TYPICAL DYNAMICS AND FLUCTUATION ANALYSIS OF SLOW-FAST SYSTEMS DRIVEN BY FRACTIONAL BROWNIAN MOTION

S. BOURGUIN, S. GAILUS, AND K. SPILOPOULOS

Boston University, Department of Mathematics and Statistics
111 Cummington Mall, Boston, MA 02215, USA

Abstract. This article studies typical dynamics and fluctuations for a slow-fast dynamical system perturbed by a small fractional Brownian noise. Based on an ergodic theorem with explicit rates of convergence, which may be of independent interest, we characterize the asymptotic dynamics of the slow component to two orders (i.e., the typical dynamics and the fluctuations). The limiting distribution of the fluctuations turns out to depend upon the manner in which the small-noise parameter is taken to zero relative to the scale-separation parameter. We study also an extension of the original model in which the relationship between the two small parameters leads to a qualitative difference in limiting behavior. The results of this paper provide an approximation to two orders of dynamical systems perturbed by small fractional Brownian noise and subject to multiscale effects.

1. Introduction

Dynamical systems exhibiting multiple characteristic scales in space or time arise naturally as models in a great variety of applied fields, including physics, chemistry, biology, neuroscience, meteorology, and mathematical finance, to name a few. It is moreover common to incorporate random perturbations into these models in order to account for imperfect information or to capture random phenomena. The particular case in which the perturbing noise is a standard Brownian motion has been studied extensively. With this choice, crucially, the Markov property and semimartingale structure of the standard Brownian motion are embedded in the system. While this simplifies the analysis and allows a host of well-developed theoretical tools to be brought to bear, the attendant features might unduly limit the flexibility of the model. For example, a physical dynamical system exhibiting long-range dependence or a particular sort of self-similarity may not be amenable to accurate description by a model driven by standard Brownian noise.

In this paper, we consider a model in which some of the random perturbation arises from a fractional Brownian motion (fBm), thereby making it possible to capture dynamical features that are out of the scope of the standard Brownian motion. More precisely, we consider \((X^\varepsilon, Y^\eta)_t = \{(X^\varepsilon_t, Y^\eta_t)\}_{0 \leq t \leq T}\) evolving in \(\mathcal{X} \times \mathcal{Y} := \mathbb{R}^m \times \mathbb{R}^{d-m}\) according to the stochastic differential equation

\[
\begin{align*}
    dX^\varepsilon_t &= c(X^\varepsilon_t, Y^\eta_t)dt + \sqrt{\varepsilon} \sigma(Y^\eta_t)dW^H_t \\
    dY^\eta_t &= \frac{1}{\eta} f(Y^\eta_t)dt + \frac{1}{\sqrt{\eta}} \tau(Y^\eta_t)dB_t \\
    X^\varepsilon_0 &= x_0 \in \mathcal{X}, \quad Y^\eta_0 = y_0 \in \mathcal{Y}.
\end{align*}
\]

Here, \(W^H\) is a fractional Brownian motion with Hurst index \(H \in (1/2, 1)\) and \(B\) is a standard Brownian motion independent of \(W^H\). The term \(dW^H\) is to be understood in the sense of pathwise integration, although this pathwise integral coincides in our framework with the analogous divergence integral, and we shall freely and frequently interpret it as such in order to apply tools of Malliavin calculus (see Remark 3 and Appendix A for a discussion of this point and for details on Malliavin calculus and integration with respect to fBm). \(\varepsilon := (\varepsilon, \eta) \in \mathbb{R}_+^2\) is a pair of small positive parameters. Note that as \(\varepsilon := (\varepsilon, \eta)\) is taken to
vanish, \( X^\varepsilon \) is the slow component and is perturbed by small noise, while \( Y^\eta \) is the fast component and feeds into the dynamics of \( X^\varepsilon \).

The main results of this work provide a rigorous description of the asymptotic behavior, to two orders, of \( X^\varepsilon \) as \( \varepsilon := (\varepsilon, \eta) \to 0 \). We first show that \( X^\varepsilon \) converges in an \( L^p \) sense, and at a particular rate, to a deterministic limiting process \( \bar{X} \), which we interpret as the typical behavior of \( X^\varepsilon \). We then derive a limit in distribution of the (appropriately-rescaled) fluctuations \( \theta^\varepsilon := \frac{1}{\sqrt{\varepsilon}}(X^\varepsilon - \bar{X}) \) about the limiting process. The limiting distribution of the fluctuations turns out to depend upon the manner in which the small asymptotic parameters are taken to vanish, even as the typical behavior does not exhibit any such dependence. In deriving the limit of the fluctuations, we assume for this reason a functional dependence \( \eta = \eta(\varepsilon) \) such that \( \lim_{\varepsilon \to 0} \frac{\sqrt{\varepsilon}}{\sqrt{\eta}} =: \lambda \in [0, \infty) \) exists.

