

where $\bar{Z}^t \in \mathcal{F}_t$ and $\tilde{Z}^t$ is independent of $\mathcal{F}_t$.

As seen in (3.4), an fBm $\hat{B}$ has a locally independent decomposition of its increments with respect to any filtration $(\mathcal{F}_t)_{t\geq 0}$ *compatible* with $\hat{B}$. By this we mean that $(\hat{W}_s)_{s\leq t} \in \mathcal{F}_t$ and $(\theta_t \hat{W}_s)_{s\geq t}$ is independent of $\mathcal{F}_t$ for any $t \geq 0$.

EXAMPLE 3.4. Let us give some further examples, which will become important later on:

(i) Let $(\hat{W}_t)_{t\geq 0}$ be a Wiener process and $\tilde{\hat{B}}_t \triangleq \alpha_{\hat{H}} \int_0^t (t-u)^{H-\frac{1}{2}} d\hat{W}_u$ be the Riemann–Liouville process. Then, for any $t \geq 0$ and $h \geq 0$,

$$(\theta_t \tilde{\hat{B}})_h = \alpha_{\hat{H}} \int_0^t \left((t+h-u)^{\hat{H}-\frac{1}{2}} - (t-u)^{\hat{H}-\frac{1}{2}} \right) d\hat{W}_u$$

$$(3.5) \qquad\qquad + \alpha_{\hat{H}} \int_t^{t+h} (t+h-u)^{\hat{H}-\frac{1}{2}} d\hat{W}_u$$

$$\triangleq Q_h^t + \tilde{\hat{B}}_h^t.$$

Thus, $\tilde{\hat{B}}$ admits a locally independent decomposition of its increments with respect to any filtration compatible with $\hat{B}$.

(ii) Another example, given in [19, 20], is the stationary fractional Ornstein–Uhlenbeck process $Z_t = \int_{-\infty}^t e^{-(t-s)} \, d\hat{B}_s$. More generally, it is clear that $Z_t = \int_{-\infty}^t \mathfrak{G}(s, t) \, d\hat{B}_s$ with a suitable kernel $\mathfrak{G}$ also has this property.

(iii) Albeit not being a direct instance of Definition 3.3, it is also interesting to observe a *fractal* property of $\hat{B}$: The smooth part of the increment has an independent decomposition as $\bar{\hat{B}}_h^t = P_h^t + Q_h^t$, where Q^t was defined in (3.5) and

$$P_h^t \triangleq \alpha_{\hat{H}} \int_{-\infty}^0 \left((t + h - u)^{\hat{H} - \frac{1}{2}} - (t - u)^{\hat{H} - \frac{1}{2}} \right) d\hat{W}_u.$$

Our argument for the quenched ergodic theorem will be based on a two step conditioning procedure making use of an explicit representation of the conditioned process. We state it for a general noise with locally independent increments.

LEMMA 3.5. *Let* $0 \leq s \leq t < t + h$ *and* $(Z_t)_{t \geq 0}$ *be a continuous stochastic process admitting a locally independent decomposition* $(\theta_t Z)_h = \tilde{Z}_h^t + \bar{Z}_h^t$ *with respect to* $(\mathcal{F}_t)_{t \geq 0}$. *Let* X *and* Y *be* $\mathcal{F}_t$-*measurable random variables. Then, for any* $F : \mathbb{R}^n \to \mathbb{R}$ *bounded measurable,*

$$\mathbb{E}\left[F\left(\Psi_{s,t+h}^X(Y, Z) \right) \mid \mathcal{F}_t \right] = \mathbb{E}\left[F\left(\Psi_h^x(y, \varsigma + \tilde{Z}^t) \right) \right]\big|_{\substack{x=X, \varsigma = \bar{Z}^t, \\ y = \Psi_{s,t}^X(Y, Z)}},$$

where Ψ *is defined in Definition* 3.2.

PROOF. This in an immediate consequence of the flow property of the equation (3.3) and standard properties of conditional expectations. $\square$

Coming back to the flow of the fast motion with frozen slow input (3.2), the following result is an easy consequence of Lemma 3.5.

LEMMA 3.6. *Let* $(\mathcal{F}_t)_{t \geq 0}$ *be a filtration which is compatible with* $\hat{B}$. *Fix* $0 \leq s \leq t < t + h$ *and let* X, Y *be* $\mathcal{F}_t$-*measurable random variables. Then, for any* $F : \mathbb{R}^n \to \mathbb{R}$ *bounded measurable,*

$$\mathbb{E}\left[F\left(\bar{\Phi}_{s,t+h}^X(Y) \right) \mid \mathcal{F}_t \right] = \mathbb{E}\left[F\left(\Psi_{\frac{h}{\varepsilon}}^x(y, \varsigma + \sigma \bar{\hat{B}}) \right) \right]\big|_{\substack{x=X, \varsigma = \varepsilon^{-\hat{H}} \sigma \bar{\hat{B}}_{\varepsilon .}^t, \\ y = \bar{\Phi}_{s,t}^X(Y)}},$$

where $\bar{\hat{B}}_{\varepsilon .}^t \triangleq (\bar{\hat{B}}_{\varepsilon h}^t)_{h \geq 0}$.

We now turn to the fine properties of the smooth part of the increment. For $\alpha > 0$ we define the set

$$(3.6) \quad \Omega_\alpha \triangleq \left\{ f \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n) \cap \mathcal{C}^2((0, \infty), \mathbb{R}^n) : \limsup_{t \to \infty} \left(t^\alpha |\dot{f}(t)| + t^{1+\alpha} |\ddot{f}(t)| \right) < \infty \right\}.$$

This space is equipped with the seminorm

$$\| f \|_{\Omega_\alpha} \triangleq \sup_{t \geq 1} t^\alpha |\dot{f}(t)| + \sup_{t \geq 1} t^{1+\alpha} |\ddot{f}(t)|.$$

We also set $\Omega_{\alpha-} \triangleq \bigcap_{\beta < \alpha} \Omega_\beta$. The motivation for this definition stems from the following lemma.

LEMMA 3.7. *Let* $\varepsilon > 0$ *and* $t \geq 0$. *Then* $\varepsilon^{-\hat{H}} \bar{\hat{B}}_{\varepsilon .}^t \stackrel{d}{=} \bar{\hat{B}}^t \stackrel{d}{=} \bar{\hat{B}} \in \Omega_{(1-\hat{H})-}$ *a.s. and* $\| \bar{\hat{B}} \|_{\Omega_\alpha} \in \bigcap_{p \geq 1} L^p$ *for any* $\alpha < 1 - \hat{H}$.

PROOF.   Let $\delta \in (0, 1 - \hat{H})$. It is enough to prove that there is a random variable $C > 0$ with moments of all orders such that

$$(3.7) \qquad |\dot{\hat{B}}_t| \leq \frac{C}{t^{1 - \hat{H} - \delta}}, \qquad |\ddot{\hat{B}}_t| \leq \frac{C}{t^{2 - \hat{H} - \delta}}$$

for all $t \geq 1$ on a set of probability one. This in turn easily follows from sample path properties of the standard Wiener process. First, we have that

$$\dot{\hat{B}}_t = \alpha_{\hat{H}} \left(\hat{H} - \frac{1}{2} \right) \int_{-\infty}^{0} (t - u)^{\hat{H} - \frac{3}{2}} \, dW_u = -\alpha_{\hat{H}} \left(\hat{H} - \frac{1}{2} \right) \left(\hat{H} - \frac{3}{2} \right) \int_{-\infty}^{0} (t - u)^{\hat{H} - \frac{5}{2}} W_u \, du$$

since $\lim_{u \to -\infty} (t - u)^{\hat{H} - \frac{3}{2}} W_u = 0$. Therefore,

$$|\dot{\hat{B}}_t| \lesssim \left(\sup_{-1 \leq s \leq 0} |W_s| \int_{-1}^{0} (t - u)^{\hat{H} - \frac{5}{2}} \, du + \sup_{s \leq -1} \frac{|W_s|}{(t - s)^{\frac{1}{2} + \delta}} \int_{-\infty}^{-1} (t - u)^{\hat{H} - 2 + \delta} \, du \right)$$

$$\leq C \left(t^{\hat{H} - \frac{5}{2}} + (t + 1)^{\hat{H} - 1 + \delta} \right).$$

The fact that C has moments of all order is an easy consequence of Fernique's theorem. In fact, the Wiener process defines a Gaussian measure on the separable Banach space

$$\mathcal{M}^{\frac{1}{2} + \delta} \triangleq \left\{ f \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n) : \|f\|_{\mathcal{M}^{\frac{1}{2} + \delta}} \triangleq \sup_{u \geq 0} \frac{|f(u)|}{(1 + u)^{\frac{1}{2} + \delta}} < \infty \right\}.$$

By Fernique's theorem, the random variable $\|W\|_{\mathcal{M}^{\frac{1}{2} + \delta}}$ has therefore Gaussian tails. The first estimate in (3.7) follows. The bound on $|\ddot{\hat{B}}_t|$ is similar. $\square$

3.2. *A universal control.*   Let $b \in S(\kappa, R, \lambda)$, $\varsigma \in \mathcal{C}_0([0, 1], \mathbb{R}^n)$, and $u \in L^\infty([0, 1], \mathbb{R}^n)$. Let us consider the following controlled ordinary differential equation:

$$(3.8) \qquad x^{\varsigma, u}(t) = x_0 + \int_0^t b\left(x^{\varsigma, u}(s)\right) ds + \varsigma(t) + \int_0^t u(s) \, ds, \quad t \in [0, 1].$$

We think of ς as an external "adversary" and of u as a control. Since b is Lipschitz continuous, it is standard that there is a unique global solution to (3.8). If $u \equiv 0$, we adopt the shorthand $x^\varsigma \triangleq x^{\varsigma, 0}$.

The aim of this section is to exhibit an $\eta \in (0, 1)$ as large as possible so that the following holds: Given $\bar{R} > 0$, there is an $M > 0$ such that, for any adversary $\varsigma \in \mathcal{C}_0([0, 1], \mathbb{R}^n)$ and any initial condition $x_0 \in \mathbb{R}^n$, we can find a control $u \in L^\infty([0, 1], \mathbb{R}^n)$ with $|u|_\infty \leq M$ ensuring that the occupation time of $x^{\varsigma, u}$ of the set $\mathbb{R}^n \setminus B_{\bar{R}}$ is at least η. It is important to emphasize that the sup-norm of the control $|u|_\infty$ may neither depend on the adversary ς nor on the initial condition x_0 (otherwise the construction of u essentially becomes trivial). We shall actually choose u as concatenation of the zero function and a *universal* control $\hat{u} \in L^\infty([0, N^{-1}], \mathbb{R}^n)$ for a sufficiently large, but universal, $N \in \mathbb{N}$.

We begin with a lemma:

LEMMA 3.8.   *There is a constant $C > 0$ independent of ς and u such that, for the solution of* (3.8),

$$\left| x^{\varsigma, u}(t) - x^\varsigma(t) \right|^2 \leq C \left(1 + |u|_\infty^2 \right) t$$

for all $t \in [0, 1]$.

PROOF. Since b is contractive on the large scale, there are constants $D, \tilde{\kappa} > 0$ such that

$$\langle b(x) - b(y), x - y \rangle \leq D - \tilde{\kappa}|x - y|^2$$

for all $x, y \in \mathbb{R}^n$; see Remark 2.10(ii). Define now $f(t) \triangleq e^{\tilde{\kappa}t}|x^{\varsigma,u}(t) - x^\varsigma(t)|^2$, then

$$f'(t) = \tilde{\kappa} f(t) + 2e^{\tilde{\kappa}t}\langle b(x^{\varsigma,u}(t)) - b(x^\varsigma(t)) + u(t), x^{\varsigma,u}(t) - x^\varsigma(t)\rangle \leq 2De^{\tilde{\kappa}} + \frac{|u(t)|^2}{\tilde{\kappa}}$$

for all $t \in [0, 1]$. Consequently, setting $C \triangleq \max(2D, \tilde{\kappa}^{-1})$, we have

$$|x^{\varsigma,u}(t) - x^\varsigma(t)|^2 \leq C \int_0^t e^{-\tilde{\kappa}(t-s)}(1 + |u(s)|^2)\, ds$$

and the lemma follows at once. □

For a piecewise constant function $u : [0, 1] \to \mathbb{R}^n$, let $\mathcal{D}_u \subset [0, 1]$ denote the finite set of discontinuities. We then have the following control result.

PROPOSITION 3.9. *Let $\eta < \frac{1}{2}$ and $\bar{R} > 0$. Then there is a value $M > 0$ such that the following holds true: For each $\varsigma \in \mathcal{C}_0([0, 1], \mathbb{R}^n)$ and each $x_0 \in \mathbb{R}^n$, we can find a piecewise constant control $u \in L^\infty([0, 1], \mathbb{R}^n)$ with $|u|_\infty + |\mathcal{D}_u| \leq M$ such that the occupation time of $x^{\varsigma,u}$ of the set $\mathbb{R}^n \setminus B_{\bar{R}}$ is greater than or equal to η.*

PROOF. We prove that there exist an integer N and a control $\hat{u} \in L^\infty([0, N^{-1}])$ with at most two constant pieces independent of both the initial condition x_0 and the adversary ς such that either

$$\mathrm{Leb}(\{t \in [0, N^{-1}] : |x^\varsigma(t)| > \bar{R}\}) \geq \frac{\eta}{N} \quad \text{or} \quad \mathrm{Leb}(\{t \in [0, N^{-1}] : |x^{\varsigma,\hat{u}}(t)| > \bar{R}\}) \geq \frac{\eta}{N}.$$

In the former case, we of course choose $u \equiv 0$, otherwise we let $u = \hat{u}$. By the flow property of well-posed ordinary differential equations, the solution to (3.8) restarted at time N^{-1} solves a similar equation (with new adversary $\tilde{\varsigma}(\cdot) = \theta_{N^{-1}}\varsigma \in \mathcal{C}_0([0, 1 - N^{-1}], \mathbb{R}^n)$ and initial condition $x^{\varsigma,u}(N^{-1})$). Upon constructing $\hat{u}$, we can thus easily deduce the proposition by iterating this construction.

Suppose that the time spent by uncontrolled solution $(x_t^\varsigma)_{t \in [0, N^{-1}]}$ in $\mathbb{R}^n \setminus B_{\bar{R}}$ is strictly less than $\frac{\eta}{N}$. We let $A_{x_0,\varsigma}$ be the set of times $t \in [0, N^{-1}]$ at which $|x^\varsigma(t)| \leq \bar{R}$. Note that $A_{x_0,\varsigma}$ is the union of a countable number of closed, disjoint intervals. By assumption, we have $\mathrm{Leb}(A_{x_0,\varsigma}) > (1 - \eta)N^{-1}$.

For $\delta \triangleq (2N)^{-1}$ and e any fixed unit vector, we define $\hat{u}$ to be the piecewise constant function

$$\hat{u}(t) = \begin{cases} \dfrac{2\bar{R} + 1}{(1 - 2\eta)\delta}e & t \in [0, \delta], \\ -\dfrac{2\bar{R} + 1}{(1 - 2\eta)\delta}e & t \in (\delta, 2\delta], \end{cases}$$

so that

$$\int_0^t \hat{u}(s)\, ds = \begin{cases} \dfrac{2\bar{R} + 1}{(1 - 2\eta)\delta}te & t \in [0, \delta], \\ \dfrac{2\bar{R} + 1}{(1 - 2\eta)\delta}(2\delta - t)e & t \in (\delta, 2\delta]. \end{cases}$$

We observe that

$$(3.9) \qquad |x^{\varsigma,\hat{u}}(t)| \geq \left|\int_0^t \hat{u}(s)\, ds\right| - |x^\varsigma(t)| - |b|_{\mathrm{Lip}}\int_0^t |x^{\varsigma,\hat{u}}(s) - x^\varsigma(s)|\, ds.$$

Moreover, owing to Lemma 3.8, we can bound

$$
\int_0^t \left| x^{\varsigma,\hat{u}}(s) - x^{\varsigma}(s) \right| ds \le \sqrt{C}\left(1 + |\hat{u}|_\infty\right) \int_0^{2\delta} \sqrt{s}\, ds
$$

(3.10)

$$
= \frac{2\sqrt{C}}{3N^{\frac{3}{2}}}\left(1 + \frac{2(2\bar{R}+1)N}{1-2\eta}\right) < |b|_{\mathrm{Lip}}^{-1},
$$

provided we choose the integer $N = N(C, \bar{R}, \eta, |b|_{\mathrm{Lip}})$ large enough. Define the set $B_{x_0,\varsigma} \triangleq A_{x_0,\varsigma} \cap [(1-2\eta)\delta, (1+2\eta)\delta]$. Combining (3.9) and (3.10), we then certainly have that $|x^{\varsigma,\hat{u}}(t)| > \bar{R}$ for all $t \in B_{x_0,\varsigma}$. Since

$$
\mathrm{Leb}(B_{x_0,\varsigma}) \ge \frac{(1-\eta)}{N} - 2(1-2\eta)\delta = \frac{\eta}{N}
$$

and $|\hat{u}|_\infty$ as well as $|\mathcal{D}_{\hat{u}}|$ only depend on N and $\bar{R}$, this finishes the proof. $\square$

We conclude our study of the deterministic controlled ODE (3.8) with the following stability result which is proven by a standard Grönwall argument.

LEMMA 3.10. *Let $x^{\varsigma,u}$ denote the solution to the controlled differential equation* (3.8) *with initial condition $x_0 \in \mathbb{R}^n$ and control $u \in L^\infty([0,1], \mathbb{R}^n)$. Then, for any $w \in \mathcal{C}_0([0,1], \mathbb{R}^n)$, we have the bound*

$$
\left| x^{\varsigma,u} - \tilde{x} \right|_\infty \le e^{|b|_{\mathrm{Lip}}} \left| \int_0^\cdot u(s)\, ds - w \right|_\infty,
$$

where $\tilde{x}$ is the unique solution to

$$
\tilde{x}(t) = x_0 + \int_0^t b\big(\tilde{x}(s)\big)\, ds + w(t) + \varsigma(t), \quad t \in [0,1].
$$

3.3. *Exponential stability of the conditional evolution.* We now turn to the conditional evolution of (3.2) derived in Lemma 3.6. For brevity, we drop the hat on the driving fBm throughout this and the next section. Remember that we have to study SDEs driven by a Riemann–Liouville process

$$
\tilde{B}_t \triangleq \alpha_H \int_0^t (t-u)^{H-\frac{1}{2}}\, dW_u,
$$

where $(W_t)_{t \ge 0}$ is a standard Wiener process. Recall from Definition 3.2 that, for $\varsigma \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$, $\Psi_{s,t}(\cdot, \varsigma + \sigma \tilde{B})$ denotes the solution flow to the equation

(3.11)
$$
dX_t = b(X_t)\, dt + d\varsigma_t + \sigma\, d\tilde{B}_t.
$$

For brevity, let us henceforth set $\Psi_{s,t}^{\varsigma}(\cdot) \triangleq \Psi_{s,t}(\cdot, \varsigma + \sigma \tilde{B})$.