The novelty of our setup lies in the nature of the small perturbing noise, which we take to be a fractional Brownian motion rather than a standard Brownian motion. Moreover, we allow the dynamics to evolve in the full Euclidean space, and apart from the diffusion coefficient in the fast component, we do not assume that coefficients are bounded or have bounded derivatives. Consequently, we rely in the proofs of our ergodic theorem, Theorem 3, and main results upon a-priori uniform bounds on both \( X^\varepsilon \) and its Malliavin derivative \( DX^\varepsilon \) with respect to the fractional Brownian noise \( W^H \). The necessary bounds are derived in Lemmata 1 and 5. In establishing the limit in distribution of the fluctuations, we also make use of recent results of [14], which carry over to our setting.

Note that in [1], we have taken \( \sigma \) to depend upon the fast variable only and not upon the slow. There are two reasons for this restriction, both relating to the fact that, by the independence of \( W^H \) and \( B \), \( D\sigma(Y^0) \equiv 0 \) whereas in general \( D\sigma(X^\varepsilon, Y^0) \) would be nontrivial (recall that \( D \) is the Malliavin derivative with respect to the fractional Brownian noise \( W^H \)). The first reason is technical in nature. As mentioned in the preceding paragraph, our proofs rely upon a-priori uniform bounds on \( X^\varepsilon \) and \( DX^\varepsilon \). To derive these bounds, we invoke a maximal inequality for the stochastic integral with respect to \( W^H \), which in turn requires us to control the Malliavin derivative of the integrand. Thus, if \( \sigma \) were allowed to depend upon the slow variable, we would encounter a closure problem in that to obtain a bound on the \( k^\text{th} \)-order Malliavin derivative \( D^kX^\varepsilon \) one would need first to obtain a bound on the \( (k+1)^\text{st} \)-order Malliavin derivative \( D^{k+1}X^\varepsilon \), and so on in a cascading fashion. In very special cases it is possible to circumvent the problem, but this would seem to be the exception rather than the rule. The second reason has to do with modelling considerations. If one would like to interpret the slow component as the solution of an ODE perturbed by a small fractional Brownian noise, it is reasonable to formulate the model in such a way as for this noise to be centered, i.e., for the stochastic integral with respect to \( W^H \) to have mean zero. In our setup, the pathwise integral coincides with the divergence integral, which is always centered (see Appendix A for details on Malliavin calculus and integration with respect to fBm). On the other hand, if one were to allow \( \sigma \) to depend upon the slow variable then the pathwise integral would not typically be centered. It is worth noting that one’s hands are tied here insofar as general results guaranteeing the existence of unique solutions of the system [1], are known only when the integral with respect to \( W^H \) is interpreted in the pathwise sense.

In the case of perturbation by standard Brownian motion, the literature on similar limiting theorems for stochastic dynamical systems is extensive. We mention here for completeness [2] [3] [4] [5] [6] [7] [8] [9] [10] [12] [13] [14] [16] [17] [22] [23] [29] [30], which contain results on related typical dynamics, central limit theorems, and large deviations. The corresponding literature in the case of perturbation by fractional Brownian motion is quite sparse. The most relevant result in our case is the recent work [14], which studies related typical behavior of systems similar to [1]. Our results on typical behavior differ from those of [14] in that we allow most coefficients to grow polynomially in the fast variable. Consequently, as discussed above, we must derive certain a-priori bounds in order to establish our ergodic theorem, Theorem 3, to which we appeal in turn in establishing our main results. In this way, we obtain an explicit rate of convergence to the typical behavior. To complement the results on typical behavior found in [14] and in this work, we then derive a limit in distribution of the fluctuations, which characterizes the limiting behavior of the slow component to next order.
In Section 5 we study the typical behavior and fluctuations limit in the context of an extended model generalized from (1). The extended model takes the form

\begin{equation}
\begin{aligned}
\frac{dX_t^\epsilon}{\sqrt{\epsilon}} &= b(X_t^\epsilon, Y_t^\epsilon)dt + c(X_t^\epsilon, Y_t^\epsilon)dt + \sqrt{\epsilon} \sigma(Y_t^\epsilon)dW_t^H \\
\frac{dY_t^\epsilon}{\sqrt{\epsilon}} &= \frac{1}{\gamma} f(Y_t^\epsilon)dt + \frac{1}{\sqrt{\epsilon}} g(Y_t^\epsilon)dt + \frac{1}{\sqrt{\epsilon}} \tau(Y_t^\epsilon)dB_t \\
X_0^\epsilon &= x_0 \in \mathcal{X}, \quad Y_0^\epsilon = y_0 \in \mathcal{Y}.
\end{aligned}
\end{equation}

Recall that in the context of the original model, the typical behavior does not depend upon the manner in which the small asymptotic parameters are taken to vanish, and that the fluctuations analysis can be carried through assuming only a functional dependence \( \eta = \eta(\epsilon) \) for which one has \( \lim_{\epsilon \to 0} \frac{\sqrt{\epsilon}}{\eta} =: \lambda \in [0, \infty) \). Moving to the extended model, however, the introduction of the terms corresponding to the coefficients \( b \) and \( g \) introduces a qualitative discrepancy between regimes that is reflected even in the typical behavior. Accordingly, when we are considering the extended model, we not only assume, from the beginning, the above functional dependence and existence of the limit \( \lambda \in [0, \infty) \), but also distinguish two possibilities:

(i) \( \lambda = 0 \), the ‘first regime’ or ‘homogenization regime’

(ii) \( \lambda \in (0, \infty) \), the ‘second regime’ or ‘averaging regime.’

Note that the homogenization regime is that in which the term \( \frac{\sqrt{\epsilon}}{\eta} b(X^\epsilon, Y^\epsilon) \) is asymptotically singular. In precise analogy to the analysis done in the case of perturbation by standard Brownian motion in [12, 30], we shall see that the limiting contribution of the asymptotically-singular term can be captured in terms of the solution of an appropriate Poisson equation. The proofs of our results for the original model \( (1) \) then carry over with minor modifications. Having already presented the full proofs for \( (1) \), we therefore describe only the adjustments necessary for \( (2) \). The extended model \( (2) \) is particularly relevant when, for example, a fast intermediate scale forms part of the slow component. The scaling in front of the term corresponding to the coefficient \( g \) is that which results in a nontrivial limiting contribution in the event that additional intermediate fast scales form part of the dominant fast component.

The rest of the paper is organized as follows. In Section 2 we introduce notation, present the conditions that hold by assumption throughout the paper, and state our main results. Section 3 contains proofs of results related to the typical behavior of \( X^\epsilon \) as \( \epsilon \to 0 \), including supporting lemmata and our ergodic theorem. Section 4 contains proofs of tightness and convergence in distribution of the (appropriately rescaled) fluctuations of \( X^\epsilon \) about the limit \( \bar{X} \) as \( \epsilon \to 0 \). Section 5 extends our results from the original model \( (1) \) to the extended model \( (2) \). Finally, for the convenience of the reader, Appendix A collects those definitions and tools related to fractional Brownian motion, Malliavin calculus, and stochastic integration with respect to fractional Brownian motion, that are used in this paper.

2. Notation, Conditions, and Main Results

In this section we introduce notation, present the conditions that we will assume throughout the paper, and state our main results.

We will denote by \( A : B \) the Frobenius inner product \( \Sigma_{i,j}[a_{i,j} \cdot b_{i,j}] \) of matrices \( A = (a_{i,j}) \) and \( B = (b_{i,j}) \). We will use single bars \( \cdot \) to denote the Frobenius (or Euclidean) norm of a matrix, and double bars \( \| \cdot \| \) to denote the operator norm.

Condition 1 imposes conditions of growth and regularity on the drift and diffusion coefficients of the model.

Condition 1.

Conditions on \( c \):
- \( \exists (K, q) \in \mathbb{R}^2_+ \), \( r \in [0, 1) \); \( |c(x, y)| \leq K(1 + |x|^r)(1 + |y|^q) \)
- \( \exists (K, q) \in \mathbb{R}^2_+ \); \( |\nabla_x c(x, y)| + |\nabla_x \nabla_x c(x, y)| \leq K(1 + |y|^q) \)
- \( c, \nabla_x c, \nabla_x \nabla_x c \) are continuous in \( (x, y) \)
- \( c, \nabla_x c, \) and \( \nabla_y \nabla_y c \) are Hölder continuous in \( y \) uniformly in \( x \)

Conditions on \( \sigma \):
- \( \exists (K, q) \in \mathbb{R}^2_+ \); \( |\sigma(y)| \leq K(1 + |y|^q) \)
- \( \sigma \sigma^T \) is uniformly nondegenerate
Conditions on \( f \) and \( \tau \):
- \( f \) and \( \tau \) are twice differentiable, and, along with their first and second derivatives, are Hölder continuous
- \( \tau \) is uniformly bounded and uniformly nondegenerate.

Condition 2 is a basic condition of recurrence type on the fast component, yielding ergodic behavior.

**Condition 2.**
\[
\lim_{|y| \to \infty} y \cdot f(y) = -\infty.
\]

To derive most of our results we shall in fact assume a stronger recurrence condition.

**Condition 3.**
For real constants \( \alpha > 0 \), \( \beta \geq 2 \), and \( \gamma > 0 \), we shall write:
- Condition 3-(\( \alpha, \beta \))): one has
\[ \begin{align*}
& y \cdot f(y) + \alpha|y|^{\beta} + \frac{1}{2} (\beta - 2 + d - m) \sup_{\tilde{y} \in \mathcal{Y}} |\tau(\tilde{y})|^2 \\
& \quad \leq 0
\end{align*} \]
for \( |y| \) sufficiently large
- Condition 3-(\( \alpha, \beta, \gamma \))): Condition 3-(\( \alpha, \beta \)) holds and, moreover, one has \( ||\nabla_x c(x, y)|| \leq \gamma |y|^\beta \) for \( |y| \) sufficiently large.