We first prove that—starting from any two initial points—the laws of the solutions converge to each other with an exponential rate. This however does not yet imply the convergence of $\mathcal{L}(\Psi_t^{\varsigma}(x))$ to the first marginal of the invariant measure π of the equation $dX_t = b(X_t)\, dt + \sigma\, dB_t$ since, even if we choose $X_0 \sim \pi$, we have $\mathcal{L}(\Psi_t^{\varsigma}(X_0)) \ne \pi$ for $t > 0$ in general.

As a preparation, we let $(\mathcal{C}_0([0,1], \mathbb{R}^n), \mathcal{H}_H, \mu_H)$ denote the abstract Wiener space induced by the Gaussian process $(\tilde{B}_t)_{t \in [0,1]}$. Recall that the Cameron–Martin space is given by $\mathcal{H}_H = \mathcal{K}_H(H_0^1)$, where

$$
\mathcal{K}_H f(t) \triangleq \begin{cases} \alpha_H \displaystyle\int_0^t (t-s)^{H-\frac{3}{2}} f(s)\, ds & H > \dfrac{1}{2}, \\[2mm] \alpha_H \dfrac{d}{dt} \displaystyle\int_0^t (t-s)^{H-\frac{1}{2}} f(s)\, ds & H < \dfrac{1}{2}, \end{cases} \quad t \in [0,1],
$$

and

$$H_0^1 \triangleq \left\{ f = \int_0^{\cdot} \dot{f}(s)\,ds : \dot{f} \in L^2([0, 1], \mathbb{R}^n) \right\}$$

is the Cameron–Martin space of the standard Wiener process. The inner product on $\mathcal{H}_H$ is defined by $\langle \mathcal{K}_H f, \mathcal{K}_H g \rangle_{\mathcal{H}_H} \triangleq \langle \dot{f}, \dot{g} \rangle_{L^2}$.

We shall make use of the following simple observation.

LEMMA 3.11. *Let $f : [0, 1] \to \mathbb{R}^n$ be piecewise linear with $f(0) = 0$. Then, for each $H \in (0, 1)$, $f \in \mathcal{H}_H$ and*

(3.12)
$$\|f\|_{\mathcal{H}_H} \lesssim |\dot{f}|_{\infty}(1 + |\mathcal{D}_{\dot{f}}|).$$

PROOF. It follows from [43], Theorem 5 (see also [47]) that the inverse of $\mathcal{K}_H$ exists on the set of Lipschitz functions and there is a numerical constant $\varrho_H > 0$ such that $\mathcal{K}_H^{-1} = \varrho_H \mathcal{K}_{1-H}$. Notice also that we have $\frac{d}{dt}\mathcal{K}_H^{-1} f = \mathcal{K}_H^{-1} \dot{f}$.

Let us first consider the case $H < \frac{1}{2}$. The bound (3.12) is an immediate consequence of

$$\left| \frac{d}{dt}\mathcal{K}_H^{-1} f(t) \right| \leq \varrho_H \int_0^t (t-s)^{-H-\frac{1}{2}} |\dot{f}(s)|\,ds \lesssim |\dot{f}|_{\infty} \quad \forall t \in [0, 1].$$

For $H > \frac{1}{2}$ we let $\tau_1, \ldots, \tau_k$ denote the jump points of $\dot{f}$ in the interval $[0, t)$. Notice that

$$\left| \frac{d}{dt}\mathcal{K}_H^{-1} f(t) \right| \leq \varrho_H \left| \frac{d}{dt} \left(\sum_{i=1}^{k-1} \int_0^{\tau_1} (t-s)^{\frac{1}{2}-H} \dot{f}(s)\,ds + \cdots + \int_{\tau_k}^t (t-s)^{\frac{1}{2}-H} \dot{f}(s)\,ds \right) \right|$$

$$\lesssim |\dot{f}|_{\infty}(1 + |\mathcal{D}_{\dot{f}}|)t^{\frac{1}{2}-H}.$$

Since $1 - 2H > -1$, we obtain

$$\|f\|_{\mathcal{H}_H} = \left\| \frac{d}{dt}\mathcal{K}_H^{-1} f \right\|_{L^2} \lesssim |\dot{f}|_{\infty}(1 + |\mathcal{D}_{\dot{f}}|),$$

as required. □

The next important lemma lifts the control result of Proposition 3.9 to solutions of SDEs with additive noise.

LEMMA 3.12. *Let $b \in S(\kappa, R, \lambda)$ and $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ be invertible. Then, for any $\bar{R} > 0$ and any $\eta \in (0, \frac{1}{2})$, there is constant $\mathfrak{a}_{\eta, \bar{R}} > 0$ such that the following holds: For each $x \in \mathbb{R}^n$ and each $\varsigma \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$, we can find an event $\mathcal{A}_{x,\varsigma}$ with $\mathbb{P}(\mathcal{A}_{x,\varsigma}) \geq \mathfrak{a}_{\eta, \bar{R}}$ such that*

$$\int_0^1 \mathbf{1}_{\{t:|\Psi_t^{\varsigma}(x)(\omega)|>\bar{R}\}}(s)\,ds > \eta \quad \forall \omega \in \mathcal{A}_{x,\varsigma}.$$

PROOF. Let $u_{x,\varsigma} \in L^{\infty}([0, 1], \mathbb{R}^n)$ be the piecewise constant control furnished by Proposition 3.9 such that the occupation time of $x^{\varsigma, u_{x,\varsigma}}$ of the set $\mathbb{R}^n \setminus B_{\bar{R}+1}$ is greater than η. We set $U_{x,\varsigma} \triangleq \int_0^{\cdot} u_{x,\varsigma}(s)\,ds$ and note that $U_{x,\varsigma}$ is piecewise linear. Lemma 3.10 allows us to choose an $\varepsilon > 0$ (independent of x and ς) such that, on the event $\mathcal{A}_{x,\varsigma} \triangleq \{|U_{x,\varsigma} - \sigma \tilde{B}|_{\infty} \leq \varepsilon\}$, the occupation time of $(\Psi_h^{\varsigma}(x))_{h \in [0,1]}$ of $\mathbb{R}^n \setminus B_{\bar{R}}$ exceeds η.

It remains to show that $\inf_{x,\varsigma} \mathbb{P}(\mathcal{A}_{x,\varsigma}) > 0$. To this end, we first note that $U_{x,\varsigma} \in \mathcal{H}_H$ by Lemma 3.11. By the Cameron–Martin formula (see, e.g., [6]),

$$\mathbb{P}(\mathcal{A}_{x,\varsigma}) \geq \mathbb{P}(|\sigma^{-1}U_{x,\varsigma} - \tilde{B}|_{\infty} \leq |\sigma|^{-1}\varepsilon)$$

$$= \exp\left(-\frac{1}{2}\|\sigma^{-1}U_{x,\varsigma}\|_{\mathcal{H}_H}^2\right) \int_{\{|x|_{\infty} \leq |\sigma|^{-1}\varepsilon\}} e^{\langle x, U_{x,\varsigma}\rangle_{\mathcal{H}_H}} \mu_H(dx).$$

Consequently, Jensen's inequality and spherical symmetry give

$$(3.13) \qquad \mathbb{P}(\mathcal{A}_{x,\varsigma}) \geq \exp\left(-\frac{1}{2}\|\sigma^{-1}U_{x,\varsigma}\|^2_{\mathcal{H}_H}\right)\mathbb{P}(|\tilde{B}|_\infty \leq |\sigma|^{-1}\varepsilon).$$

Combining Proposition 3.9 and Lemma 3.11, we obtain that $\sup_{x,\varsigma}\|U_{x,\varsigma}\|_{\mathcal{H}_H} \lesssim M(1+M)$. This concludes the proof. $\square$

PROPOSITION 3.13.   *Let* $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ *be invertible. Then, for any* $\kappa, R > 0$ *and any* $p \geq 1$, *there exists a number* $\Lambda = \Lambda(\kappa, R, p) \in (0, \kappa)$ *such that the following holds: If* $b \in$ S(κ, R, Λ), *there are constants* $c, C > 0$ *such that, for any* $\varsigma \in \mathcal{C}_0([0,1], \mathbb{R}^n)$,

$$\mathcal{W}^p\big(\mathcal{L}(\Psi^\varsigma_t(Y)), \mathcal{L}(\Psi^\varsigma_t(\tilde{Y}))\big) \leq C\mathcal{W}^p(\mathcal{L}(Y), \mathcal{L}(\tilde{Y}))e^{-ct}$$

for all $t \geq 0$.

PROOF.   Write $X_t \triangleq \Psi^\varsigma_t(Y)$ and $Z_t = \Psi^\varsigma_t(\tilde{Y})$. Let $\mu_t \triangleq \mathcal{L}(X_t)$ and $\nu_t \triangleq \mathcal{L}(Z_t)$, thus (X_t, Z_t) is a synchronous coupling of μ_t and ν_t. Our strategy for proving the exponential convergence of $t \mapsto \mathcal{W}^p(\mu_t, \nu_t)$ is to show that, for any $t > 0$, the evolution of $(X_s)_{s \in [t, t+1]}$ conditional on $\mathcal{F}_t$ spends a sufficient amount of time in the contractive region $\{|x| > R\}$. As noted in Example 3.4(i), there is an independent increment decomposition $(\theta_t\tilde{B})_h = Q^t_h + \tilde{B}^t_h$ for the Riemann–Liouville process. Using this and the conditional evolution derived in Lemma 3.5, we find

$$
\begin{aligned}
&\mathbb{E}\big[|X_{t+1} - Z_{t+1}|^p\big] \\
&= \mathbb{E}\big[\mathbb{E}\big[|\Psi^\varsigma_{t,t+1}(X_t) - \Psi^\varsigma_{t,t+1}(Z_t)|^p|\mathcal{F}_t\big]\big] \\
(3.14) \qquad &= \mathbb{E}\big[\mathbb{E}\big[|\Psi_1(X_t, \theta_t\varsigma + \sigma\theta_t\tilde{B}) - \Psi_1(Z_t, \theta_t\varsigma + \sigma\theta_t\tilde{B})|^p|\mathcal{F}_t\big]\big] \\
&= \mathbb{E}\big[\mathbb{E}\big[|\Psi_1(X_t, \theta_t\varsigma + \sigma Q^t + \sigma\tilde{B}^t) - \Psi_1(Z_t, \theta_t\varsigma + \sigma Q^t + \sigma\tilde{B}^t)|^p|\mathcal{F}_t\big]\big] \\
&= \mathbb{E}\big[\mathbb{E}\big[|\Psi^{\theta_t\varsigma+\ell}_1(x) - \Psi^{\theta_t\varsigma+\ell}_1(z)|^p\big]|_{\substack{x=X_t, z=Z_t, \\ \ell=\sigma Q^t}}\big],
\end{aligned}
$$

where in the last step we also used that $(\tilde{B}^t_h)_{h\geq 0} \overset{d}{=} (\tilde{B}_h)_{h\geq 0}$.

By assumption, the drift b does not expand by more than a factor of Λ on all of $\mathbb{R}^n$. We therefore have the pathwise estimate

$$(3.15) \qquad |\Psi^{\theta_t\varsigma+\ell}_{s,t}(x) - \Psi^{\theta_t\varsigma+\ell}_{s,t}(z)|^p \leq e^{p(t-s)\Lambda}|x-z|^p$$

for all $0 \leq s < t \leq 1$. Let $\eta \in (0, \frac{1}{2})$ and $\bar{\kappa} \in (0, \kappa)$ be such that $\Xi \triangleq \bar{\kappa}\eta - \Lambda(1-\eta) > 0$ (recall that we assume $\Lambda < \kappa$). Let $\bar{R} > R$ be the corresponding radius furnished by Lemma 2.9. For any $x \in \mathbb{R}^n$ and any $\varsigma, \ell \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$, let $\mathcal{A}_{x,\theta_t\varsigma+\ell}$ be the event from Lemma 3.12. Recall that $\mathbb{P}(\mathcal{A}_{x,\theta_t\varsigma+\ell}) \geq \mathfrak{a}_{\eta,\bar{R}} > 0$ and

$$\int_0^1 \mathbf{1}_{\{s:|\Psi^{\theta_t\varsigma+\ell}_s(x)(\omega)|>\bar{R}\}}(r)\,dr > \eta \quad \forall\omega \in \mathcal{A}_{x,\theta_t\varsigma+\ell}.$$

Since $\Xi > 0$, by possibly decreasing Λ we can also ensure that

$$(3.16) \qquad 0 < \Lambda < \frac{1}{p}\log\left(\frac{1 - \mathfrak{a}_{\eta,\bar{R}}e^{-p\Xi}}{1 - \mathfrak{a}_{\eta,\bar{R}}}\right).$$

Owing to pathwise continuity of $h \mapsto \Psi^{\theta_t\varsigma+\ell}_h(x)$, there are random times $t_1, \ldots, t_{2N(\omega)}$ such that, for all $\omega \in \mathcal{A}_{x,\theta_t\varsigma+\ell}$:

- $0 \leq t_1(\omega) < \cdots < t_{2N(\omega)}(\omega) \leq 1$,
- $\sum_{i=1}^{N(\omega)}(t_{2i}(\omega) - t_{2i-1}(\omega)) \geq \eta$, and
- $\bigcup_{i=1}^{N(\omega)}(t_{2i-1}(\omega), t_{2i}(\omega)) \subset \{h \in [0,1] : |\Psi_h^{\theta_t \varsigma + \ell}(x)(\omega)| > \bar{R}\}$.

Together with (3.15) it follows that, on the event $\mathcal{A}_{x,\theta_t \varsigma + \ell}$,

$$
\begin{aligned}
\big|\Psi_1^{\theta_t \varsigma + \ell}(x) &- \Psi_1^{\theta_t \varsigma + \ell}(z)\big|^p \\
&= \big|\Psi_{t_{2N},1}^{\theta_t \varsigma + \ell}\big(\Psi_{t_{2N}}^{\theta_t \varsigma + \ell}(x)\big) - \Psi_{t_{2N},1}^{\theta_t \varsigma + \ell}\big(\Psi_{t_{2N}}^{\theta_t \varsigma + \ell}(z)\big)\big|^p \\
&\leq e^{p(1-t_{2N})\Lambda}\big|\Psi_{t_{2N}}^{\theta_t \varsigma + \ell}(x) - \Psi_{t_{2N}}^{\theta_t \varsigma + \ell}(z)\big|^p \\
&\leq e^{p(1-t_{2N})\Lambda}e^{-p(t_{2N}-t_{2N-1})\bar{\kappa}}\big|\Psi_{t_{2N-1}}^{\theta_t \varsigma + \ell}(x) - \Psi_{t_{2N-1}}^{\theta_t \varsigma + \ell}(z)\big|^p \\
&\leq \cdots \leq \exp\bigg[p\bigg(\Lambda\sum_{i=0}^{N}(t_{2i+1}-t_{2i}) - \bar{\kappa}\sum_{i=0}^{N}(t_{2i}-t_{2i-1})\bigg)\bigg]|x-z|^p \\
&\leq e^{-p\Xi}|x-z|^p,
\end{aligned}
$$

where we have set $t_{2N+1} \triangleq 1$ for convenience. On the complementary event $\Omega \setminus \mathcal{A}_{x,\theta_t \varsigma + \ell}$, we apply the trivial estimate (3.15). Inserting these bounds back into (3.14), we conclude that

$$
\mathbb{E}\big[|X_{t+1} - Z_{t+1}|^p\big] \leq \big((1 - \mathfrak{a}_{\eta,\bar{R}})e^{p\Lambda} + \mathfrak{a}_{\eta,\bar{R}}e^{-p\Xi}\big)\mathbb{E}\big[|X_t - Z_t|^p\big] \triangleq \rho\,\mathbb{E}\big[|X_t - Z_t|^p\big].
$$

Observe that $\rho < 1$ by (3.16). Finally, a straight-forward induction shows that

(3.17)
$$
\begin{aligned}
\mathcal{W}^p\big(\mathcal{L}(\Psi_t^\varsigma(Y)), \mathcal{L}(\Psi_t^\varsigma(\tilde{Y}))\big) &\leq \|X_t - Z_t\|_{L^p} \leq e^{\Lambda}\rho^{[t]}\|Y - \tilde{Y}\|_{L^p} \\
&\leq \frac{e^{\Lambda}}{\rho}e^{-|\log\rho|t}\|Y - \tilde{Y}\|_{L^p},
\end{aligned}
$$

where $[\cdot]$ denotes the integer part. Minimize over the set of couplings of $\mathcal{L}(Y)$ and $\mathcal{L}(\tilde{Y})$ to conclude the proof. $\square$

A more explicit expression for the threshold value $\Lambda(\kappa, R, p)$ can be derived by the method outlined in Remark 3.25 below. We abstain from including further details in this work. Let us however introduce the following notation.

DEFINITION 3.14. Let $\kappa, R > 0$ and $p \geq 1$. We abbreviate $S_p(\kappa, R) \triangleq S(\kappa, R, \Lambda(\kappa, R, p))$ with the constant from Proposition 3.13.

By Lemma 3.6, the Wasserstein bound of Proposition 3.13 lifts to bounds on the fast motion with frozen slow input (3.2). We obtain the following Lipschitz dependence of the flow $\bar{\Phi}$ on the initial value.

COROLLARY 3.15. *Let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration compatible with the fBm $\hat{B}$. Let $0 \leq s \leq t$ and let X, Y, and $\tilde{Y}$ be $\mathcal{F}_s$-measurable random variables. Suppose that there are $\kappa, R > 0$ such that $b(x, \cdot) \in S_1(\kappa, R)$ for every $x \in \mathbb{R}^d$. Then there is a constant $c > 0$ such that, for any Lipschitz continuous function $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$,*

$$
\big|\mathbb{E}\big[h(X, \bar{\Phi}_{s,t}^X(Y)) - h(X, \bar{\Phi}_{s,t}^X(\tilde{Y}))|\mathcal{F}_s\big]\big| \lesssim |h|_{\mathrm{Lip}}|Y - \tilde{Y}|e^{-c\frac{|t-s|}{\varepsilon}}.
$$

If, in addition, $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$, then also

$$
\big\|\bar{\Phi}_{s,t}^X(Y) - \bar{\Phi}_{s,t}^X(\tilde{Y})\big\|_{L^p} \lesssim \|Y - \tilde{Y}\|_{L^p}e^{-c\frac{|t-s|}{\varepsilon}}.
$$

PROOF. The first estimate is an immediate consequence of Lemma 3.6 and Kantorovich–Rubinstein duality. The second bound follows from the fact that we used a synchronous coupling in the proof of Proposition 3.13. □

The proof of Proposition 3.13 shows that its conclusion actually holds if $\tilde{B}$ is replaced by another process Z with similar properties.

REMARK 3.16. Let Z be a process with locally independent increment decomposition $\theta_t Z = \bar{Z}^t + \tilde{Z}^t$. Assume that:

(i) the $\mathcal{F}_t$-adapted part $\bar{Z}^t$ takes values in $\mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$ and
(ii) there is a unit vector $e \in \mathbb{R}^n$ such that, for each $t \geq 0$, $\mathcal{L}((\tilde{Z}_h^t \cdot e)_{h \in [0,1]})$ is supported on all of $\mathcal{C}_0([0,1])$.