**Remark 1.** Clearly, Condition 2 is implied by Condition 3-(\( \alpha, \beta \)), which in turn is implied by the stronger condition
\[
\lim_{|y| \to \infty} y \cdot f(y) + \alpha|y|^{\beta} = -\infty.
\]

One has the infinitesimal generator
\[
\mathcal{L} := f \cdot \nabla_y + \frac{1}{2} (\tau \tau^T) : \nabla_y^2
\]
for the rescaled fast dynamics. Conditions 1 and 2 are enough to guarantee that one has on \( \mathcal{Y} \) a unique invariant measure \( \mu \) corresponding to the operator \( \mathcal{L} \), as discussed for example in [27].

**Remark 2.** Besides Conditions 1 and 2 or 3, we also assume throughout that the system (1) has a unique strong solution. Sufficient conditions for this have been derived, for example, in [10], if one assumes that the coefficients and some of their partial derivatives satisfy a global Lipschitz condition and have at most linear growth (see [10] and [18] for more details).

We now state our main results, the first of which concerns the typical behavior of \( X^\varepsilon \) as \( \varepsilon \to 0 \). We prove in Theorem 4 that \( X^\varepsilon \) converges in an \( L^p \) sense, and at a particular rate, to a deterministic limiting process \( \bar{X} \). This implies in particular that one has convergence in probability. The proof is deferred to Section 3.

**Theorem 1.** Assume Conditions 1 and 3-(\( \alpha, \beta, \gamma \)), where \( \alpha \geq 0 \), \( \beta \geq 2 \), \( \gamma \geq 0 \), and \( T \beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2 < 2\alpha \). For any \( 0 < p < \frac{2\alpha}{T \beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2} \), there is a constant \( \tilde{K} \) such that for \( \varepsilon := (\varepsilon, \eta) \) sufficiently small,
\[
E \sup_{0 \leq t \leq T} |X^\varepsilon_t - \bar{X}_t|^p \leq \tilde{K} \left( \varepsilon^p + \sqrt{\eta^p} \right),
\]
where \( \bar{X} \) is the (deterministic) solution of the integral equation
\[
\bar{X}_t = x_0 + \int_0^t \bar{c}(\bar{X}_s)ds,
\]
where \( \bar{c} \) is the averaged function
\[
\bar{c}(x) := \int_{\mathcal{Y}} c(x, y)\mu(dy).
\]
Our second main result concerns the asymptotic behavior of the (appropriately rescaled) fluctuations of \(X^\varepsilon\) around \(X\) as \(\varepsilon \to 0\). We prove in Theorem 2 that the fluctuations converge in distribution to a particular limit, which we characterize explicitly. The proof is deferred to Section 3. In order to state the theorem, we note that by Theorem 3 in [23], the equations
\[
\begin{align*}
\mathcal{L}\Phi(x, y) &= -(c(x, y) - \bar{c}(x)) \\
\int_y \Phi(x, y) \mu(dy) &= 0
\end{align*}
\]
admite a unique solution \(\Phi\) in the class of functions that grow at most polynomially in \(|y|\) as \(y \to \infty\).

**Theorem 2.** Suppose that \(\eta = \eta(\varepsilon)\) and that \(\lim_{\varepsilon \to 0} \sqrt{\varepsilon^2} =: \lambda \in [0, \infty)\). Assume Conditions 1 and 3 (\(\alpha, \beta, \gamma\), where \(\alpha \geq 0\), \(\beta \geq 2\), \(\gamma \geq 0\), and \(T\beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2 < 2\alpha\)). With \(\Phi\) as in (4), set \(\Sigma_\Phi := \left((\nabla_y \Phi^\tau)(\nabla_y \Phi^\tau)^T\right)^{1/2}\). The family of processes \(\{\hat{\theta}^\varepsilon\}_\varepsilon\) converges in distribution on the space \(C([0, T]; \mathcal{X})\) (endowed, as usual, with the topology of uniform convergence) as \(\varepsilon \to 0\) to the law of the solution \(\theta\) of the mixed SDE
\[
\begin{align*}
d\theta_t &= \int_0^t (\nabla_x \bar{c})(\bar{X}_s) \cdot \theta_s ds + \lambda \int_0^t \Sigma_\Phi(\bar{X}_s) d\tilde{B}_s + \bar{\sigma}\tilde{W}_t^H \\
\theta_0 &= 0,
\end{align*}
\]
where \(\bar{\sigma} := \int_0^T \sigma(y) \mu(dy)\), \(\tilde{W}^H\) is a fractional Brownian motion with Hurst index \(H\), and \(\tilde{B}\) is a standard Brownian motion independent of \(\tilde{W}^H\).