Then a statement similar to Proposition 3.13 holds.

EXAMPLE 3.17. Suppose that $\tilde{Z}_h^t = \int_t^{t+h} \mathfrak{G}(t+h-s)\, dW_s$ for some kernel $\mathfrak{G} : \mathbb{R}_+ \to L(\mathbb{R}^n, \mathbb{R}^n)$ which is square integrable at the origin and continuous on $(0, \infty)$. Then the requirement (ii) in Remark 3.16 holds if $\int_0^t |\mathfrak{G}(s)|\, ds > 0$ for each $t > 0$. Indeed, this can be shown by a clever application of Titmarsh's convolution theorem as in [12], Lemma 2.1.

The example shows that in particular an fBm of any Hurst parameter $H \in (0, 1)$ falls in the regime of Remark 3.16. Hence, we have the following corollary to Proposition 3.13.

COROLLARY 3.18. *Let $p \geq 1$ and suppose that $b \in S_p(\kappa, R)$ for some $\kappa, R > 0$. Let $(X_t)_{t \geq 0}$ be the solution to*

$$(3.18) \qquad dX_t = b(X_t)\, dt + \sigma\, dB_t$$

started in the generalized initial condition μ, where $(B_t)_{t \geq 0}$ is an fBm with Hurst parameter $H \in (0, 1)$ and $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ is invertible. Then there is a unique invariant measure $\mathcal{I}_\pi \in \mathcal{P}(\mathbb{R}^n \times \mathcal{H}_H)$ for the equation (3.18) in the sense of Definition 2.1. Moreover, writing $\pi = \Pi^*_{\mathbb{R}^n} \mathcal{I}_\pi$ for the first marginal, there are constants $c, C > 0$ such that

$$(3.19) \qquad \mathcal{W}^p\big(\mathcal{L}(X_t), \pi\big) \leq C\mathbb{W}^p(\mu, \mathcal{I}_\pi) e^{-ct}$$

for all $t \geq 0$.

PROOF. By Proposition 2.4, we know that there is an invariant measure $\mathcal{I}_\pi$ to (3.18) with moments of all orders. The Wasserstein estimate (3.19) then follows by the very same arguments as in Proposition 3.13. The only difference is that we now have to specify a generalized initial condition $\nu \in \mathcal{P}((\mathbb{R}^n \times \mathcal{H}_H)^2)$ for the coupling (X_t, Z_t); see Section 2.1. Unlike for the conditioned dynamics, we have $Z_t \sim \pi$ if we start Z in the invariant measure $\mathcal{I}_\pi$. In order for our previous argument to apply, we need to ensure that the past of the noises in the synchronous coupling coincide. In (3.17) we can thus only minimize over couplings in the set

$$\big\{\rho \in \mathcal{P}\big((\mathbb{R}^n \times \mathcal{H}_H)^2\big) : \rho\big(\mathbb{R}^n \times \mathbb{R}^n \times \Delta_{\mathcal{H}_H}\big) = 1\big\},$$

which precisely yields (3.19). □

3.4. *Quenched convergence to the invariant measure.* The other distance, which will play a role in Section 4.2 below, is between $\mathcal{L}(\Psi_t^\varsigma(Y))$ and the stationary law π of the equation (3.18). We stress that—contrarily to the proof of Corollary 3.18—we cannot simply start the process in the invariant measure. In fact, the measure π is not stationary for (3.11) since the increments of $\tilde{B}$ are not stationary. It is therefore necessary to wait for a sufficient decay of the deterministic "adversary" ς, whence we only find an algebraic rate of convergence. Before we state the result, let us first illustrate that there is indeed no hope for an exponential rate.

EXAMPLE 3.19. Let

$$dX_t = -X_t\,dt + d\tilde{B}_t, \qquad dY_t = -Y_t\,dt + dB_t.$$

If we start both X and Y in the generalized initial condition $\delta_0 \otimes \mathsf{W}$, then $\mathcal{L}(X_t) = N(0, \Sigma_t^2)$ and $\mathcal{L}(Y_t) = N(0, \bar{\Sigma}_t^2)$ where

$$\Sigma_t^2 = \bar{\Sigma}_t^2 - \mathbb{E}\left[\left|\int_0^t e^{-(t-s)}\dot{\bar{B}}_s\,ds\right|^2\right].$$

In particular, $\mathcal{W}^2(\mathcal{L}(X_t), \mathcal{L}(Y_t)) = |\Sigma_t - \bar{\Sigma}_t| \gtrsim t^{-(1-\hat{H})}$ uniformly in $t \geq 1$. Since it is easy to see that $\mathcal{W}^2(\mathcal{L}(Y_t), \pi) \lesssim e^{-t}$, it follows that $\mathcal{W}^2(\mathcal{L}(X_t), \pi) \gtrsim t^{-(1-\hat{H})}$.

PROPOSITION 3.20. *Suppose that $b \in S_p(\kappa, R)$ for some $\kappa, R > 0$ and $\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)$ is invertible. Let $p \geq 1$, $\varsigma \in \Omega_\alpha$ for some $\alpha > 0$, and Y be an $\mathcal{F}_0$-measurable random variable. Then, for each $\beta < \min(\alpha, 1 - H)$, there is a constant $C > 0$ such that*

$$(3.20) \qquad \mathcal{W}^p\big(\mathcal{L}(\Psi_t^\varsigma(Y)), \pi\big) \leq C\frac{(1 + \|\varsigma\|_{\Omega_\beta})(1 + \mathcal{W}^p(\mathcal{L}(Y), \pi))}{t^\beta}$$

for all $t > 0$.

PROOF. Fix $t \geq 1$, abbreviate $X \triangleq \Psi^\varsigma(Y)$, and let Z be the stationary solution to the equation (3.18). We assume that X and Z are driven by the same Wiener process. Let us first consider the case $p \geq 2$. Recall the following locally independent decompositions from Section 3.1:

$$\theta_t B = \bar{B}^t + \tilde{B}^t, \qquad \theta_t \tilde{B} = Q^t + \tilde{B}^t.$$

Remember also that the "smooth" part of the fBm increment can be further decomposed as $\bar{B}^t = P^t + Q^t$; see Example 3.4(iii). Therefore,

$$\mathbb{E}[|X_{t+1} - Z_{t+1}|^p]$$

$$= \mathbb{E}\big[\mathbb{E}[|\Psi_{t,t+1}(X_t, \varsigma + \sigma\tilde{B}) - \Psi_{t,t+1}(Z_t, \sigma B)|^p|\mathcal{F}_t]\big]$$

$$(3.21) \qquad = \mathbb{E}\big[\mathbb{E}[|\Psi_1(X_t, \theta_t\varsigma + \sigma Q^t + \sigma\tilde{B}^t) - \Psi_1(Z_t, \sigma P^t + \sigma Q^t + \sigma\tilde{B}^t)|^p|\mathcal{F}_t]\big]$$

$$= \mathbb{E}\big[|\Psi_1^{\theta_t\varsigma+\ell}(x) - \Psi_1^{\bar{\ell}+\ell}(z)|^p|_{\substack{x=X_t, z=Z_t, \\ \ell=\sigma Q^t, \bar{\ell}=\sigma P^t}}\big].$$

Write $R_h \triangleq \Psi_h^{\theta_t\varsigma+\ell}(x)$ and $S_h \triangleq \Psi_h^{\bar{\ell}+\ell}(z)$. Notice that, since ς and $\bar{\ell}$ are differentiable,

$$\frac{d}{dh}|R_h - S_h|^p = p\langle\dot{\varsigma}_{t+h} - \dot{\bar{\ell}}_h + b(R_h) - b(S_h), R_h - S_h\rangle|R_h - S_h|^{p-2}$$

$$\leq p(\Lambda + \gamma)|R_h - S_h|^p + \left(\frac{p-1}{\gamma p}\right)^{p-1}(|\dot{\varsigma}_{t+h}| + |\dot{\bar{\ell}}_h|)^p$$

for any $\gamma > 0$, where $\Lambda = \Lambda(\kappa, R, p)$ is the expansion threshold derived in Proposition 3.13. It follows that, for any $0 \le h_1 \le h_2 \le 1$,

$$
\begin{aligned}
|R_{h_2} &- S_{h_2}|^p \\
&\le |R_{h_1} - S_{h_1}|^p e^{p(\Lambda+\gamma)(h_2-h_1)} \\
(3.22) \qquad &\quad + \left(\frac{p-1}{\gamma p}\right)^{p-1} \int_{h_1}^{h_2} e^{p(\Lambda+\gamma)(h_2-s)} \big(|\dot{\varsigma}_{t+s}| + |\dot{\bar{\ell}}_s|\big)^p \, ds \\
&\le |R_{h_1} - S_{h_1}|^p e^{p(\Lambda+\gamma)(h_2-h_1)} + C_\gamma (h_2 - h_1),
\end{aligned}
$$

where we abbreviated

$$
C_\gamma \triangleq \left(\frac{p-1}{\gamma p}\right)^{p-1} \left(\frac{\|\varsigma\|_{\Omega_\beta}}{t^\beta} + |\dot{\bar{\ell}}|_\infty\right)^p.
$$

We now argue similar to Proposition 3.13: Pick $\eta \in (0, \frac{1}{2})$ and $\bar{\kappa} \in (0, \kappa)$ such that $\Xi \triangleq \eta\bar{\kappa} - (1-\eta)\Lambda > 0$. Let $\bar{R} > 0$ be the corresponding constant of Lemma 2.9 and $\mathcal{A}_{x,\theta_t\varsigma+\ell}$ be the event furnished by Lemma 3.12. As before, we write $t_1, \ldots, t_{2N(\omega)}$ for the random times characterizing the excursions of $(R_h)_{h\in[0,1]}$ outside of $B_{\bar{R}}$, see Proposition 3.13. By an argument similar to (3.22),

$$
(3.23) \qquad |R_{t_{2i}} - S_{t_{2i}}|^p \le |R_{t_{2i-1}} - S_{t_{2i-1}}|^p e^{p(\gamma-\bar{\kappa})(t_{2i}-t_{2i-1})} + C_\gamma (t_{2i} - t_{2i-1})
$$

for all $i = 1, \ldots, N(\omega)$ on the set $\mathcal{A}_{x,\theta_t\varsigma+\ell}$. Combining (3.22) and (3.23), we further find on this set

$$
\begin{aligned}
|R_1 - S_1|^p &\le e^{p(\Lambda+\gamma)(1-t_{2k})} |R_{t_{2k}} - S_{t_{2k}}|^p + C_\gamma(1 - t_{2k}) \\
&\le e^{p(\Lambda+\gamma)(1-t_{2k})} e^{p(\gamma-\bar{\kappa})(t_{2k}-t_{2k-1})} |R_{t_{2k-1}} - S_{t_{2k-1}}|^p + C_\gamma(1 - t_{2k-1}) \\
&\le \cdots \le e^{p(\Lambda+\gamma)(1-\eta)+p(\gamma-\bar{\kappa})\eta} |x - z|^p + C_\gamma \le e^{-p(\Xi-\gamma)} |x - z|^p + C_\gamma.
\end{aligned}
$$

Choose $\gamma > 0$ sufficiently small such that simultaneously $\Xi - \gamma > 0$ and

$$
\rho \triangleq (1 - \mathfrak{a}_{\eta,\bar{R}}) e^{p(\Lambda+\gamma)} + \mathfrak{a}_{\eta,\bar{R}} e^{-p(\Xi-\gamma)} < 1.
$$

This shows that

$$
(3.24) \qquad \mathbb{E}\big[|R_1 - S_1|^p\big] \le \rho |x - y|^p + C_\gamma.
$$

It is clear that the estimate (3.24) also holds for $p < 2$ with the constant

$$
C_\gamma = \frac{1}{(2\gamma)^{\frac{p}{2}}} \left(\frac{\|\varsigma\|_{\Omega_\beta}}{t^\beta} + |\dot{\bar{\ell}}|_\infty\right)^p
$$

and a slightly increased $\rho < 1$. Since $P^t = \bar{B}_{t+\cdot}$, Lemma 3.7 and the identity (3.21) show that

$$
\mathbb{E}\big[|X_{t+1} - Y_{t+1}|^p\big] \le \rho \mathbb{E}\big[|X_t - Y_t|^p\big] + \frac{C(1 + \|\varsigma\|_{\Omega_\beta}^p)}{t^{p\beta}}
$$

for some numerical constant $C > 0$ independent of t and ς. Therefore, iterating this bound we find

$$
(3.25) \qquad \mathbb{E}\big[|X_t - Y_t|^p\big] \lesssim e^{-ct} \mathbb{E}\big[|X_0 - Y_0|^p\big] + C\big(1 + \|\varsigma\|_{\Omega_\beta}^p\big) \sum_{i=0}^{[t]-2} \frac{\rho^i}{(t-1-i)^{p\beta}}.
$$

The last sum is easily seen to be $\lesssim t^{-p\beta}$ uniformly in $t \ge 2$ and the claim follows at once. $\qquad\square$

By a strategy inspired by [39], Section 7 (see also [21]), we can lift Proposition 3.20 to a total variation bound. Since the exposition of Panloup and Richard does not immediately transfer to the problem at hand, we choose to include the necessary details. Consider the system

$$(3.26) \qquad \begin{aligned} dX_s &= b(X_s)\,ds + d\varsigma_s + \sigma\,d\tilde{B}_s, \\ dZ_s &= b(Z_s)\,ds + \sigma\,dB_s + \sigma\varphi^t(s)\,ds, \end{aligned}$$

where X_0 is an arbitrary initial condition and Z is the stationary solution of the first equation. Our aim is to exhibit an adapted integrable function $\varphi^t : [0, t+1] \to \mathbb{R}^n$ which vanishes on $[0, t]$ and ensures that $X_{t+1} = Z_{t+1}$. To this end, we define

$$(3.27) \qquad \varphi^t(s) \triangleq \begin{cases} \left(2\dfrac{|X_t - Z_t|^{\frac{1}{2}}}{|X_s - Z_s|^{\frac{1}{2}}} + \lambda \right) \sigma^{-1}(X_s - Z_s) - \dot{\tilde{B}}_s + \sigma^{-1}\dot{\varsigma}_s & s \in [t, t+1], \\ 0 & \text{otherwise.} \end{cases}$$

LEMMA 3.21.   *Let $t \geq 1$, $\varsigma \in \Omega_\alpha$, $b \in S(\kappa, R, \lambda)$, and consider the system (3.26) with φ^t defined in (3.27). Then $X_{t+1} = Z_{t+1}$ and, for any $\beta < \alpha \wedge (1 - H)$,*

$$(3.28) \qquad \begin{aligned} |\varphi^t|_\infty &\lesssim |X_t - Z_t| + \frac{\|\varsigma\|_{\Omega_\beta} + \|\bar{B}\|_{\Omega_\beta}}{t^\beta}, \\ |\dot{\varphi}^t|_\infty &\lesssim |X_t - Z_t|^{\frac{1}{2}} + |X_t - Z_t| + \frac{\|\varsigma\|_{\Omega_\beta} + \|\bar{B}\|_{\Omega_\beta}}{t^{1+\beta}}, \end{aligned}$$

where the derivative of φ^t is understood as right- and left-sided derivative at the boundaries t and $t + 1$, respectively.

PROOF.   The argument is a minor modification of [21], Lemma 5.8: Abbreviate $f(s) \triangleq |X_s - Z_s|^2$, then

$$f'(s) = 2\langle b(X_s) - b(Z_s) + \dot{\varsigma}_s - \sigma\dot{\tilde{B}}_s - \sigma\varphi^t(s), X_s - Z_s \rangle \leq -4|X_t - Z_t|^{\frac{1}{2}} f(s)^{\frac{3}{4}}$$

since $b \in S(\kappa, R, \lambda)$. It follows that

$$|X_s - Z_s|^{\frac{1}{2}} \leq |X_t - Z_t|^{\frac{1}{2}} - (s - t)|X_t - Z_t|^{\frac{1}{2}} \quad \forall s \in [t, t+1],$$

whence $X_{t+1} = Z_{t+1}$. This also implies

$$\left| \frac{d}{ds}(X_s - Z_s) \right| \leq (|b|_{\text{Lip}} + 2 + \lambda)|X_t - Z_t|^{\frac{1}{2}}|X_s - Z_s|^{\frac{1}{2}}$$

and consequently

$$\left| \frac{d}{ds}\left(\frac{X_s - Z_s}{|X_s - Z_s|^{\frac{1}{2}}} \right) \right| \leq \frac{3}{2} \frac{|\frac{d}{ds}(X_s - Z_s)|}{|X_s - Z_s|^{\frac{1}{2}}} \lesssim |X_t - Z_t|^{\frac{1}{2}}.$$

The bounds (3.28) follow at once. $\square$

REMARK 3.22.   We stress that the bound on $|\dot{\varphi}^t|_\infty$ only holds for a Lipschitz continuous drift b.

It is now easy to prove the following result.

PROPOSITION 3.23. *Assume the conditions of Proposition* 3.20 *for* $p = 1$. *Then, for any* $\beta < \alpha \wedge (1 - H)$, *it holds that*

$$\left\| \mathcal{L}(\Psi_t^\varsigma(Y)) - \pi \right\|_{\mathrm{TV}} \lesssim t^{-\frac{\beta}{3}} (1 + \|\varsigma\|_{\Omega_\beta})(1 + \mathcal{W}^1(\mathcal{L}(Y), \pi)) \quad \forall t > 0.$$

PROOF. Let B and B' be H-fBms built from underlying two-sided Wiener processes W and W', see (2.1). Recall that $\tilde{B}$ is the Riemann–Liouville process associated with B. Let X and Z solve

$$(3.29) \qquad \begin{aligned} dX_s &= b(X_s)\,ds + d\varsigma_s + \sigma\,d\tilde{B}_s, \\ dZ_s &= b(Z_s)\,ds + \sigma\,dB'_s, \end{aligned}$$

where $X_0 \overset{d}{=} Y$ and Z is the stationary solution. Fix $t > 1$. We shall use the bound

$$\left\| \mathcal{L}(\Psi_{t+1}^\varsigma(Y)) - \pi \right\|_{\mathrm{TV}} = \inf_{(\tilde{B}, B')} \mathbb{P}(X_{t+1} \neq Z_{t+1}) \leq \inf_{(W, W')} \mathbb{P}(X_{t+1} \neq Z_{t+1})$$

$$(3.30) \qquad \leq \inf_{(W, W')} \mathbb{P}(X_{t+1} \neq Z_{t+1}, |X_t - Z_t| \leq \delta)$$

$$+ \inf_{(W, W')} \mathbb{P}(|X_t - Z_t| > \delta).$$

Taking W and W' equal, we are in the setting of Proposition 3.20. The estimate (3.25) thus shows that, for any $\delta \in (0, 1]$,

$$\inf_{(W, W')} \mathbb{P}(|X_t - Z_t| > \delta) \leq \frac{C(1 + \|\varsigma\|_{\Omega_\beta})(1 + \mathcal{W}^1(\mathcal{L}(Y), \pi))}{\delta t^\beta}.$$

To bound the first term in (3.30) we exploit the fact that X_t and Z_t are already close so that we can couple them at time $t + 1$ with a controlled cost. Let φ^t be the function from Lemma 3.21; in particular $\varphi^t(s) = 0$ for $s < t$. We observe that $B' = B + \int_0^\cdot \varphi^t(s)\,ds$ on $[0, t+1]$ if and only if $W' = W + \int_{-\infty}^\cdot \psi^t(s)\,ds$ on $(-\infty, t+1]$, where for a suitable constant $\gamma_H \in \mathbb{R}$,

$$\psi^t(s) = \begin{cases} \gamma_H \dfrac{d}{ds} \displaystyle\int_t^s (s - u)^{\frac{1}{2} - H} \varphi^t(u)\,du & s \in [t, t+1], \\ 0 & s \in (-\infty, t), \end{cases}$$

see [21], Lemma 4.2, for details. Let $\mathcal{T}$ be the linear transformation $\mathcal{T}(w) = w + \int_{-\infty}^\cdot \psi^t(s)\,ds$, $w \in \mathcal{C}((-\infty, t+1], \mathbb{R}^n)$. We return to the equations (3.29). By the construction, we have $X_{t+1} = Z_{t+1}$ for those realizations for which $W' = \mathcal{T}(W)$. In particular,

$$\inf_{(W, W')} \mathbb{P}(X_{t+1} \neq Z_{t+1}, |X_t - Z_t| \leq \delta) \leq \inf_{(W, W')} \mathbb{P}(W' \neq \mathcal{T}(W), |X_t - Z_t| \leq \delta).$$

Throughout this proof, all stochastic processes are considered only up to time $t + 1$. Let $\hat{\mathbb{P}}(\cdot) = \mathbb{P}(\cdot \mid |X_t - Z_t| \leq \delta)$. Let μ and ν denote the laws under $\hat{\mathbb{P}}$ of W and $\mathcal{T}(W)$, respectively. We take an optimal coupling $(\hat{W}, \hat{W}')$ achieving the total variation distance $\|\mu - \nu\|_{\mathrm{TV}}$. Then the Pinsker–Csiszar inequality, a consequence of the Girsanov theorem, shows that

$$\inf_{(W, W')} \mathbb{P}(X_{t+1} \neq Z_{t+1}, |X_t - Z_t| \leq \delta) \leq \|\mu - \nu\|_{\mathrm{TV}} \mathbb{P}(|X_t - Z_t| \leq \delta)$$

$$\leq \frac{1}{2} \mathbb{E}_{\hat{\mathbb{P}}} \left[\int_t^{t+1} |\psi^t(s)|^2\,ds \right]^{\frac{1}{2}} \mathbb{P}(|X_t - Z_t| \leq \delta).$$

On integration by parts we find

$$\int_t^{t+1} |\psi^t(s)|^2\,ds \lesssim \begin{cases} |\varphi^t|_\infty^2 & H < \dfrac{1}{2}, \\ |\varphi^t|_\infty^2 + |\dot{\varphi}^t|_\infty^2 & H > \dfrac{1}{2}. \end{cases}$$

In either case, (3.28) yields

$$\int_t^{t+1} |\psi^t(s)|^2 \, ds \lesssim \delta + \frac{\|\varsigma\|_{\Omega_\beta}^2 + \|\bar{B}\|_{\Omega_\beta}^2}{t^{2\beta}}$$

on the event $\{|X_t - Z_t| \le \delta\}$ and therefore

$$\inf_{(W,W')} \mathbb{P}(X_{t+1} \ne Z_{t+1}, |X_t - Z_t| \le \delta) \lesssim \sqrt{\delta} + \frac{\|\varsigma\|_{\Omega_\beta} + \|\|\bar{B}\|_{\Omega_\beta}\|_{L^2}}{t^\beta}.$$

Combining this with (3.30) and Lemma 3.7, we have proven

$$(3.31) \qquad \|\mathcal{L}(\Psi_{t+1}^\varsigma(Y)) - \pi\|_{\mathrm{TV}} \lesssim (1 + \|\varsigma\|_{\Omega_\beta})(1 + \mathcal{W}^1(\mathcal{L}(Y), \pi))\left(\sqrt{\delta} + \frac{1}{\delta t^\beta}\right),$$

which is minimized for $\delta = t^{-\frac{2\beta}{3}}$. $\quad\square$

By duality and Lemma 3.6, we obtain the following ergodic theorem as a corollary to Propositions 3.20 and 3.23. It provides the fundamental estimates for our proof of the averaging principle for the fractional slow-fast system with feedback dynamics.

COROLLARY 3.24. *Let $0 \le s \le t$ and let X, Y be $\mathcal{F}_s$-measurable random variables. Suppose that there are κ, $R > 0$ such that $b(x, \cdot) \in S_1(\kappa, R)$ for every $x \in \mathbb{R}^d$. Then, for any $\zeta < 1 - \hat{H}$ and:*

(i) *any Lipschitz function $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$,*

$$|\mathbb{E}[h(X, \bar{\Phi}_{s,t}^X(Y)) - \bar{h}(X) \,|\, \mathcal{F}_s]| \lesssim |h|_{\mathrm{Lip}}(1 + \|\varepsilon^{-\hat{H}}\bar{B}_{\varepsilon \cdot}^s\|_{\Omega_\zeta})(1 + |Y|)\left(1 \wedge \frac{\varepsilon^\zeta}{|t - s|^\zeta}\right).$$

(ii) *any bounded measurable function $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$,*

$$|\mathbb{E}[h(X, \bar{\Phi}_{s,t}^X(Y)) - \bar{h}(X) \,|\, \mathcal{F}_s]| \lesssim |h|_\infty (1 + \|\varepsilon^{-\hat{H}}\bar{B}_{\varepsilon \cdot}^s\|_{\Omega_\zeta})(1 + |Y|)\left(1 \wedge \frac{\varepsilon^{\frac{\zeta}{3}}}{|t - s|^{\frac{\zeta}{3}}}\right).$$

Here, as usual, $\bar{h}(x) = \int_{\mathbb{R}^n} h(x, y) \pi^x(dy)$.

3.5. *Geometric ergodicity for SDEs driven by fractional Brownian motion.* Applying the arguments of Propositions 3.13 and 3.23 to the equation

$$(3.32) \qquad\qquad dY_t = b(Y_t) \, dt + \sigma \, dB_t,$$

we obtain an exponential rate of convergence improving the known results:

PROOF OF THEOREM 1.3. In Corollary 3.18 we have already proven the Wasserstein decay (1.7):

$$\mathcal{W}^p(\mathcal{L}(Y_t), \pi) \le C e^{-ct} \mathbb{W}^p(\mu, \pi), \quad \forall t \ge 0.$$

The total variation rate (1.8) then follows by a similar Girsanov coupling as in the proof of Proposition 3.23. In fact, we now consider

$$dX_s = b(X_s) \, ds + \sigma \, dB_s,$$
$$dZ_s = b(Z_s) \, ds + \sigma \, dB_s + \sigma \varphi^t(s) \, ds,$$

where X is started in the generalized initial condition μ and Z is the stationary solution. Let us define

$$\varphi^t(s) \triangleq -\left(\frac{4|X_t - Z_t|^{\frac{1}{2}}}{|X_s - Z_s|^{\frac{1}{2}}} + \lambda\right)\sigma^{-1}(X_s - Z_s)\mathbf{1}_{[t,t+1]}(s).$$

It can then be checked similar to Lemma 3.21 that $X_{t+1} = Y_{t+1}$ and

$$|\varphi^t|_\infty \lesssim |X_t - Z_t|, \qquad |\dot{\varphi}^t|_\infty \lesssim |X_t - Z_t|^{\frac{1}{2}} + |X_t - Z_t|.$$

Consequently, the estimate (3.31) becomes

$$(3.33) \qquad \|\mathcal{L}(Y_{t+1}) - \pi\|_{\mathrm{TV}} \lesssim \mathbb{W}^1(\mu, \pi)\left(\sqrt{\delta} + \frac{e^{-ct}}{\delta}\right)$$

and choosing $\delta = e^{-\frac{ct}{2}}$ shows a geometric decay of the total variation distance at a fixed time. To get asserted decay on the path space (1.8), we observe that, by the very same argument as in [39], Proposition 7.2(iii), φ^t actually induces a coupling on the path space with a similar cost. Hence, $\|\mathcal{L}(Y_{t+\cdot}) - \mathbb{P}_\pi\|_{\mathrm{TV}}$ is still bounded by a quantity proportional to the right-hand side of (3.33) and (1.8) follows at once. $\quad\square$

REMARK 3.25.   The admissible repulsivity strength $\Lambda(\kappa, R, p)$ obtained in the proof of Theorem 1.3 is certainly not optimal. We therefore abstain from deriving a quantitative upper bound. Let us however indicate one way to obtain such an estimate: Start from (3.13) in the proof of Lemma 3.12 and recall a standard result (see, e.g., [44], Theorem D.4) saying that

$$\mathbb{P}(|\tilde{B}|_\infty \le |\sigma|^{-1}\varepsilon) \ge 1 - K(|\sigma|^{-1}\varepsilon)^{\frac{1}{H}}e^{-H(|\sigma|^{-1}\varepsilon)^2}$$

for a known numerical constant $K > 0$. Finally optimize over all constants involved.

Let us finally sketch the main differences for a more general Gaussian driving noise G in equation (3.32). We assume that G has continuous sample paths and a moving average representation similar to (2.1) with a kernel $\mathfrak{G} : \mathbb{R} \to L(\mathbb{R}^n, \mathbb{R}^n)$ which vanishes on $(-\infty, 0]$, is continuous on $(0, \infty)$, and satisfies

$$\int_{-\infty}^t |\mathfrak{G}(t - u) - \mathfrak{G}(-u)|^2 \, du < \infty$$

for each $t > 0$. Then

$$G_t = \int_{-\infty}^t \mathfrak{G}(t - u) - \mathfrak{G}(-u) \, dW_u, \quad t \ge 0,$$

has the locally independent increment decomposition

$$(\theta_t G)_h = \int_{-\infty}^t \mathfrak{G}(t + h - u) - \mathfrak{G}(t - u) \, dW_u + \int_t^{t+h} \mathfrak{G}(t + h - u) \, dW_u \triangleq \bar{G}_h^t + \tilde{G}_h^t$$

with respect to any compatible filtration. Moreover, we require that

$$\int_0^\delta |\mathfrak{G}(u)| \, du > 0$$

for each $\delta > 0$. We remark that (up to a time-shift) this is certainly implied by the assumptions of Panloup and Richard; see [39], Condition (**C₂**). As we have seen in Example 3.17, the Cameron–Martin space of $(\tilde{G}_h)_{h\in[0,1]}$ then densely embeds into $\mathcal{C}_0([0, 1], \mathbb{R}^n)$. Thus Remark 3.16 applies and we obtain a geometric rate in Wasserstein distance, provided that there is a stationary measure for the equation $dY_t = b(Y_t) \, dt + \sigma \, dG_t$.

4. The fractional averaging principle. Let us remind the reader of the setup of Theorem 1.2: We consider the slow-fast system

$$(4.1) \qquad dX_t^\varepsilon = f(X_t^\varepsilon, Y_t^\varepsilon)\, dt + g(X_t^\varepsilon, Y_t^\varepsilon)\, dB_t, \quad X_0^\varepsilon = X_0,$$

$$(4.2) \qquad dY_t^\varepsilon = \frac{1}{\varepsilon} b(X_t^\varepsilon, Y_t^\varepsilon)\, dt + \frac{1}{\varepsilon^{\hat{H}}} \sigma\, d\hat{B}_t, \quad Y_0^\varepsilon = Y_0,$$

driven by independent d-dimensional and n-dimensional fractional Brownian motions B and $\hat{B}$ with Hurst parameters $H \in (\frac{1}{2}, 1)$ and $\hat{H} \in (1 - H, 1)$, respectively. We claim that X_t^ε converges to the solution of the naïvely averaged equation (1.3) as $\varepsilon \to 0$.

Let us also introduce the following filtrations for later reference:

$$\mathcal{G}_t \triangleq \sigma(B_s, s \le t), \qquad \hat{\mathcal{G}}_t \triangleq \sigma(\hat{B}_s, s \le t), \qquad \mathcal{F}_t \triangleq \mathcal{G}_t \vee \hat{\mathcal{G}}_t.$$

To be utterly precise, we actually use the right-continuous completion of $\mathcal{F}$ in order to ensure that hitting time of an open sets by a continuous, adapted process is a stopping time. Observe that $\mathcal{F}$ is compatible with the fBm $\hat{B}$; see Section 3.1.

We shall first convince ourselves that, under the conditions of Theorem 1.2, the pathwise solution of the slow-fast system (4.1)–(4.2) exists globally. If the drift vector field $b : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ in (4.2) were *globally* Lipschitz continuous, this would be an easy consequence of the standard Young bound [50]:

$$(4.3) \qquad \left| \int_s^t f_r\, d\mathfrak{h}_r \right| \lesssim |f|_{C^\beta} |\mathfrak{h}|_{C^\alpha} |t - s|^{\alpha+\beta} + |f_s| |\mathfrak{h}|_{C^\alpha} |t - s|^\alpha,$$

provided that $\alpha + \beta > 1$. We shall also prove a bound on the moments of the Hölder norm of the solution for any fixed scale ε. The main technical estimates in the proof of Theorem 1.2 are delegated to Section 4.2, allowing us to easily conclude the argument in Section 4.3 by appealing to Lê's stochastic sewing lemma [29].

4.1. *A solution theory for the slow-fast system.* We shall begin with a deterministic (pathwise) existence and uniqueness result. Fix a terminal time $T > 0$ and let $\mathfrak{h} = (\mathfrak{h}^1, \mathfrak{h}^2) \in C^{\alpha_1}([0, T], \mathbb{R}^m) \times C^{\alpha_2}([0, T], \mathbb{R}^n)$, where $\alpha_1 > \frac{1}{2}$ and $\alpha_2 > 1 - \alpha_1$. We consider the Young differential equation

$$(4.4) \qquad z(t) = \begin{pmatrix} z^1(t) \\ z^2(t) \end{pmatrix} = z_0 + \int_0^t \begin{pmatrix} F_1(z(s)) \\ F_2(z(s)) \end{pmatrix} ds + \int_0^t G(z(s))\, d\mathfrak{h}_s.$$

We impose the following assumptions on the data.

CONDITION 4.1. We shall assume all of the following:

(i) $F_1 : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$ is bounded and globally Lipschitz continuous.

(ii) $F_2 : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz continuous and of linear growth, that is, $|F_2(z, x)| \lesssim 1 + |x| + |z|$ for all $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^d$. Moreover, there are $\kappa, D > 0$ such that

$$\langle F_2(z, x) - F_2(z, y), x - y \rangle \le D - \kappa |x - y|^2 \quad \forall x, y \in \mathbb{R}^n, \forall z \in \mathbb{R}^d.$$

(iii) $G : \mathbb{R}^d \times \mathbb{R}^n \to L(\mathbb{R}^{m+n}, \mathbb{R}^{d+n})$ is of the form $G = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix}$ with $G_1 \in C_b^2(\mathbb{R}^d \times \mathbb{R}^n, L(\mathbb{R}^m, \mathbb{R}^d))$ and $G_2 \in L(\mathbb{R}^d, \mathbb{R}^d)$ is constant.

Our proof for the well-posedness of (4.4) and the nonexplosiveness is based on the following comparison lemma, versions of which will be of repeated use in the sequel.

LEMMA 4.2.   *Let $F_2 : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ satisfy Condition 4.1(ii) and let $\varsigma \in \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^n)$, $z \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$.*

(i) *Then, for any $x_0 \in \mathbb{R}^n$, there are unique global solutions to*

$$x(t) = x_0 + \int_0^t F_2\big(z(s), x(s)\big)\, ds + \varsigma_t, \qquad y(t) = x_0 - \int_0^t y(s)\, ds + \varsigma_t.$$

Furthermore, on any finite time interval $[0, T]$, the difference of the solutions satisfies the bound

$$(4.5) \qquad |x(t) - y(t)|^2 \lesssim \int_0^t e^{-\kappa(t-s)}\big(1 + |y(s)| + |z(s)|\big)^2\, ds$$

for all $t \in [0, T]$. In particular,

$$(4.6) \qquad |x|_\infty \lesssim 1 + |x_0| + |\varsigma|_\infty + |z|_\infty.$$

(ii) *If, in addition, $\varsigma \in \mathcal{C}^\alpha([0, T], \mathbb{R}^n)$ for some $\alpha > 0$, then $x \in \mathcal{C}^\alpha([0, T], \mathbb{R}^n)$ and the following bound holds:*

$$(4.7) \qquad |x|_{\mathcal{C}^\alpha} \lesssim 1 + |x_0| + |z|_\infty + |\varsigma|_{\mathcal{C}^\alpha}.$$

PROOF.   Since F_2 is locally Lipschitz, it is clear that uniqueness holds for the equation defining x. To see existence, first notice that $\tilde{x}(t) \triangleq x(t) - \varsigma_t$ solves

$$\tilde{x}(t) = x_0 + \int_0^t F_2\big(z(s), \tilde{x}(s) + \varsigma_s\big)\, ds.$$

Set $\Upsilon(s, x) = F_2(z(s), x + \varsigma_s)$. This function is jointly continuous in (s, x). Therefore, a local solution exists by the Carathéodory theorem.

On the other hand, global existence and uniqueness of y is standard. Consequently, the required nonexplosion statement follows easily upon establishing (4.5). To this end, we first observe that, for all $z \in \mathbb{R}^d$ and all $x, y \in \mathbb{R}^n$, the off-diagonal large scale contraction property and the linear growth of F furnish the following bound:

$$\langle F_2(z, x) + y, x - y\rangle \le D - \kappa|x - y|^2 + \langle F_2(z, y) + y, x - y\rangle$$

$$\le D - \frac{\kappa}{2}|x - y|^2 + \frac{C}{\kappa}\big(1 + |z| + |y|\big)^2$$

for some uniform constant $C > 0$, where we also used Young's inequality. Consequently, the function $h(t) \triangleq e^{\kappa t}|x(t) - y(t)|^2$ satisfies

$$h'(t) \lesssim e^{\kappa t}\big(1 + |y(t)| + |z(t)|\big)^2$$

and (4.5) follows at once.

The bound (4.7) is an immediate consequence of (4.5) together with the fact that

$$|x|_{\mathcal{C}^\alpha} \lesssim |F_2(z, x)|_\infty T^{1-\alpha} + |\varsigma|_{\mathcal{C}^\alpha} \lesssim \big(1 + |z|_\infty + |x|_\infty\big)T^{1-\alpha} + |\varsigma|_{\mathcal{C}^\alpha}. \qquad \square$$

The announced existence and uniqueness result for (4.4) is as follows.