3. **FIRST-ORDER LIMIT OR TYPICAL BEHAVIOR**

In this section, we focus on proving Theorem 1 which establishes the first-order limit, or typical behavior, of the slow component \(X^\varepsilon\) as \(\varepsilon \to 0\). In order to make the exposition easier to follow, we present several supporting lemmata leading up to an ergodic theorem, Theorem 3 which is the essential ingredient in the proof of Theorem 1.

**Remark 3.** Recall that while we interpret the stochastic integral \(\int_0^t \sigma(Y^\gamma_s) dW^H_s\) appearing in (1) in the pathwise sense so that one may appeal to existence and uniqueness results in the literature, this pathwise integral coincides in our framework with the divergence integral \(\int_0^t \sigma(Y^\gamma_s) \delta W^H_s\) (see Appendix A for details on Malliavin calculus and integration with respect to fBm). Indeed, in view of (20), the two integrals coincide as soon as the integrand is in the kernel of the Malliavin derivative associated with the fractional Brownian motion \(W^H\), which is of course true in our case as \(\sigma(Y^\gamma)\) and \(W^H\) are independent. In what follows, we work mainly with the divergence integral in place of the pathwise integral so that we may make use of powerful results from the Malliavin stochastic calculus of variations.

We begin by stating a maximal inequality for the divergence integral \(\int_0^t \sigma(Y^\gamma_s) \delta W^H_s\).

**Lemma 1.** Assume Conditions 1 and 3. For any \(H^{-1} < p < \infty\), there is a constant \(\tilde{K}\) such that for \(\eta\) sufficiently small,
\[
E \sup_{0 \leq t \leq T} \left| \int_0^t \sigma(Y^\gamma_s) \delta W^H_s \right|^p \leq \tilde{K}.
\]

**Proof.** Recalling that \(\sigma\) is polynomially bounded in its argument and appealing to Lemma 1 in [23], the claim follows from the maximal inequality stated after (2.14) in [19].

We next obtain a preliminary uniform bound on the slow component.

**Lemma 2.** Assume Conditions 1 and 3. For any \(0 < p < \infty\), there is a constant \(\tilde{K}\) such that for \(\varepsilon := (\varepsilon, \eta)\) sufficiently small,
\[
E \sup_{0 \leq t \leq T} |X^\varepsilon_t|^p \leq \tilde{K}.
\]

**Proof.** It is enough to prove the lemma for \(p \geq 2\). Recall that
\[
X^\varepsilon_t = x_0 + \int_0^t c(X^\varepsilon_s, Y^\gamma_s) ds + \sqrt{\varepsilon} \int_0^t \sigma(Y^\gamma_s) \delta W^H_s.
\]
Lemma 3. Assume Conditions 1 and 2. Let \( \mathcal{L} \) that for \( K, r, q > 0 \), in Lemma 1 in [23], we now show that polynomially-bounded measurable functions of \( \Box \) The proof is complete upon applying the Grönwall inequality.

Proof. Note that, using our assumption on the function \( \nu_\beta \) be a standard Brownian motion and let \( \tilde{Y} \) denote the solution of the stochastic differential equation

\[
\left\{ \begin{array}{l}
d\tilde{Y}_t = f(\bar{Y}_t)dt + \tau(\bar{Y}_t)d\tilde{B}_t \\
\bar{Y}_0 = \eta_0.
\end{array} \right.
\]

Taking together the bound on the slow component in Lemma 2 and the bound on the fast component in Lemma 1 in [23], we now show that polynomially-bounded measurable functions of \( X^\varepsilon \) and \( Y^\eta \) represent classes in \( L^p(\Omega \times [0, T]) \).

Lemma 3. Assume Conditions 1 and 2. Let \( h \) be a measurable function on \( X \times Y \) and suppose that constants \( K, r, q > 0 \) exist for which \( |h(x, y)| \leq K(1 + |x|^r)(1 + |y|^q) \). For any \( 0 < p < \infty \), there is a constant \( \bar{K} \) such that for \( \varepsilon := (\epsilon, \eta) \) sufficiently small,

\[
E \int_0^T |h(X_t^\varepsilon, Y_t^\eta)|^p dt \leq \bar{K}.
\]

Proof. Note that, using our assumption on the function \( h \), we can write

\[
E \int_0^T |h(X_t^\varepsilon, Y_t^\eta)|^p dt \leq E \int_0^T K^p(1 + |X_t^\varepsilon|^r)(1 + |Y_t^\eta|^q)^p dt
\]

\[
\leq \frac{K^p}{2} \left( E \int_0^T (1 + |X_t^\varepsilon|^r)^{2p} dt + E \int_0^T (1 + |Y_t^\eta|^q)^{2p} dt \right).
\]

The terms inside the parentheses are bounded respectively by Lemma 2 above and Lemma 1 in [23], concluding the proof.