PROPOSITION 4.3.   *Under Condition 4.1, for any $T > 0$ and any $\beta < \alpha_1 \wedge \alpha_2$, (4.4) has a unique global solution in $\mathcal{C}^\beta([0, T], \mathbb{R}^{d+n})$.*

PROOF.   Owing to Lemma 4.2, it is enough to derive an a priori bound on $|z^1|_{\mathcal{C}^{\tilde{\alpha}}}$, $\tilde{\alpha} \in [\beta, \alpha_1)$, to conclude with a standard Picard argument.

Let $\delta \in (0, 1)$. By the Young bound (4.3), we see that

$$|z^1|_{\mathcal{C}^{\tilde{\alpha}}} \lesssim |F_1|_\infty \delta^{1-\tilde{\alpha}} + (|G_1(z^1, z^2)|_{\mathcal{C}^{\tilde{\alpha} \wedge \alpha_2}} + |G_1|_\infty)|\mathfrak{h}^1|_{\mathcal{C}^{\tilde{\alpha}}}$$
$$\lesssim (1 + |z^1|_{\mathcal{C}^{\tilde{\alpha}}} + |z^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}})\delta^{\alpha_1 - \tilde{\alpha}},$$

where the prefactor is proportional to $M \triangleq |F_1|_\infty + |G|_\infty + |G|_{\text{Lip}}$. We may apply Lemma 4.2 to z^2 to further find

$$|z^1|_{\mathcal{C}^{\tilde{\alpha}}} \lesssim (1 + |z^1|_{\mathcal{C}^{\tilde{\alpha}}} + |z_0| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}})\delta^{\alpha_1 - \tilde{\alpha}}.$$

Here, we take the Hölder norms of z^1, z^2 over the interval $[0, \delta]$, whereas we use the full interval $[0, T]$ for $\mathfrak{h}^1$ and $\mathfrak{h}^2$. For $\delta > 0$ small enough, we therefore get

(4.8) $|z^1|_{\mathcal{C}^{\tilde{\alpha}}([0,\delta])} \lesssim (1 + |z_0| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}}).$

Combining this with Lemma 4.2, we can find a constant $C > 0$ such that

$$|z(\delta)| \le |z_0| + |z^1|_{\mathcal{C}^{\tilde{\alpha}}([0,\delta])} + |z^2|_{\mathcal{C}^{\alpha_2}([0,\delta])} \le C(1 + |z_0| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}}).$$

This bound can now be easily iterated and together with (4.8) we see that there is a (increased) constant C such that

$$|z^1|_{\mathcal{C}^{\tilde{\alpha}}([t,t+\delta])} \lesssim (1 + |z_t| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}})$$
$$\le C^{[\frac{t}{\delta}]+1}(1 + |z_0| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}})^{[\frac{t}{\delta}]+2}$$

for each $t \in [0, T - \delta]$. Since $|\cdot|_{\mathcal{C}^{\tilde{\alpha}}([0,T])} \le 2\delta^{\tilde{\alpha}-1} \sup_t |\cdot|_{\mathcal{C}^{\tilde{\alpha}}([t,t+\delta])}$, we get that

(4.9) $|z^1|_{\mathcal{C}^{\tilde{\alpha}}([0,T])} \le \dfrac{2C^{[\frac{t}{\delta}]+1}}{\delta^{1-\tilde{\alpha}}}(1 + |z_0| + |\mathfrak{h}^2|_{\mathcal{C}^{\alpha_2}})(1 + |\mathfrak{h}^1|_{\mathcal{C}^{\alpha_1}})^{[\frac{T}{\delta}]+2}.$

Local existence and uniqueness of a solution to (4.4) is a classical consequence of the Young bound. Indeed, if we define

$$A_\delta \triangleq \{f \in \mathcal{C}^\beta([0, \delta], \mathbb{R}^{d+n}): f(0) = z_0 \text{ and } |f|_{\mathcal{C}^\beta} \le 1\},$$

then, for $\delta > 0$ small enough, the operator $\mathcal{A}_\delta : A_\delta \to A_\delta$,

$$(\mathcal{A}_\delta z)(t) \triangleq z_0 + \int_0^t \begin{pmatrix} F_1(z(s)) \\ F_2(z(s)) \end{pmatrix} ds + \int_0^t G(z(s)) \, d\mathfrak{h}_s,$$

is contracting on a complete metric space. Abbreviating $\gamma \triangleq \alpha_1 \wedge \alpha_2$, this in turn follows from the well-known bounds

$$\left| \int_0^\cdot G(z(s)) \, d\mathfrak{h}_s \right|_{\mathcal{C}^\beta} \lesssim (|G|_{\text{Lip}} + |G|_\infty)(|z|_{\mathcal{C}^\beta} + 1)|\mathfrak{h}|_{\mathcal{C}^\gamma} \delta^{\gamma-\beta},$$

$$\left| \int_0^\cdot G(z(s)) - G(\bar{z}(s)) \, d\mathfrak{h}_s \right|_{\mathcal{C}^\beta} \lesssim (|G|_{\text{Lip}} + |DG|_{\text{Lip}})|\mathfrak{h}|_{\mathcal{C}^\gamma} \delta^{\gamma-\beta}|z - \bar{z}|_{\mathcal{C}^\beta},$$

$$\left| \int_0^\cdot \begin{pmatrix} F_1(z(s)) \\ F_2(z(s)) \end{pmatrix} ds \right|_{\mathcal{C}^\beta} \le (|F_1|_{\infty; B_{\delta^\beta}(z_0)} + |F_2|_{\infty; B_{\delta^\beta}(z_0)})\delta^{1-\beta},$$

$$\left| \int_0^\cdot \begin{pmatrix} F_1(z(s)) - F_1(\bar{z}(s)) \\ F_2(z(s)) - F_2(\bar{z}(s)) \end{pmatrix} ds \right|_{\mathcal{C}^\beta} \le (|F_1|_{\text{Lip}} + |F_2|_{\text{Lip}; B_{\delta^\beta}(z_0)})\delta|z - \bar{z}|_{\mathcal{C}^\beta}$$

for all $z, \bar{z} \in A_\delta$, where $|\cdot|_{\infty; A}$ and $|\cdot|_{\text{Lip}; A}$ denote the respective norms of the function restricted to the set A. Here, we also used that $\max(|z - z_0|_\infty, |\bar{z} - z_0|_\infty) \le \delta^\beta$ since $z, \bar{z} \in A_\delta$

by assumption. Consequently, there is a unique solution to (4.4) in $\mathcal{C}^\beta([0,\delta],\mathbb{R}^{d+n})$. Global existence and uniqueness follow from the a priori estimates (4.7) and (4.9) by a standard maximality argument. $\square$

We now bring the randomness back in the picture. To this end, let $\alpha > 0$, $p \geq 1$, and $T > 0$. We define the space

$$\mathcal{B}_{\alpha,p}([0,T],\mathbb{R}^d)$$
$$\triangleq \{X : [0,T] \times \Omega \to \mathbb{R}^d : X \text{ is } (\mathcal{F}_t)_{t\in[0,T]}\text{-adapted and } \|X\|_{\mathcal{B}_{\alpha,p}([0,T],\mathbb{R}^d)} < \infty\},$$

where we introduced the seminorm

$$\|X\|_{\mathcal{B}_{\alpha,p}([0,T],\mathbb{R}^d)} \triangleq \sup_{s\neq t\in[0,T]} \frac{\|X_t - X_s\|_{L^p}}{|t-s|^\alpha}.$$

If the terminal time T and the dimension d are clear from the context, we shall also write $\mathcal{B}_{\alpha,p}$ for brevity. By Kolmogorov's continuity theorem, we have the continuous embeddings

$$(4.10) \qquad L^p(\Omega, \mathcal{C}^{\alpha+\delta}([0,T],\mathbb{R}^d)) \hookrightarrow \mathcal{B}_{\alpha,p}([0,T],\mathbb{R}^d) \hookrightarrow L^p(\Omega, \mathcal{C}^{\alpha-\delta-\frac{1}{p}}([0,T],\mathbb{R}^d))$$

for any $\delta > 0$. Finally, let us also introduce the Besov-type space

$$W_0^{\alpha,\infty}([0,T],\mathbb{R}^d) \triangleq \{f : [0,T] \to \mathbb{R}^d : |f|_{\alpha,\infty} < \infty\},$$
$$|f|_{\alpha,\infty} \triangleq \sup_{t\in[0,T]} \left(|f(t)| + \int_0^t \frac{|f(t)-f(s)|}{|t-s|^{\alpha+1}} \, ds\right).$$

Nualart and Răşcanu proved the following classical result.

PROPOSITION 4.4 ([38], Theorem 2.1.II). *Let $f : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$ be bounded Lipschitz continuous and $g : \mathbb{R}^d \times \mathbb{R}^n \to L(\mathbb{R}^m, \mathbb{R}^d)$ be of class $\mathcal{C}_b^2$. Let $(Y_t)_{t\in[0,T]}$ be a stochastic process with sample paths in $\mathcal{C}^\gamma([0,T],\mathbb{R}^n)$ for some $\gamma > 1 - H$ and let B be an fBm with Hurst parameter $H > \frac{1}{2}$. Then there is a unique global solution to the equation*

$$X_t = X_0 + \int_0^t f(X_s, Y_s) \, ds + \int_0^t g(X_s, Y_s) \, dB_s$$

and, provided that $X_0 \in L^\infty$, we also have that

$$|X|_{\alpha,\infty} \in \bigcap_{p\geq 1} L^p$$

for each $\alpha < \frac{1}{2} \wedge \gamma$.

COROLLARY 4.5. *Fix the scale parameter $\varepsilon > 0$ and a terminal time $T > 0$. Let $\alpha < H \wedge \hat{H}$. There is a unique pathwise solution $(X^\varepsilon, Y^\varepsilon) \in \mathcal{C}^\alpha([0,T],\mathbb{R}^{d+n})$ to the slow-fast system (4.1)–(4.2). Moreover, for any $p \geq 1$ and any $\beta < \frac{1}{2} \wedge \hat{H}$, we have that*

$$(4.11) \qquad\qquad\qquad \|X^\varepsilon\|_{\mathcal{B}_{\beta,p}} < \infty.$$

PROOF.   The first part is an immediate consequence of Proposition 4.3. We stress that the bound (4.11) does not follow from our a priori estimate (4.9) since, by Fernique's theorem,

$$\mathbb{E}[\exp(a|B|_{\mathcal{C}^\beta}^2)] < \infty$$

if and only if $a > 0$ is sufficiently small. Instead, we employ Proposition 4.4: Since $Y^\varepsilon \in \mathcal{C}^{\hat{H}-}([0, T], \mathbb{R}^n)$ by Lemma 4.2, we see that, for each $\alpha < \frac{1}{2} \wedge \hat{H}$, $|X^\varepsilon|_{\alpha, \infty} \in \bigcap_{p \geq 1} L^p$, it is clear that

$$W_0^{\alpha, \infty}([0, T], \mathbb{R}^d) \hookrightarrow \mathcal{C}^{\alpha - \delta}([0, T], \mathbb{R}^d)$$

for any $\delta > 0$. Combine this with the continuous embedding (4.10) to conclude (4.11). □

REMARK 4.6. We finally record that Proposition 4.3 and Corollary 4.5 are the only places in the proof of Theorem 1.2 which require a linear growth of the drift b, see Condition 3.1. In fact, the remainder of the argument would still work, mutatis mutandis, under the weaker assumption of a polynomially growing drift, that is, $|b(x, y)| \lesssim 1 + |x|^N + |y|^N$ for some $N \in \mathbb{N}$. It is however unclear whether the solution to (4.1)–(4.2) exists globally in this case.

4.2. *Uniform bounds on the slow motions.* Our strategy in proving Theorem 1.2 is as follows: The integrals in (4.1) are approximated by suitable Riemann sums, on which we then aim to establish uniform bounds. These estimates translate into bounds on the integrals in view of Lê's stochastic sewing lemma [29].

Fix a terminal time $T > 0$ and let $\mathcal{S}^p$ denote the set of adapted two-parameter processes on the simplex with finite pth moments; in symbols:

$$\mathcal{S}^p \triangleq \{A : [0, T]^2 \times \Omega \to \mathbb{R}^d : A_{s,t} = 0 \text{ for } s \geq t \text{ and } A_{s,t} \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}) \text{ for all } s, t \geq 0\}.$$

Given $\eta, \bar{\eta} > 0$, we define the spaces

$$H_\eta^p \triangleq \left\{ A \in \mathcal{S}^p : \|A\|_{H_\eta^p} \triangleq \sup_{0 \leq s < t \leq T} \frac{\|A_{st}\|_{L^p}}{|s - t|^\eta} < \infty \right\},$$

$$\bar{H}_{\bar{\eta}}^p \triangleq \left\{ A \in \mathcal{S}^p : \|A\|_{\bar{H}_{\bar{\eta}}^p} \triangleq \sup_{0 \leq s < u < t \leq T} \frac{\|\mathbb{E}[\delta A_{\text{sut}} | \mathcal{F}_s]\|_{L^p}}{|s - t|^{\bar{\eta}}} < \infty \right\},$$

where we have set $\delta A_{s,u,t} \triangleq A_{s,t} - A_{s,u} - A_{u,t}$. With this notation we have the following version of the stochastic sewing lemma.

PROPOSITION 4.7 (Stochastic sewing lemma [29], Theorem 2.1 and Propostion 2.7). *Let $p \geq 2$, $\eta > \frac{1}{2}$, and $\bar{\eta} > 1$. Suppose that $A \in H_\eta^p \cap \bar{H}_{\bar{\eta}}^p$. Then, for every $0 \leq s \leq t \leq T$, the limit*

$$I_{s,t}(A) \triangleq \lim_{|P| \to 0} \sum_{[u,v] \in P} A_{u,v}$$

along partitions P of $[s, t]$ with mesh $|P| \triangleq \max_{[u,v] \in P} |v - u|$ tending to zero exists in L^p. The limiting process $I(A)$ is additive in the sense that $I_{s,u}(A) + I_{u,t}(A) = I_{s,t}(A)$ for all $0 \leq s \leq u \leq t \leq T$. Furthermore, there is a constant $C = C(p, \eta, \bar{\eta})$ such that

$$\|I_{s,t}(A)\|_{L^p} \leq C(\|A\|_{\bar{H}_{\bar{\eta}}^p} |t - s|^{\bar{\eta}} + \|A\|_{H_\eta^p} |t - s|^\eta)$$

for all $0 \leq s \leq t \leq T$. Moreover, if $\|\mathbb{E}[A_{s,t} | \mathcal{F}_s]\|_{L^p} \lesssim |t - s|^{\bar{\eta}}$, then $I(A) \equiv 0$.

Recall our notation of the fast motion's flow from (3.1) and (3.2), respectively. We are ultimately going to apply Proposition 4.7 with the two-parameter process

$$(4.12) \qquad A_{s,t}^\varepsilon \triangleq \int_s^t \left(g(X_s^\varepsilon, \bar{\Phi}_{s,r}^{X_s^\varepsilon}(\Phi_{0,s}^{X_s^\varepsilon}(Y_0))) - \bar{g}(X_s^\varepsilon) \right) dB_r, \quad 0 \leq s < t,$$

where, thanks to the conditional independence of the integrand and $\sigma(B_r - B_s, r \in [s, t])$ given $\mathcal{F}_s$, the integral is well defined in the (mixed) Wiener–Young sense as detailed in Section 4.2.1 below. There we also show that the integral $I(A^\varepsilon)$ constructed in Proposition 4.7 actually agrees with the Young integral in (4.1). It will be clear that our bounds on A^ε also apply to the Riemann summands for the drift term in (4.1), whence we exclude it from our considerations for now.

4.2.1. *A priori integral estimates.* We will use the notion of *mixed Wiener–Young integrals*: If $F : (s, t] \to L(\mathbb{R}^m, \mathbb{R}^d)$ is a (sufficiently regular) random function *independent* of $(\tilde{B}_r^s)_{r \in [0, t-s]}$, we can make the definition

$$
(4.13) \qquad \int_s^t F_r \, dB_r \triangleq \int_s^t F_r \dot{\tilde{B}}_{r-s}^s \, dr + \int_0^{t-s} F_{r+s} \, d\tilde{B}_r^s,
$$

where the integral with respect to $\tilde{B}^s$ is well defined in the Wiener sense (after all $\tilde{B}^s$ is a Gaussian process). The Hölder norm of negative exponent $-\kappa$, $\kappa \in [0, 1]$, is defined by

$$
|F|_{-\kappa} \triangleq \sup_{u, v \in (s, t]} \frac{1}{|v - u|^{1-\kappa}} \left| \int_u^v F_r \, dr \right|.
$$

Note that for $\kappa = 0$ we of course recover the usual sup-norm $|F|_\infty$.

In terms of this norm, one can then prove the following fundamental estimate on (4.13).

LEMMA 4.8 ([23], Lemma 3.4). *Let $2 \le p < q$. Fix $\kappa \in [0, H - \frac{1}{2})$ and $0 \le s \le t \le T$. Suppose that $F : (s, t] \to L(\mathbb{R}^m, \mathbb{R}^d)$ is independent of $(\tilde{B}_r^s)_{r \in [0, t-s]}$ and $|F|_{-\kappa} \in L^q$. Then one has the bound*

$$
\left\| \int_s^t F_r \, dB_r \right\|_{L^p} \lesssim \| |F|_{-\kappa} \|_{L^q} |t - s|^{H-\kappa},
$$

where the prefactor is independent of F and $0 \le s \le t \le T$.

We also have the the following estimate, which is a simple consequence of [23], Lemmas 3.10 and 3.12.

LEMMA 4.9. *Let $p \ge 2$ and $\alpha > 1 - H$. Let X be an $(\mathcal{F}_t)_{t \in [0, T]}$-adapted stochastic process with α-Hölder sample paths. Moreover assume that $X \in \mathcal{B}_{\alpha, p}$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a bounded Lipschitz continuous function. Then we have the following bound on the Young integral:*

$$
\left\| \int_s^t f(X_r) \, dB_r \right\|_{\mathcal{B}_{H,p}} \lesssim (|f|_\infty + |f|_{\text{Lip}})(1 + \|X\|_{\mathcal{B}_{\alpha, p}}),
$$

uniformly in $0 \le s < t \le T$.

It is of course fundamental for our argument that the "integral" furnished by Proposition 4.7 indeed coincides with the Young integral. This is ensured by the two lemmas below.