As we have mentioned, the proof of our ergodic theorem relies on having first obtained uniform bounds not only on the slow component \( X^\varepsilon \) but also on its Malliavin derivative \( DX^\varepsilon \) with respect to the fractional Brownian motion \( W^H \). The next lemma provides appropriate technical uniform bounds on exponential moments of the fast component, which we will then use to establish the necessary bound on the Malliavin derivative.

Lemma 4. Assume Conditions 1 and 3, \( \alpha, \beta \), where \( \alpha > 0 \) and \( \beta \geq 2 \). For any \( \nu \geq 0 \) such that \( \nu \beta \sup_{y \in Y} ||\tau(y)||^2 < 2\alpha \), there is a constant \( \bar{K} \) such that for all \( \eta > 0 \),

\[
\sup_{0 \leq t \leq T} E e^{\nu |Y_t^\eta|^\beta} \leq \bar{K}.
\]

Proof. Let \( \tilde{B} \) be a standard Brownian motion and let \( \bar{Y} \) denote the solution of the stochastic differential equation

\[
\left\{ \begin{array}{l}
d\bar{Y}_t = f(\bar{Y}_t)dt + \tau(\bar{Y}_t)d\tilde{B}_t \\
\bar{Y}_0 = \eta_0.
\end{array} \right.
\]
Since \( \{Y^\alpha_t\}_{0 \leq t \leq T} \) has the same law as \( \{\tilde{Y}_{t/n}\}_{0 \leq t \leq T} \), the claim of the lemma is equivalent to the statement that
\[
(5) \quad \sup_{0 \leq t < \infty} E e^{\nu|Y_t|^\alpha} < \infty.
\]

Fix \( N \in \mathbb{N} \) and put \( t_N := \inf\{t \in [0, \infty) ; Y_t \geq N\} \). Let \( \tilde{Y}_N \) denote the process obtained by halting \( \tilde{Y} \) at time \( t_N \), i.e., let \( \{\tilde{Y}_{N,t}\}_{t \geq 0} := \{\tilde{Y}_{t,n} : t \geq 0\} \). We will show that there is a constant \( \tilde{K} \) such that for all \( N \in \mathbb{N} \),
\[
(6) \quad \sup_{0 \leq t < \infty} E e^{\nu|\tilde{Y}_{N,t}|^\alpha} < \tilde{K}
\]
and that
\[
(7) \quad P(\lim_{N \to \infty} t_N = \infty) = 1.
\]

Taking together (6) and (7), (5) follows easily.

To establish (6), choose \( \ell > 1 \) such that \( \ell \nu / \beta \sup_{y \in \mathcal{Y}} \|\tau(y)\|^2 \leq 2\alpha \) and apply Itô’s lemma to obtain
\[
(8) \quad E e^{\nu|\tilde{Y}_{N,t}|^\alpha} = e^{\nu|y_0|^\alpha} E \left[ e^{\nu \int_0^t \nabla (|\tilde{Y}_{N,s}|^\alpha) ds} \left( \tilde{Y}_{N,t} \right) \right] = e^{\nu|y_0|^\alpha} E \left[ e^{I_1 \cdot e^{I_2}} \right],
\]
where, with \( 1_{d-m} \) denoting the \((d - m) \times (d - m)\) identity matrix,
\[
I_1 := \nu \int_0^t \left( \beta |\tilde{Y}_{N,s}|^{\beta - 2} \tilde{Y}_{N,s} \cdot f(\tilde{Y}_{N,s}) + \frac{\nu \ell}{2(\ell - 1)} \beta^2 |\tilde{Y}_{N,s}|^{2\beta - 4} |\tilde{Y}_{N,s} \cdot \tau(\tilde{Y}_{N,s})|^2 
\]
\[
+ \frac{1}{2} \left( \beta(\beta - 2)|\tilde{Y}_{N,s}|^{\beta - 4} \tilde{Y}_{N,s} \otimes \tilde{Y}_{N,s} + \beta|\tilde{Y}_{N,s}|^{\beta - 2} 1_{d-m} \right) : (\tau^T)(\tilde{Y}_{N,s}) \right) ds,
\]
\[
I_2 := \nu \int_0^t \beta |\tilde{Y}_{N,s}|^{\beta - 2} \tilde{Y}_{N,s} \cdot \tau(\tilde{Y}_{N,s}) d\tilde{B}_s - \frac{\nu \ell}{2(\ell - 1)} \int_0^t \beta^2 |\tilde{Y}_{N,s}|^{2\beta - 4} \tilde{Y}_{N,s} \cdot \tau(\tilde{Y}_{N,s}) |^2 ds.
\]