LEMMA 4.10. *Let $X = (X_t)_{t \in [0, T]}$ be a continuous process with values in $\mathbb{R}^d$. Let $b : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ be of linear growth and satisfy*

$$
\langle b(z, x) - b(z, y), x - y \rangle \le D - \kappa |x - y|^2 \quad \forall x, y \in \mathbb{R}^n, \forall z \in \mathbb{R}^d.
$$

Then, for any $p \ge 2$ and any random variable $Y \in L^p$, the following holds:

$$
\sup_{\varepsilon \in (0, 1]} \sup_{0 \le s \le t \le T} \| \Phi_{s,t}^X(Y) \|_{L^p} \lesssim 1 + \|Y\|_{L^p} + \sup_{0 \le t \le T} \|X_t\|_{L_p}.
$$

PROOF.   It is clear that we can assume $s = 0$ without loss of generality. Let Z^ε solve

$$Z_t^\varepsilon = Y - \frac{1}{\varepsilon} \int_0^t Z_s^\varepsilon \, ds + \frac{1}{\varepsilon \hat{H}} \sigma \hat{B}_t.$$

By (4.5), we have

$$\big|\Phi_{0,t}^X(Y) - Z_t^\varepsilon\big|^2 \lesssim \int_0^{\frac{t}{\varepsilon}} e^{-\kappa(\frac{t}{\varepsilon} - s)} \big(1 + |X_{\varepsilon s}| + |Z_{\varepsilon s}^\varepsilon|\big)^2 \, ds$$

for all $t \in [0, T]$. Since $(Z_{\varepsilon h}^\varepsilon)_{h \geq 0} \stackrel{d}{=} (Z_h^1)_{h \geq 0}$ and $\sup_{t \geq 0} \|Z_t^1\|_{L^p} \lesssim 1 + \|Y\|_{L^p}$, the lemma follows at once. $\square$

LEMMA 4.11.   *Let $H > \frac{1}{2}$ and let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function. Let $p > 2$ and $\alpha > 1 - H$. Let X be an $\mathbb{R}^d$-valued, $(\mathcal{F}_t)_{t \in [0,T]}$-adapted process with $\sup_{t \in [0,T]} \|X_t\|_{L^p} < \infty$ and sample paths in $\mathcal{C}^\alpha([0, T], \mathbb{R}^d)$. Let $Y_0 \in L^p$. Define*

$$A_{s,t} \triangleq \int_s^t h\big(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\big) \, dB_r,$$

where the integration is understood in the mixed Wiener–Young sense, see (4.13). If $A \in H_\eta^2 \cap \bar{H}_{\bar{\eta}}^2$ for some $\eta > \frac{1}{2}$ and $\bar{\eta} > 1$, then, for any $\varepsilon > 0$ and any $0 \leq s \leq t \leq T$,

$$\lim_{|P| \to 0} \sum_{[u,v] \in P([s,t])} A_{u,v} = \int_s^t h\big(X_r, \Phi_{0,r}^X(Y_0)\big) \, dB_r,$$

where the right-hand side is the Young integral.

PROOF.   We first note that, by Lemma 4.2, the process $\Phi_{0,\cdot}^X(Y_0)$ takes values in $\mathcal{C}^\beta([0, T], \mathbb{R}^d)$ for any $\beta < \hat{H}$. The pathwise Young integral $\int h(X_r, \Phi_{0,r}^X(Y_0)) \, dB_r$ is thus well defined and is given by the limit of the Riemann sums of

$$\tilde{A}_{s,t} \triangleq h\big(X_s, \Phi_{0,s}^X(Y_0)\big)(B_t - B_s)$$

along any sequence of partitions. By the last part of Proposition 4.7, it now suffices to show that $\|A_{s,t} - \tilde{A}_{s,t}\|_{L^2} \lesssim |t - s|^{\bar{\eta}}$ for some $\bar{\eta} > 1$.

To see this, we apply Lemma 4.8 with $\kappa = 0$ to find that, for each $\beta < \hat{H}$,

$$\|A_{s,t} - \tilde{A}_{s,t}\|_{L^2} = \left\| \int_s^t \big(h\big(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\big) - h\big(X_s, \Phi_{0,s}^X(Y_0)\big)\big) \, dB_r \right\|_{L^2}$$

$$\leq \left\| \sup_{s \leq r \leq t} \big|h\big(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\big) - h\big(X_s, \Phi_{0,s}^X(Y_0)\big)\big| \right\|_{L^p} |t - s|^H$$

$$\leq |h|_{\mathrm{Lip}} \big\| |\bar{\Phi}_{s,\cdot}^{X_s}(\Phi_{0,s}^X(Y_0))|_{\mathcal{C}^\beta} \big\|_{L^p} |t - s|^{H+\beta}.$$

Since $H + \hat{H} > 1$, we can conclude with Lemmas 4.2 and 4.10. $\square$

Our interest in Lemma 4.11 is of course in applying it to the slow motion (4.1) and the Riemann summands $A_{s,t}^\varepsilon$ defined in (4.12). We have already seen in Corollary 4.5 that $X^\varepsilon \in \bigcap_{p \geq 1} \mathcal{B}_{\alpha,p}$ for any $\alpha < \frac{1}{2} \wedge \hat{H}$. We are therefore left to check that $A^\varepsilon \in H_\eta^p \cap \bar{H}_{\bar{\eta}}^p$ for some $\eta > \frac{1}{2}, \bar{\eta} > 1$, and $p \geq 2$. Since these estimates are somewhat technically involved and require longer computations, we devote a subsection to each of the norms $\|\cdot\|_{H_\eta^p}$ and $\|\|\cdot\|\|_{\bar{H}_{\bar{\eta}}^p}$, respectively.

4.2.2. *Controlling the increment $A_{s,t}^\varepsilon$.* Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$. Recall that write $\bar{h}(x) = \int h(x, y) \pi^x (dy)$ for its average with respect to the first marginal of the invariant measure of the process $\bar{\Phi}^x$, see (3.2) and Definition 2.1. The following lemma exploits the convergence rates derived in Section 3. (The reader should observe that without further notice we assume that the conditions of Theorem 1.2 on the drift $b : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ are in place.)

LEMMA 4.12.   *Let $q > 1$. Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be a bounded measurable function and let $X, Y \in L^q$ be $\mathcal{F}_s$-measurable random variables. Then, for any $0 \le s \le t$, any $p \ge 2$, and any $\zeta < 1 - \hat{H}$, we have that*

$$\left\| \int_s^t \left(h\left(X, \bar{\Phi}_{s,r}^X(Y)\right) - \bar{h}(X)\right) dr \right\|_{L^p} \lesssim |h|_\infty \left(1 + \|Y\|_{L^q}^{\frac{1}{p}} + \|X\|_{L^q}^{\frac{1}{p}}\right) \varepsilon^{\frac{\zeta}{3p}} |t - s|^{1 - \frac{\zeta}{3p}}.$$

PROOF.   There is no loss of generality in assuming that $\bar{h} \equiv 0$. Notice also that the trivial estimate $\|\int_s^t h(X, \bar{\Phi}_{s,r}^X(Y)) dr\|_{L^\infty} \le |h|_\infty |t - s|$. By interpolation, we can therefore restrict ourselves to the case $p = 2$. Clearly,

$$\mathbb{E}\left[\left|\int_s^t h\left(X, \bar{\Phi}_{s,r}^X(Y)\right) dr\right|^2\right] = 2 \int_s^t \int_s^v \mathbb{E}\left[h\left(X, \bar{\Phi}_{s,r}^X(Y)\right) h\left(X, \bar{\Phi}_{s,v}^X(Y)\right)\right] dr \, dv.$$

For $r < v$ we condition the integrand on $\mathcal{F}_r$, and use Corollary 3.24(ii) together with Lemmas 3.7 and 4.10 to find

$$\left|\mathbb{E}\left[h\left(X, \bar{\Phi}_{s,r}^X(Y)\right) h\left(X, \bar{\Phi}_{s,v}^X(Y)\right)\right]\right|$$

$$\le |h|_\infty \mathbb{E}\left[\left|\mathbb{E}\left[h\left(X, \bar{\Phi}_{s,v}^X(Y)\right) | \mathcal{F}_r\right]\right|\right]$$

$$\lesssim |h|_\infty^2 \mathbb{E}\left[\left(1 + \|\varepsilon^{-\hat{H}} \bar{B}_{\varepsilon \cdot}^r\|_{\Omega_\zeta}\right)\left(1 + |\bar{\Phi}_{s,r}^X(Y)|\right)\right]\left(\frac{\varepsilon}{v - r}\right)^{\frac{\zeta}{3}}$$

$$\lesssim |h|_\infty^2 \left(1 + \|Y\|_{L^q} + \|X\|_{L^q}\right)\left(\frac{\varepsilon}{v - r}\right)^{\frac{\zeta}{3}}. \qquad \square$$

We can now establish the required estimate on the H_η^p-norm of $A_{s,t}$.

PROPOSITION 4.13.   *Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be bounded measurable and let X be an $(\mathcal{F}_t)_{t \in [0,T]}$-adapted, continuous process with $\sup_{t \in [0,T]} \|X_t\|_{L^q} < \infty$ for some $q \ge 1$. Define*

$$A_{s,t} \triangleq \int_s^t \left[h\left(X_s, \bar{\Phi}_{s,r}^{X_s}\left(\Phi_{0,s}^X(Y_0)\right)\right) - \bar{h}(X_s)\right] dB_r, \quad 0 \le s \le t \le T,$$

in the mixed Wiener–Young sense, see (4.13). Let $\kappa \in (0, H - \frac{1}{2})$ and set $\eta = H - \kappa$. Then $A \in H_\eta^p$ for each $p \ge 2$, and any $\varepsilon > 0$. Moreover, there is a $\gamma > 0$ such that

$$\|A\|_{H_\eta^p} \lesssim |h|_\infty \left(1 + \sup_{0 \le t \le T} \|X_t\|_{L^q}\right) \varepsilon^\gamma.$$

PROOF.   Again, we may assume that $\bar{h} \equiv 0$ without any loss of generality. Since X is $(\mathcal{F}_t)_{t \in [0,T]}$-adapted, we can use Lemma 4.8 to obtain that, for $\tilde{q} > p$ and $\kappa \in [0, H - \frac{1}{2})$,

$$\|A_{s,t}\|_{L^p} \lesssim \||h\left(X_s, \bar{\Phi}_{s,\cdot}^{X_s}\left(\Phi_{0,s}^X(Y_0)\right)\right)|_{-\kappa}\|_{L^{\tilde{q}}} |t - s|^{H - \kappa}.$$

By Lemmas 4.10 and 4.12, we obtain

$$\left\| \int_u^v h\left(X_s, \bar{\Phi}_{s,r}^{X_s}\left(\Phi_{0,s}^X(Y_0)\right)\right) dr \right\|_{L^{\tilde{q}}} \lesssim |h|_\infty \left(1 + \|Y_0\|_{L^q}^{\frac{1}{\tilde{q}}} + \sup_{0 \le r \le s} \|X_r\|_{L^q}^{\frac{1}{\tilde{q}}}\right) \varepsilon^{\frac{\zeta}{3\tilde{q}}} |v - u|^{1 - \frac{\zeta}{3\tilde{q}}}$$

for all $u, v \in [s, t]$ and any $\zeta < 1 - \hat{H}$. Therefore, Kolmogorov's continuity theorem shows that

$$\big\| |h\big(X_s, \bar{\Phi}_{s,\cdot}^{X_s}(\Phi_{0,s}^X(Y_0))\big)|_{-\kappa}\big\|_{L^{\tilde{q}}} \lesssim |h|_\infty \Big(1 + \|Y_0\|_{L^q}^{\frac{1}{\tilde{q}}} + \sup_{0 \le t \le T} \|X_t\|_{L^q}^{\frac{1}{\tilde{q}}}\Big)\varepsilon^{\frac{\zeta}{3\tilde{q}}},$$

provided that we choose $\tilde{q} > \kappa^{-1}(1 + \frac{\zeta}{3})$, and the final result follows. □

4.2.3. *Continuity of the invariant measures.* Let $\varepsilon > 0$ and $s < t$. We write

$$P([s, t]; \varepsilon) \triangleq \begin{cases} \Big\{\Big\{s + k\varepsilon : k = 0, \ldots, \Big[\dfrac{t-s}{\varepsilon}\Big]\Big\} \cup \{t\} & t - s - \varepsilon\Big[\dfrac{t-s}{\varepsilon}\Big] \ge \dfrac{\varepsilon}{2}, \\ \Big\{s + k\varepsilon : k = 0, \ldots, \Big[\dfrac{t-s}{\varepsilon}\Big] - 1\Big\} \cup \Big\{t - \dfrac{\varepsilon}{2}, t\Big\} & t - s - \varepsilon\Big[\dfrac{t-s}{\varepsilon}\Big] < \dfrac{\varepsilon}{2}. \end{cases}$$

Notice that the distance between two subsequent points $(t_i, t_{i+1}) \in P([s, t]; \varepsilon)$ satisfies $|t_{i+1} - t_i| \in [\frac{\varepsilon}{2}, \varepsilon]$. Recall from Condition 3.1 that the drift b is assumed to be locally Lipschitz uniformly with respect to the second argument. We write

$$|b|_{\mathrm{Lip}; K} \triangleq \sup_{\substack{|x_1|, |x_2| \le K \\ y \in \mathbb{R}^n}} |b(x_1, y) - b(x_2, y)|$$

for $K > 0$. In order to keep the statements of the next lemmas concise, we shall freely absorb quantities independent of $0 \le s \le t$ and $\varepsilon \in (0, 1]$ into the prefactor hidden beneath $\lesssim$.

LEMMA 4.14. *Let $p \ge 1$ and suppose that $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $X, \bar{X} \in L^\infty$, and $Y \in L^p$ be $\mathcal{F}_s$-measurable random variables. Then*

$$\big\|\bar{\Phi}_{s,t}^X(Y) - \bar{\Phi}_{s,t}^{\bar{X}}(Y)\big\|_{L^p} \lesssim \|X - \bar{X}\|_{L^p}.$$

PROOF.   We abbreviate $\Lambda \triangleq \Lambda(\kappa, R, p)$ and observe that, for any $s \le u \le r$,

$$\begin{aligned} \frac{d}{dr}\big|\bar{\Phi}_{u,r}^X(Y) - \bar{\Phi}_{u,r}^{\bar{X}}(Y)\big|^2 &= \frac{2}{\varepsilon}\langle b(X, \bar{\Phi}_{u,r}^X(Y)) - b(\bar{X}, \bar{\Phi}_{u,r}^{\bar{X}}(Y)), \bar{\Phi}_{u,r}^X(Y) - \bar{\Phi}_{u,r}^{\bar{X}}(Y)\rangle \\ &\le \frac{2(\Lambda + 1)}{\varepsilon}\big|\bar{\Phi}_{u,r}^X(Y) - \bar{\Phi}_{u,r}^{\bar{X}}(Y)\big|^2 \\ &\quad + \frac{|b|_{\mathrm{Lip}; \|X\|_{L^\infty} \vee \|\bar{X}\|_{L^\infty}}^2}{2\varepsilon}|X - \bar{X}|^2 \end{aligned}$$

with probability 1. It follows that

(4.14) $$\big|\bar{\Phi}_{u,r}^X(Y) - \bar{\Phi}_{u,r}^{\bar{X}}(Y)\big| \lesssim |b|_{\mathrm{Lip}; \|X\|_{L^\infty} \vee \|\bar{X}\|_{L^\infty}} e^{(\Lambda+1)\frac{|r-u|}{\varepsilon}}|X - \bar{X}|.$$

This bound is of course only useful on a time interval with length of order ε. We therefore expand

$$\big\|\bar{\Phi}_{s,t}^X(Y) - \bar{\Phi}_{s,t}^{\bar{X}}(Y)\big\|_{L^p} \le \sum_{(t_i, t_{i+1}) \in P([s,t];\varepsilon)} \big\|\bar{\Phi}_{t_{i+1},t}^{\bar{X}}(\bar{\Phi}_{s,t_{i+1}}^X(Y)) - \bar{\Phi}_{t_i,t}^{\bar{X}}(\bar{\Phi}_{s,t_i}^X(Y))\big\|_{L^p}.$$

Corollary 3.15 shows that

$$\big\|\bar{\Phi}_{t_{i+1},t}^{\bar{X}}(\bar{\Phi}_{s,t_{i+1}}^X(Y)) - \bar{\Phi}_{t_i,t}^{\bar{X}}(\bar{\Phi}_{s,t_i}^X(Y))\big\|_{L^p} \lesssim \big\|\bar{\Phi}_{s,t_{i+1}}^X(Y) - \bar{\Phi}_{t_i,t_{i+1}}^{\bar{X}}(\bar{\Phi}_{s,t_i}^X(Y))\big\|_{L^p} e^{-c\frac{|t-t_{i+1}|}{\varepsilon}}$$

$$\lesssim \|X - \bar{X}\|_{L^p} e^{-c\frac{|t-t_{i+1}|}{\varepsilon}},$$

where the last inequality uses (4.14) together with $|t_{i+1} - t_i| \asymp \varepsilon$. Consequently,

$$\left\| \bar{\Phi}_{s,t}^X(Y) - \bar{\Phi}_{s,t}^{\bar{X}}(Y) \right\|_{L^p} \lesssim \|X - \bar{X}\|_{L^p} \sum_{(t_i, t_{i+1}) \in P([s,t]; \varepsilon)} e^{-c\frac{|t - t_{i+1}|}{\varepsilon}} \lesssim \|X - \bar{X}\|_{L^p}$$

uniformly in $0 \leq s \leq t$ and $\varepsilon \in (0, 1]$. $\square$

Lemma 4.14 implies the local Lipschitz continuity of the invariant measure π^x in the parameter $x \in \mathbb{R}^d$:

PROPOSITION 4.15.   *Let $p \geq 1$ and $K > 0$. Suppose that $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$. Then*

$$\mathcal{W}^p\left(\pi^{x_1}, \pi^{x_2}\right) \lesssim |x_1 - x_2|,$$

uniformly for $|x_1|, |x_2| \leq K$.

PROOF.   Owing to Theorem 1.3, it follows that

$$\mathcal{W}^p\left(\pi^{x_1}, \pi^{x_2}\right) \leq \limsup_{\varepsilon \to 0} \left\| \bar{\Phi}_{0,1}^{x_1}(0) - \bar{\Phi}_{0,1}^{x_2}(0) \right\|_{L^p}$$

and we conclude with Lemma 4.14. $\square$

The simple proof of the following corollary is left to the reader.

COROLLARY 4.16.   *Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$ be Lipschitz continuous. Then $\bar{h} : \mathbb{R}^d \to \mathbb{R}^d$ is locally Lipschitz.*

4.2.4. *Controlling the second order increment $\delta A_{s,u,t}^\varepsilon$.*   Uniform bounds on the second order increments are difficult to obtain even for the Markovian fast dynamic. The first technical estimate of this subsection is the following.