Applying Young’s inequality with conjugate exponents \( \ell \) and \( \frac{1}{\ell} \),
\[
(9) \quad e^{\nu|y_0|^\alpha} E \left[ e^{I_1 \cdot e^{I_2}} \right] \leq e^{\nu|y_0|^\alpha} \left( \frac{1}{\ell} E e^{I_1 \cdot I_2} + E e^{\frac{\ell - 1}{\ell} I_1} \right).
\]

Note that, on the one hand, by Condition (4) \((\alpha, \beta)\), there is a constant \( C \) independent of \( N \) such that \( E e^{I_1} \leq C \), and on the other hand, \( e^{\frac{\ell - 1}{\ell} I_2} \) is an exponential martingale with unit mean. Consequently,
\[
(10) \quad e^{\nu|y_0|^\alpha} \left( \frac{1}{\ell} E e^{I_1 \cdot I_2} + E e^{\frac{\ell - 1}{\ell} I_1} \right) \leq e^{\nu|y_0|^\alpha} \left( \frac{1}{\ell} C + \frac{\ell - 1}{\ell} \right).
\]

Putting together (8), (9), and (10), we obtain (6) with \( \tilde{K} := \frac{C + \ell - 1}{\ell} e^{\nu|y_0|^\alpha} \).

It remains to verify (7). By [15] Chapter 6, Theorem 4.1, one may realize on some probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\) real-valued stochastic processes \( \{\Xi_t\}_{t \geq 0} \) and \( \{\Xi^+_t\}_{t \geq 0} \) such that \( \Xi_0 = \Xi^+_0 = |y_0|^2 \),
\[
\tilde{P} \left( \forall t \geq 0, \sup_{0 \leq s \leq t} \Xi_s \leq \sup_{0 \leq s \leq t} \Xi^+_s \right) = 1,
\]
\( \Xi \) is equal in law to \( |\tilde{Y}|^2 \), and \( \Xi^+ \) is an Itô diffusion with generator \( a(\xi) \left( b(\xi) dT + \frac{1}{2} \left( \tau(\xi) \right)_d \right) \), where, for \( \xi > 0 \),
\[
a(\xi) := \sup_{y \in \mathcal{Y}; |y|^2 = \xi} \left( 4y^T \tau(y) \tau^T(y) y \right),
\]
\[
b(\xi) := \sup_{y \in \mathcal{Y}; |y|^2 = \xi} \left( 2y \cdot f(y) + |\tau(y)|^2 \right).
\]

It is enough, then, to show that \( \tilde{P} \lim_{N \to \infty} \inf\{t \in [0, \infty) ; \Xi^+_t \geq N\} = \infty \). If this were not the case, one would have some fixed \( \tilde{T} > 0 \) for which the event
\[
\tilde{A} := \left\{ \sup_{N \in \mathbb{N}} \left( \inf\{t \in [0, \tilde{T}] ; \Xi^+_t \geq N\} < \tilde{T} \right) \right\}
\]
is such that \( \tilde{P}(\tilde{A}) > 0 \). One would then have that for all \( N \in \mathbb{N} \), \( E \sup_{0 \leq t \leq \tilde{T}} |\Xi^+_t| \geq N \tilde{P}(\tilde{A}) \), whence immediately \( E \sup_{0 \leq t \leq \tilde{T}} |\Xi^+_t| = \infty \). It therefore suffices to show that for each fixed \( \tilde{T} > 0 \), \( E \sup_{0 \leq t \leq \tilde{T}} |\Xi^+_t| < \infty \).

By Condition 9 (\( \alpha, \beta \)), one sees in particular that \( a(\xi)b(\xi) \) is negative for \( \xi \) sufficiently large, say \( \xi > \xi_0 \).

In light of this observation, by [15, Chapter 6, Theorem 1.1] coupled with a stopping-and-starting argument it suffices to consider an Itô diffusion \( \tilde{Z} \) with initial value \( \tilde{Z}_0 = |\eta_0|^2 \wedge \xi_0 \) and generator \( \frac{1}{2}a(\xi)\frac{\partial^2}{\partial \xi^2} \), and to show that for each fixed \( \tilde{T} > 0 \), \( E \sup_{0 \leq t \leq \tilde{T}} |\tilde{Z}_t| < \infty \).

To this end, suppose that

\[
\tilde{Z}_t = \tilde{Z}_0 + \int_0^t \sqrt{a(\tilde{Z}_s)}dV_s,
\]

where \( V_s \) is a standard Brownian motion in one dimension. By the Burkholder-Davis-Gundy inequality and the fact that \( a(\xi) \leq 4(\sup_{y \in Y} ||\eta(y)||^2)(1 + \xi^2) \), we have, for \( K := 16 \sup_{y \in Y} ||\eta(y)||^2 \),

\[
E \sup_{0 \leq s \leq t} \tilde{Z}_s^2 \leq 2 \tilde{Z}_0^2 + 2E \sup_{0 \leq s \leq t} \left| \int_0^s \sqrt{a(\tilde{Z}_u)}dB_u \right|^2 \\
\leq 2 \tilde{Z}_0^2 + 2Kt + 2K \int_0^t E \sup_{0 \leq u \leq s} \tilde{Z}_u^2 du,
\]

whence Grönwall’s inequality gives

\[
E \sup_{0 \leq t \leq \tilde{T}} \tilde{Z}_t^2 \leq \left( 2 \tilde{Z}_0^2 + 2K\tilde{T} \right) e^{2KT}.
\]

This completes the proof of the lemma. \( \square \)

With Lemma 4 in hand, we are now in a position to establish the necessary bound on the Malliavin derivative \( DX^\varepsilon \) of the slow component \( X^\varepsilon \).