LEMMA 4.17.   *Let $1 \leq p < q$ and suppose that $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function with $\bar{h} \equiv 0$. Suppose that $X, \bar{X} \in L^\infty$ and $Y \in L^q$ are $\mathcal{F}_s$-measurable random variables. Then, for any $\rho \in (0, 1)$, there is a $\gamma > 0$ such that*

$$\left\| \mathbb{E}\left[h\left(X, \bar{\Phi}_{s,t}^X(Y)\right) - h\left(\bar{X}, \bar{\Phi}_{s,t}^{\bar{X}}(Y)\right) | \mathcal{F}_s\right] \right\|_{L^p} \lesssim |h|_{\text{Lip}}\left(1 + \|Y\|_{L^q}\right)\|X - \bar{X}\|_{L^p}^\rho \left(1 \wedge \frac{\varepsilon^\gamma}{|t - s|^\gamma}\right).$$

PROOF.   By Corollary 3.24 i and Hölder's inequality, we certainly have

$$(4.15) \quad \left\| \mathbb{E}\left[h\left(X, \bar{\Phi}_{s,t}^X(Y)\right) - h\left(\bar{X}, \bar{\Phi}_{s,t}^{\bar{X}}(Y)\right) | \mathcal{F}_s\right] \right\|_{L^p} \lesssim |h|_{\text{Lip}}\left(1 + \|Y\|_{L^q}\right)\left(1 \wedge \frac{\varepsilon^\zeta}{|t - s|^\zeta}\right).$$

On the other hand, by the continuity lemma (Lemma 4.14),

$$\left\| \mathbb{E}\left[h\left(X, \bar{\Phi}_{s,t}^X(Y)\right) - h\left(\bar{X}, \bar{\Phi}_{s,t}^{\bar{X}}(Y)\right) | \mathcal{F}_s\right] \right\|_{L^p}$$

$$\lesssim |h|_{\text{Lip}}\left(\|X - \bar{X}\|_{L^p} + \left\| \bar{\Phi}_{s,t}^X(Y) - \bar{\Phi}_{s,t}^{\bar{X}}(Y) \right\|_{L^p}\right) \lesssim |h|_{\text{Lip}}\|X - \bar{X}\|_{L^p}.$$

Finally, we interpolate this bound with (4.15). $\square$

Our remaining task is to derive an estimate on the distance between $\Phi_{s,t}^Z$ and $\bar{\Phi}_{s,t}^{Z_s}$. This is based on the following version of Lemma 4.14.

LEMMA 4.18.   *Let $p \geq 1$ and suppose that $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $Y \in L^p$ be $\mathcal{F}_s$-measurable and Z be a continuous process. Assume that $|Z|_\infty \in L^\infty$. Then*

$$\left\| \bar{\Phi}_{s,t}^{Z_s}(Y) - \Phi_{s,t}^Z(Y) \right\|_{L^p} \lesssim \left\| \sup_{r \in [s,t]} |Z_r - Z_s| \right\|_{L^p}.$$

PROOF.   The reader can easily check that the very same argument we gave at the beginning of the proof of Lemma 4.14 also shows that, for $0 \leq s \leq u \leq r \leq T$,

$$\left| \bar{\Phi}_{u,r}^{Z_s}(Y) - \Phi_{u,r}^Z(Y) \right| \lesssim |b|_{\text{Lip}; \||Z|_\infty\|_{L^\infty}} \left(\int_{\frac{u}{\varepsilon}}^{\frac{r}{\varepsilon}} e^{2(\Lambda+1)(\frac{r}{\varepsilon}-v)} |Z_{\varepsilon v} - Z_s|^2 \, dv \right)^{\frac{1}{2}}$$

$$\lesssim \sup_{v \in [u,r]} |Z_v - Z_s| e^{(\Lambda+1)\frac{|r-u|}{\varepsilon}}.$$

The asserted bound then follows along the same lines as Lemma 4.14.   □

The following estimate is now an easy consequence.

LEMMA 4.19.   *Let $p \geq 1$ and suppose that $b(x, \cdot) \in S_p(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be Lipschitz continuous. Assume furthermore that X and Y are $\mathcal{F}_u$- and $\mathcal{F}_s$-measurable random variables, respectively. Moreover, let $Z \in \mathcal{B}_{\alpha, p}([0, T], \mathbb{R}^d)$ for some $\alpha > 0$ and assume that $|Z|_\infty \in L^\infty$. Then*

(4.16)
$$\left\| \mathbb{E}\left[h\left(X, \bar{\Phi}_{u,t}^X(\bar{\Phi}_{s,u}^{Z_s}(Y))\right) - h\left(X, \bar{\Phi}_{u,t}^X(\Phi_{s,u}^Z(Y))\right) | \mathcal{F}_u \right] \right\|_{L^p}$$
$$\lesssim |h|_{\text{Lip}} \|Z\|_{\mathcal{B}_{\alpha,p}} |u - s|^\alpha e^{-c\frac{|t-u|}{\varepsilon}}.$$

PROOF.   By Corollary 3.15, we have that

$$\left\| \mathbb{E}\left[h\left(X, \bar{\Phi}_{u,t}^X(\bar{\Phi}_{s,u}^{Z_s}(Y))\right) - h\left(X, \bar{\Phi}_{u,t}^X(\Phi_{s,u}^Z(Y))\right) | \mathcal{F}_u \right] \right\|_{L^p}$$
$$\lesssim |h|_{\text{Lip}} \left\| \bar{\Phi}_{s,u}^{Z_s}(Y) - \Phi_{s,u}^Z(Y) \right\|_{L^p} e^{-c\frac{|t-u|}{\varepsilon}}.$$

By Lemma 4.18,

$$\left\| \bar{\Phi}_{s,u}^{Z_s}(Y) - \Phi_{s,u}^Z(Y) \right\|_{L^p} \lesssim \|Z\|_{\mathcal{B}_{\alpha,p}} |u - s|^\alpha.   □$$

Finally, we can establish the second estimate needed for the application of Proposition 4.7.

PROPOSITION 4.20.   *Let $1 \leq p < q$ and suppose that $b(x, \cdot) \in S_q(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function. Assume that $X \in \mathcal{B}_{\alpha, p}$ for some $\alpha > 1 - H$ and $|X|_\infty \in L^\infty$. Define*

$$A_{s,t} \triangleq \int_s^t \left(h\left(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\right) - \bar{h}(X_s) \right) dB_r,$$

in the mixed Wiener–Young sense, see (4.13). Then $A \in \bar{H}_{\bar{\eta}}^p$ for any $\bar{\eta} < \alpha + H$ and any $\varepsilon > 0$. Moreover, there is a $\gamma > 0$ such that

$$\|A\|_{\bar{H}_{\bar{\eta}}^p} \lesssim |h|_{\text{Lip}}(1 \vee \||X|_\infty\|_{L^\infty})(1 \vee \|X\|_{\mathcal{B}_{\alpha,p}})\varepsilon^\gamma.$$

PROOF.   Fix $1 < \bar{\eta} < \alpha + H$ and choose $\rho \in (0, 1)$ such that $\bar{\eta} < H + \alpha\rho^2$ and $p \leq \rho q$. Since $|X|_\infty \in L^\infty$, owing to Corollary 4.16 we may assume that $\bar{h} \equiv 0$ without any loss of generality. Recall that $\delta A_{s,u,t} = A_{s,t} - A_{s,u} - A_{u,t}$, so

$$\delta A_{s,u,t} = \int_u^t \left(h\left(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\right) - h\left(X_u, \bar{\Phi}_{u,r}^{X_u}(\Phi_{0,u}^X(Y_0))\right) \right) dB_r.$$

We condition on $\mathcal{F}_u$ instead of $\mathcal{F}_s$. This gives

$$\|\mathbb{E}[\delta A_{s,u,t} \mid \mathcal{F}_s]\|_{L^p}$$

$$\leq \left\| \int_u^t \mathbb{E}\big[h\big(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\big) - h\big(X_u, \bar{\Phi}_{u,r}^{X_u}(\Phi_{0,u}^X(Y_0))\big)|\mathcal{F}_u\big]\dot{B}_r^u \, dr \right\|_{L^p}$$

$$\leq (\mathrm{I}) + (\mathrm{II})$$

with

$$(\mathrm{I}) \triangleq \left\| \int_u^t \mathbb{E}\big[h\big(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))\big) - h\big(X_u, \bar{\Phi}_{u,r}^{X_u}(\bar{\Phi}_{s,u}^{X_s}(\Phi_{0,s}^X(Y_0)))\big)|\mathcal{F}_u\big]\dot{B}_r^u \, dr \right\|_{L^p},$$

$$(\mathrm{II}) \triangleq \left\| \int_u^t \mathbb{E}\big[h\big(X_u, \bar{\Phi}_{u,r}^{X_u}(\bar{\Phi}_{s,u}^{X_s}(\Phi_{0,s}^X(Y_0)))\big) - h\big(X_u, \bar{\Phi}_{u,r}^{X_u}(\Phi_{0,u}^X(Y_0))\big)|\mathcal{F}_u\big]\dot{B}_r^u \, dr \right\|_{L^p}.$$

These terms are now bounded individually. Let us begin with the bound on (I). Thanks to Lemma 4.10, this term falls in the regime of Lemma 4.17. By Hölder's inequality, we therefore find

$$(\mathrm{I}) \lesssim |h|_{\mathrm{Lip}} \|X_s - X_u\|_{L^\infty}^{\rho(1-\rho)} \|X_s - X_u\|_{L^p}^{\rho^2} \int_u^t \|\dot{B}_r^u\|_{L^{\frac{p}{1-\rho}}} \left(1 \wedge \frac{\varepsilon^\gamma}{|r-u|^\gamma}\right) dr$$

$$\lesssim |h|_{\mathrm{Lip}} \||X|_\infty\|_{L^\infty}^{\rho(1-\rho)} \|X\|_{\mathcal{B}_{\alpha,p}}^{\rho^2} \varepsilon^\delta |t-s|^{\bar{\eta}}$$

for $\delta > 0$ sufficiently small. Here, the last inequality used that, for any $p \geq 1$, $\|\dot{B}_r^u\|_{L^p} \lesssim |r-u|^{H-1}$ together with the elementary fact

$$\int_u^t \frac{1}{|r-u|^{1-H}} \left(1 \wedge \frac{\varepsilon^\gamma}{|r-u|^\gamma}\right) dr \lesssim \varepsilon^\delta |t-u|^{H-\delta}$$

for any $\delta \in (0, \gamma]$.

The term (II) can be handled similarly in view of Lemma 4.19. $\square$

4.3. *Proof of Theorem 1.2.* The estimates of the previous two subsection furnish the following fundamental estimates.

PROPOSITION 4.21. *Let $2 \leq p < q$ and suppose that $b(x, \cdot) \in S_q(\kappa, R)$ for all $x \in \mathbb{R}^d$. Let $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ be a bounded Lipschitz continuous function. Assume that there is an $\alpha > 1 - H$ such that X has α-Hölder sample paths and $X \in \mathcal{B}_{\alpha,p}$. If, in addition, $|X|_\infty \in L^\infty$, then, for any $\eta < H$ and any $\bar{\eta} < \alpha + H$, there is a $\gamma > 0$ such that*

$$(4.17) \quad \left\| \int_0^\cdot \big(h(X_r, \Phi_{0,s}^X(Y_0)) - \bar{h}(X_r)\big) dB_r \right\|_{\mathcal{B}_{\eta,p}}$$

$$\lesssim (|h|_\infty + |h|_{\mathrm{Lip}})(1 + \||X|_\infty\|_{L^\infty})(1 + \|X\|_{\mathcal{B}_{\alpha,p}})\varepsilon^\gamma,$$

and

$$(4.18) \quad \left\| \int_0^\cdot h(X_r, \Phi_{0,r}^X(Y_0)) dB_r \right\|_{\mathcal{B}_{\eta,p}} \lesssim (|h|_\infty + |h|_{\mathrm{Lip}})(1 + \||X|_\infty\|_{L^\infty})(1 + \|X\|_{\mathcal{B}_{\alpha,p}}),$$

uniformly in $0 \leq s < t \leq T$ and $\varepsilon \in (0, 1]$. Here, the integrals are both taken in the Young sense.

PROOF. First note that, by Lemma 4.11, the Young integrals in both (4.17) and (4.18) coincide with the processes $I(A^i)$ obtained by "sewing" the Riemann summands

$$A_{s,t}^1 \triangleq \int_s^t \big(h(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))) - \bar{h}(X_s)\big) dB_r, \qquad A_{s,t}^2 \triangleq \int_s^t h(X_s, \bar{\Phi}_{s,r}^{X_s}(\Phi_{0,s}^X(Y_0))) dB_r,$$

where integration is now understood in the mixed Wiener–Young sense (see (4.13)), with the help of Proposition 4.7. Consequently, the estimate (4.17) follows immediately from combining Propositions 4.13 and 4.20. Owing to Lemma 4.9 and Corollary 4.16, (4.18) is then an easy consequence of the first bound. $\quad\square$

For $\varepsilon > 0$ and $M > 0$, let us define the $(\mathcal{F}_t)_{t\geq 0}$-stopping time $\tau_M^\varepsilon \triangleq \inf\{t \geq 0 : |X_t^\varepsilon| > M\}$. Applying the previous proposition to the slow-fast system (4.1)–(4.2), we can deduce relative compactness of the stopped slow motion $X^{\varepsilon,M} \triangleq X_{\cdot \wedge \tau_M^\varepsilon}^\varepsilon$.

COROLLARY 4.22. *Consider the slow-fast system* (4.1)–(4.2) *with Condition* 3.1 *in place. Let* $\beta < \frac{1}{2} \wedge \hat{H}$ *and* $p \geq 2$. *Suppose that there are* $\kappa, R > 0$ *and* $q > p$ *such that* $b(x, \cdot) \in S_q(\kappa, R)$ *for each* $x \in \mathbb{R}^d$. *Then, for any* $M > 0$,

$$\sup_{\varepsilon \in (0,1]} \|X^{\varepsilon,M}\|_{\mathcal{B}_{\beta,p}} < \infty.$$

PROOF. Recall from Corollary 4.5 that, for each $\varepsilon > 0$, there is a unique global solution X^ε to (4.1) with values in $\mathcal{C}^\alpha([0, T], \mathbb{R}^d)$ for some $\alpha > 1 - H$. Moreover, since the Hölder norm of the stopped solution $X^{\varepsilon,M}$ is controlled by the Hölder norm of X^ε, the argument of Corollary 4.5 also shows that $\|X^{\varepsilon,M}\|_{\mathcal{B}_{\beta,p}} < \infty$ for each $\beta < \frac{1}{2} \wedge \hat{H}$ and $p \geq 1$. Employing Proposition 4.21, we obtain that, for any $\gamma < H - \beta$ and any $\delta \in (0, T]$,

$$\|X^{\varepsilon,M} \upharpoonright_{[0,\delta]}\|_{\mathcal{B}_{\beta,p}} \lesssim (|g|_\infty + |g|_{\mathrm{Lip}})(1 + \|X^{\varepsilon,M} \upharpoonright_{[0,\delta]}\|_{\mathcal{B}_{\beta,p}})\delta^\gamma + |f|_\infty \delta^{1-\beta},$$

uniformly in $\varepsilon \in (0, 1]$. Hence, choosing $\delta > 0$ sufficiently small, the proof is concluded by a standard iteration argument. $\quad\square$

Now we can finish the proof of Theorem 1.2 by localizing the argument of Hairer and Li. To this end, we rely on the following deterministic residue lemma.

LEMMA 4.23 (Residue lemma). *Let* $F : \mathbb{R}^d \to \mathbb{R}^d$ *be Lipschitz continuous,* $G : \mathbb{R}^d \to L(\mathbb{R}^m, \mathbb{R}^d)$ *be of class* $\mathcal{C}_b^2$, *and* $\mathfrak{h} \in \mathcal{C}^\alpha([0, T], \mathbb{R}^n)$ *for some* $\alpha > \frac{1}{2}$. *Moreover, let* $Z, \bar{Z} \in \mathcal{C}^{\tilde{\alpha}}([0, T], \mathbb{R}^d)$ *for some* $\tilde{\alpha} \in (1 - \alpha, \alpha]$ *with* $Z_0 = \bar{Z}_0$. *Then there is a constant* C *depending only on* F, G, *and the terminal time* T *such that*

$$|z - \bar{z}|_{\mathcal{C}^{\tilde{\alpha}}} \leq C \exp(C|\mathfrak{h}|_{\mathcal{C}^\alpha}^{\frac{1}{\alpha}} + C|Z|_{\mathcal{C}^{\tilde{\alpha}}}^{\frac{1}{\tilde{\alpha}}} + C|\bar{Z}|_{\mathcal{C}^{\tilde{\alpha}}}^{\frac{1}{\tilde{\alpha}}})|Z - \bar{Z}|_{\mathcal{C}^{\tilde{\alpha}}},$$

where z *and* $\bar{z}$ *are the solutions to the equations*

$$z_t = Z_t + \int_0^t F(z_s)\,ds + \int_0^t G(z_s)\,d\mathfrak{h}_s, \qquad \bar{z}_t = \bar{Z}_t + \int_0^t F(z_s)\,ds + \int_0^t F(\bar{z}_s)\,d\mathfrak{h}_s.$$

Albeit the statement of Lemma 4.23 is slightly stronger than [23], Lemma 2.2, it is straightforward to show that the very same proof still applies. We therefore omit the details and finally turn to the proof of the main result of this article.

PROOF OF THEOREM 1.2. First observe that, by the assumptions of the theorem and Corollary 4.16, there exists a unique global solution to the averaged equation (1.3), see [35, 36, 38]. We fix $\bar{\alpha} \in (\alpha, H)$ with $(\bar{\alpha} - \alpha)^{-1} < p$. Choose $\beta \in (1 - H, \hat{H} \wedge \frac{1}{2})$. By Corollary 4.22, $\sup_{\varepsilon \in (0,1]} \|X^{\varepsilon,M}\|_{\mathcal{B}_{\beta,p}} < \infty$ for each $M > 0$. Consequently, by Proposition 4.21,

we deduce that

$$\left\| \int_0^\cdot \left(g(X_r^{\varepsilon,M}, \Phi_{0,r}^{X^{\varepsilon,M}}(Y_0)) - \bar{g}(X_r^{\varepsilon,M}) \right) dB_r \right\|_{\mathcal{B}_{\tilde{\alpha},p}} \lesssim \varepsilon^\gamma,$$

$$\left\| \int_0^\cdot \left(f(X_r^{\varepsilon,M}, \Phi_{0,r}^{X^{\varepsilon,M}}(Y_0)) - \bar{f}(X_r^{\varepsilon,M}) \right) dr \right\|_{\mathcal{B}_{\tilde{\alpha},p}} \lesssim \varepsilon^\gamma.$$

Therefore, $\|\hat{X}^{\varepsilon,M} - \bar{X}^{\varepsilon,M}\|_{\mathcal{B}_{\tilde{\alpha},p}} \lesssim \varepsilon^\gamma$, where

$$\hat{X}_t^{\varepsilon,M} \triangleq X_0 + \int_0^t f(X_r^{\varepsilon,M}, \Phi_{0,r}^{X^{\varepsilon,M}}(Y_0)) \, dr + \int_0^t g(X_r^{\varepsilon,M}, \Phi_{0,r}^{X^{\varepsilon,M}}(Y_0)) \, dB_r,$$

$$\bar{X}_t^{\varepsilon,M} \triangleq X_0 + \int_0^t \bar{f}(X_r^{\varepsilon,M}) \, dr + \int_0^t \bar{g}(X_r^{\varepsilon,M}) \, dB_r.$$

In particular, $|\hat{X}^{\varepsilon,M} - \bar{X}^{\varepsilon,M}|_{\mathcal{C}^\alpha} \to 0$ in probability by the embedding (4.10). Note also the decomposition

$$X_t^{\varepsilon,M} = \hat{X}_t^{\varepsilon,M} - \bar{X}_t^{\varepsilon,M} + X_0 + \int_0^t \bar{f}(X_r^\varepsilon) \, dr + \int_0^t \bar{g}(X_r^\varepsilon) \, dB_r, \quad t \in [0, \tau_M^\varepsilon \wedge T],$$

whence Lemma 4.23 furnishes the bound

(4.19) $|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0, \tau_M^\varepsilon \wedge T])} \le C \exp(C|B|_{\mathcal{C}^\alpha}^{\frac{1}{\alpha}} + C|\hat{X}^{\varepsilon,M} - \bar{X}^{\varepsilon,M}|_{\mathcal{C}^\alpha}^{\frac{1}{\alpha}})|\hat{X}^{\varepsilon,M} - \bar{X}^{\varepsilon,M}|_{\mathcal{C}^\alpha}.$

As we have seen above, for each $M > 0$, the right-hand side goes to 0 in probability as $\varepsilon \to 0$. Hence, we also have that $|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0, \tau_M^\varepsilon \wedge T])} \to 0$ in probability.

On the other hand, note that

$$\mathbb{P}(\tau_M^\varepsilon < T) \le \mathbb{P}\left(\sup_{t \in [0, \tau_M^\varepsilon]} |X_t^\varepsilon| \ge M, \tau_M^\varepsilon < T \right) \le \mathbb{P}\left(|X^\varepsilon|_{\mathcal{C}^\gamma([0, \tau_M^\varepsilon \wedge T])} \ge T^{-\gamma}(M - \|X_0\|_{L^\infty}) \right)$$

(4.20) $\le \mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\gamma([0, \tau_M^\varepsilon \wedge T])} \ge T^{-\gamma}(M - \|X_0\|_{L^\infty}) - |\bar{X}|_{\mathcal{C}^\gamma([0,T])} \right)$

$$\le \mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\gamma([0, \tau_M^\varepsilon \wedge T])} \ge 1 \right) + \mathbb{P}\left(|\bar{X}|_{\mathcal{C}^\gamma([0,T])} > T^{-\gamma}(M - \|X_0\|_{L^\infty}) - 1 \right)$$

for each $\gamma > 0$. By Proposition 4.4, we know that $|\bar{X}|_{\mathcal{C}^\gamma([0,T])} \in L^1$ provided that $\gamma < \frac{1}{2}$. We fix such a γ.

It is now easy to finish the proof. Let $\delta_1, \delta_2 \in (0, 1)$ be given. Then we can find a $M > 0$ such that

$$\mathbb{P}\left(|\bar{X}|_{\mathcal{C}^\gamma([0,T])} > T^{-\gamma}(M - \|X_0\|_{L^\infty}) - 1 \right) \le \frac{\delta_2}{2}.$$

For this M, we can also find an $\varepsilon_0 > 0$ such that

$$\mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0, \tau_M^\varepsilon \wedge T])} > \delta_1 \right) \le \frac{\delta_2}{4} \quad \forall \varepsilon \in (0, \varepsilon_0).$$

The estimate (4.20) therefore yields that

$$\mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0,T])} > \delta_1 \right) \le \mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0, \tau_M^\varepsilon \wedge T])} > \delta_1, \tau_M^\varepsilon \ge T \right) + \mathbb{P}(\tau_M^\varepsilon < T)$$

$$\le 2\mathbb{P}\left(|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0, \tau_M^\varepsilon \wedge T])} > \delta_1 \right) + \frac{\delta_2}{2} \le \delta_2$$

for all $\varepsilon \in (0, \varepsilon_0)$. Hence, $|X^\varepsilon - \bar{X}|_{\mathcal{C}^\alpha([0,T])} \to 0$ in probability as $\varepsilon \to 0$, as required. $\square$

REMARK 4.24.   The proof above shows that we can choose

$$\lambda_0 = \inf_{x \in \mathbb{R}^d} \Lambda(\kappa, R, p)$$

for any $p > \max(2, (H - \alpha)^{-1})$ in Theorem 1.2. Here, Λ is the constant from Proposition 3.13.

4.4. *Smoothness of the averaged coefficients.* Let us finally show that an *everywhere contractive* fast process falls in the regime of Theorem 1.2. While smoothness of $\bar{g}$ also holds under less restrictive conditions, the proof becomes much more involved. To keep this article concise, we chose to report on these results in future work.

COROLLARY 4.25. *Suppose that*:

- $g \in \mathcal{C}_b^3(\mathbb{R}^d \times \mathbb{R}^n, L(\mathbb{R}^m, \mathbb{R}^d))$,
- *there is a $\kappa > 0$ such that $b(x, \cdot) \in S(\kappa, 0, 0)$ for every $x \in \mathbb{R}^d$,*
- $b \in \mathcal{C}^3(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^d)$ *is globally Lipschitz continuous and there is an $N \in \mathbb{N}$ such that, for each $i, j, k \in \{x, y\}$,*

$$\left|D_{i,j}^2 b(x, y)\right| + \left|D_{i,j,k}^3 b(x, y)\right| \lesssim 1 + |y|^N \quad \forall x \in \mathbb{R}^d, \forall y \in \mathbb{R}^n.$$

Then the conclusion of Theorem 1.2 holds.

EXAMPLE 4.26. Let $V \in \mathcal{C}^4(\mathbb{R}^d \times \mathbb{R}^n)$. If $\inf_{x,y} D_{y,y}^2 V(x, y) \geq \kappa$, $|D_{x,y}^2 V|_\infty + |D_{y,y}^2 V|_\infty < \infty$, and, for each $i, j, k \in \{x, y\}$,

$$\left|D_{i,j,y}^3 V(x, y)\right| + \left|D_{i,j,k,y}^4 V(x, y)\right| \lesssim 1 + |y|^N \quad \forall x \in \mathbb{R}^d, \forall y \in \mathbb{R}^n,$$

then $b = -D_y V$ falls in the regime of Corollary 4.25. To give a concrete example, we can choose $V(x, y) = (2 + \sin(x))(y^2 + \sin(y))$, which furnishes the drift $b(x, y) = -(2 + \sin(x))(2y + \cos(y))$.

PROOF OF COROLLARY 4.25. In order to apply Theorem 1.2 it is enough to show that, for any $g \in \mathcal{C}_b^3(\mathbb{R}^n)$, the function

$$\bar{h}(x) \triangleq \int_{\mathbb{R}^n} g(y)\,\pi^x(dy)$$

is again of class $\mathcal{C}_b^2(\mathbb{R}^d)$. To this end, we define $h_t(x) \triangleq \mathbb{E}[g(Y_t^x)]$ where Y^x is the solution to the SDE

$$dY_t^x = b(x, Y_t^x)\,dt + \sigma\,d\hat{B}$$

started in the generalized initial condition $\delta_0 \otimes W$. Note that $h_t \to \bar{h}$ pointwise as $t \to \infty$ by Theorem 1.3. Since $h_t \in \mathcal{C}_b^2(\mathbb{R}^d)$ for each $t \geq 0$, it thus suffices to show that

(4.21) $$\sup_{t \geq 0}\left(|Dh_t|_\infty + \left|D^2 h_t\right|_\infty\right) < \infty$$

and both Dh_t and $D^2 h_t$ converge locally uniformly along a subsequence. By a straightforward "diagonal sequence" argument, we actually only need to prove uniform convergence on a fixed compact $K \subset \mathbb{R}^d$.

Under the assumptions of the corollary, it is easy to see that the mapping $x \mapsto Y_t^x$ is three-times differentiable for each $t \geq 0$ and it holds that

$$D_x Y_t^x = \int_0^t J_{s,t} D_x b(x, Y_s^x)\,ds,$$

$$D_{x,x}^2 Y_t^x(u \otimes v) = \int_0^t J_{s,t}\left(D_{x,x}^2 b(x, Y_s^x)(u \otimes v) + 2D_{x,y}^2 b(x, Y_s^x)(u \otimes D_x Y_s^x(v))\right.$$

$$\left. + D_{y,y}^2 b(x, Y_s^x)(D_x Y_s^x(u) \otimes D_x Y_s^x(v))\right) ds,$$

where $J_{s,t}$ solves the homogeneous problem

$$J_{s,t} = \text{id} + \int_s^t D_y b(x, Y_r^x) J_{s,r} \, dr.$$

Since $b(x, \cdot) \in S(\kappa, 0, 0)$, it is not hard to see that, for each $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^n$, $D_y b(x, y) \leq -\kappa$ in the sense of quadratic forms. In particular, the operator norm of J satisfies the bound

$$|J_{s,t}| \leq e^{-\kappa(t-s)}.$$

By an argument similar to Lemma 4.10, it follows that, for any $p \geq 1$,

$$\sup_{t \geq 0} \sup_{x \in \mathbb{R}^d} \| D_x Y_t^x \|_{L^p} < \infty \quad \text{and} \quad \sup_{t \geq 0} \sup_{x \in \mathbb{R}^d} \| D_{x,x}^2 Y_t^x \|_{L^p} < \infty.$$

Based on this, it is straight-forward to verify (4.21). Consequently, by the Arzela–Ascoli theorem, there is a subsequence of times along which Dh converges uniformly on K. By a similar—albeit more tedious—computation, the reader can easily check that also

$$\sup_{t \geq 0} \sup_{x \in \mathbb{R}^d} \| D_{x,x,x}^3 Y_t^x \|_{L^p} < \infty.$$

In particular, $D^3 h$ is uniformly bounded, whence we can pass to a further subsequence along which $D^2 h$ also converges uniformly on K. Therefore, $\bar{h} \in \mathcal{C}_b^2(\mathbb{R}^d)$ as required. $\square$

REFERENCES

[1] ABOURASHCHI, N. and VERETENNIKOV, A. YU. (2010). On stochastic averaging and mixing. *Theory Stoch. Process.* **16** 111–129. MR2779833

[2] ASHKENAZY, Y., BAKER, D. R., GILDOR, H. and HAVLIN, S. (2003). Nonlinearity and multifractality of climate change in the past 420,000 years. *Geophys. Res. Lett.* **30**.

[3] BAI, S. and TAQQU, M. S. (2016). Short-range dependent processes subordinated to the Gaussian may not be strong mixing. *Statist. Probab. Lett.* **110** 198–200. MR3474757 https://doi.org/10.1016/j.spl.2015.12.010

[4] BARBOZA, L., LI, B., TINGLEY, M. P. and VIENS, F. G. (2014). Reconstructing past temperatures from natural proxies and estimated climate forcings using short- and long-memory models. *Ann. Appl. Stat.* **8** 1966–2001. MR3292486 https://doi.org/10.1214/14-AOAS785

[5] BERGLUND, N. and GENTZ, B. (2006). *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*: *A Sample-Paths Approach. Probability and Its Applications* (*New York*). Springer, London. MR2197663

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

[7] BOGOLYUBOV, N. N. and MITROPOL'SKIĬ, YU. A. (1955). *Asimptotičeskie Metody v Teorii Nelineĭnyh Kolebaniĭ*. Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow. MR0073023

[8] BOURGUIN, S., GAILUS, S. and SPILIOPOULOS, K. (2021). Typical dynamics and fluctuation analysis of slow-fast systems driven by fractional Brownian motion. *Stoch. Dyn.* **21** Paper No. 2150030, 30. MR4323165 https://doi.org/10.1142/S0219493721500301

[9] BOURGUIN, S., GAILUS, S. and SPILIOPOULOS, K. (2021). Discrete-time inference for slow-fast systems driven by fractional Brownian motion. *Multiscale Model. Simul.* **19** 1333–1366. MR4301917 https://doi.org/10.1137/20M135813X

[10] CHAK, M., KANTAS, N. and PAVLIOTIS, G. A. (2020). On the generalised langevin equation for simulated annealing.

[11] CHERIDITO, P., KAWAGUCHI, H. and MAEJIMA, M. (2003). Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.* **8** no. 3, 14. MR1961165 https://doi.org/10.1214/EJP.v8-125

[12] CHERNY, A. (2008). Brownian moving averages have conditional full support. *Ann. Appl. Probab.* **18** 1825–1830. MR2432181 https://doi.org/10.1214/07-AAP502

[13] DAVIDSEN, J. and GRIFFIN, J. (2010). Volatility of unevenly sampled fractional Brownian motion: An application to ice core records. *Phys. Rev. E* **81** 016107.

[14] DEYA, A., PANLOUP, F. and TINDEL, S. (2019). Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise. *Ann. Probab.* **47** 464–518. MR3909974 https://doi.org/10.1214/18-AOP1265

[15] EBERLE, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041 https://doi.org/10.1007/s00440-015-0673-1

[16] EICHINGER, K., KUEHN, C. and NEAMŢU, A. (2020). Sample paths estimates for stochastic fast-slow systems driven by fractional Brownian motion. *J. Stat. Phys.* **179** 1222–1266. MR4123355 https://doi.org/10.1007/s10955-020-02485-4

[17] FONTBONA, J. and PANLOUP, F. (2017). Rate of convergence to equilibrium of fractional driven stochastic differential equations with some multiplicative noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 503–538. MR3634264 https://doi.org/10.1214/15-AIHP724

[18] FREIDLIN, M. I. and WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. Springer, Heidelberg. MR2953753 https://doi.org/10.1007/978-3-642-25847-3

[19] GEHRINGER, J. and LI, X.-M. (2020). Diffusive and rough homogenisation in fractional noise field. Available at arXiv:2006.11544.

[20] GEHRINGER, J. and LI, X.-M. (2022). Functional Limit Theorems for the Fractional Ornstein-Uhlenbeck Process. *J. Theoret. Probab.* **35** 426–456. MR4379471 https://doi.org/10.1007/s10959-020-01044-7

[21] HAIRER, M. (2005). Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.* **33** 703–758. MR2123208 https://doi.org/10.1214/009117904000000892

[22] HAIRER, M. (2009). Ergodic properties of a class of non-Markovian processes. In *Trends in Stochastic Analysis*. *London Mathematical Society Lecture Note Series* **353** 65–98. Cambridge Univ. Press, Cambridge. MR2562151

[23] HAIRER, M. and LI, X.-M. (2020). Averaging dynamics driven by fractional Brownian motion. *Ann. Probab.* **48** 1826–1860. MR4124526 https://doi.org/10.1214/19-AOP1408

[24] HAIRER, M. and OHASHI, A. (2007). Ergodic theory for SDEs with extrinsic memory. *Ann. Probab.* **35** 1950–1977. MR2349580 https://doi.org/10.1214/009117906000001141

[25] HAN, M., XU, Y. and PEI, B. (2021). Mixed stochastic differential equations: Averaging principle result. *Appl. Math. Lett.* **112** Paper No. 106705, 7. MR4145479 https://doi.org/10.1016/j.aml.2020.106705

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

[27] HEUNIS, A. J. and KOURITZIN, M. A. (1994). Strong convergence in the stochastic averaging principle. *J. Math. Anal. Appl.* **187** 134–155. MR1296610 https://doi.org/10.1006/jmaa.1994.1349

[28] KÄRNER, O. (2002). On nonstationarity and antipersistency in global temperature series. *J. Geophys. Res., Atmos.* **107** ACL–1.

[29] LÊ, K. (2020). A stochastic sewing lemma and applications. *Electron. J. Probab.* **25** Paper No. 38, 55. MR4089788 https://doi.org/10.1214/20-ejp442

[30] LEIMKUHLER, B. and SACHS, M. (2022). Efficient Numerical Algorithms for the Generalized Langevin Equation. *SIAM J. Sci. Comput.* **44** A364–A388. MR4379630 https://doi.org/10.1137/20M138497X

[31] LÉVY, P. (1953). Random functions: General theory with special reference to Laplacian random functions. *Univ. California Publ. Statist.* **1** 331–390. MR0055607

[32] LI, X.-M. (2018). Perturbation of conservation laws and averaging on manifolds. In *Computation and Combinatorics in Dynamics, Stochastics and Control. Abel Symp.* **13** 499–550. Springer, Cham. MR3967395

[33] LIU, S.-J. and KRSTIC, M. (2012). *Stochastic Averaging and Stochastic Extremum Seeking. Communications and Control Engineering Series*. Springer, London. MR2952849 https://doi.org/10.1007/978-1-4471-4087-0

[34] LIU, W., RÖCKNER, M., SUN, X. and XIE, Y. (2020). Averaging principle for slow-fast stochastic differential equations with time dependent locally Lipschitz coefficients. *J. Differential Equations* **268** 2910–2948. MR4047972 https://doi.org/10.1016/j.jde.2019.09.047

[35] LYONS, T. and QIAN, Z. (2002). *System Control and Rough Paths. Oxford Mathematical Monographs*. Oxford Univ. Press, Oxford. MR2036784 https://doi.org/10.1093/acprof:oso/9780198506485.001.0001

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

[37] 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

[38] NUALART, D. and RĂŞCANU, A. (2002). Differential equations driven by fractional Brownian motion. *Collect. Math.* **53** 55–81. MR1893308

[39] PANLOUP, F. and RICHARD, A. (2020). Sub-exponential convergence to equilibrium for Gaussian driven stochastic differential equations with semi-contractive drift. *Electron. J. Probab.* **25** Paper No. 62, 43. MR4112766 https://doi.org/10.1214/20-ejp464

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

[41] PEI, B., INAHAMA, Y. and XU, Y. (2020). Pathwise unique solutions and stochastic averaging for mixed stochastic partial differential equations driven by fractional brownian motion and brownian motion. *Probability*.

[42] PEI, B., INAHAMA, Y. and XU, Y. (2021). Averaging principle for fast-slow system driven by mixed fractional Brownian rough path. *J. Differential Equations* **301** 202–235. MR4305934 https://doi.org/10. 1016/j.jde.2021.08.006

[43] PICARD, J. (2011). Representation formulae for the fractional Brownian motion. In *Séminaire de Probabilités XLIII. Lecture Notes in Math.* **2006** 3–70. Springer, Berlin. MR2790367 https://doi.org/10.1007/ 978-3-642-15217-7_1

[44] PITERBARG, V. I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields. Translations of Mathematical Monographs* **148**. Amer. Math. Soc., Providence, RI. MR1361884 https://doi.org/10. 1090/mmono/148

[45] RÖCKNER, M. and XIE, L. (2021). Averaging principle and normal deviations for multiscale stochastic systems. *Comm. Math. Phys.* **383** 1889–1937. MR4244263 https://doi.org/10.1007/s00220-021-04069-z

[46] RÖCKNER, M., XIE, L. and YANG, L. (2020). Asymptotic behavior of multiscale stochastic partial differential equations.

[47] SAMKO, S. G., KILBAS, A. A. and MARICHEV, O. I. (1993). *Fractional Integrals and Derivatives*: *Theory and Applications*. Gordon & Breach, Yverdon. MR1347689

[48] 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

[49] SONECHKIN, D. M. (1998). Climate dynamics as a nonlinear Brownian motion. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **8** 799–803.

[50] YOUNG, L. C. (1936). An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67** 251–282. MR1555421 https://doi.org/10.1007/BF02401743

[51] YUAN, N., FU, Z. and LIU, S. (2014). Extracting climate memory using fractional integrated statistical model: A new perspective on climate prediction. *Sci. Rep.* **4** 6577. https://doi.org/10.1038/srep06577