**Lemma 5.** Assume Conditions 1 and 3–(\( \alpha, \beta, \gamma \)), where \( \alpha > 0 \), \( \beta \geq 2 \), \( \gamma > 0 \), and \( T\beta \gamma \sup_{y \in Y} ||\eta(y)||^2 < 2\alpha \).

For any \( 0 < p < \frac{2\alpha}{T\beta \gamma \sup_{y \in Y} ||\eta(y)||^2} \), there is a constant \( \tilde{K} \) such that for \( \varepsilon := (\epsilon, \eta) \) sufficiently small,

\[
\sup_{0 \leq s, t \leq T} E |D_s X^\varepsilon| p \leq \tilde{K}.
\]

**Proof.** It is enough to prove the lemma for \( p > 1 \). We begin by noting that

\[
D_s X^\varepsilon_t = \int_0^t \nabla_x c(X^\varepsilon_u, Y^\varepsilon_u)D_s X^\varepsilon_u du + \sqrt{\epsilon} \sigma(Y^\varepsilon_u) \chi_{[0,t]}(s).
\]

Hence, for \( t < s \), \( D_s X^\varepsilon_t = 0 \), while for \( t \geq s \),

\[
|D_s X^\varepsilon_t|^p = \left| \sqrt{\epsilon}\sigma(Y^\varepsilon_u) \right|^p + \int_s^t |D_s X^\varepsilon_u|^p - 2D_s X^\varepsilon_u : (\nabla_x c(X^\varepsilon_u, Y^\varepsilon_u)D_s X^\varepsilon_u) du \\
\leq \left| \sqrt{\epsilon}\sigma(Y^\varepsilon_u) \right|^p + \int_s^t |D_s X^\varepsilon_u|^p ||\nabla_x c(X^\varepsilon_u, Y^\varepsilon_u)|| du.
\]

Applying Grönwall’s inequality and Young’s inequality with conjugate exponents \( \frac{\ell}{\ell - 1} \) and \( \ell \) yet to be determined, we obtain, for \( t \geq s \),

\[
|D_s X^\varepsilon_t|^p \leq |\sqrt{\epsilon}\sigma(Y^\varepsilon_u)|^p e^{p \int_s^t ||\nabla_x c(X^\varepsilon_u, Y^\varepsilon_u)|| du} \\
\leq I + II_t,
\]

where

\[
I := \frac{\ell - 1}{\ell} |\sqrt{\epsilon}\sigma(Y^\varepsilon_u)|^p, \\
II_t := \frac{1}{\ell} e^{p \int_s^t ||\nabla_x c(X^\varepsilon_u, Y^\varepsilon_u)|| du}.
\]

For any given choice of \( \ell > 1 \), the expected value \( E(I) \) of the first summand is easily handled by [22, Corollary 1], so we proceed to consider the expected value of the second summand. Recalling that, by
assumption, \( \|\nabla c(x,y)\| \leq \gamma |y|^2 \) for \( |y| \) sufficiently large, we have, for some constant \( C > 0 \), applying Jensen’s inequality,
\[
E \ell_t \leq E \frac{1}{\ell} e^{\ell p \int_0^T \|\nabla c(X_t^\alpha,Y_t^\gamma)\| du} \\
\leq \frac{1}{\ell} C \ell + E \frac{1}{\ell} e^{\ell p \int_0^T |Y_t^\gamma|^2 du} \\
\leq \frac{1}{T} C \ell + E \frac{1}{T} \int_0^T e^{\ell p \int_0^T |Y_t^\gamma|^2 du} \\
\leq \frac{1}{T} C \ell + \frac{1}{\ell} \sup_{0 \leq t \leq T} E e^{\ell p \int_0^T |Y_t^\gamma|^2},
\]
whence the proof is complete upon choosing \( \ell > 1 \) small enough that \( \ell T p \gamma sup_{y \in Y} ||\tau(y)||^2 \leq 2\alpha \) and appealing to Lemma 4.

\[ \square \]

The technical ingredients for the ergodic theorem are now in place. Before moving on to the theorem, we present a version of Itô’s lemma adapted to our framework.

**Lemma 6.** Let \( F \) be a function of class \( C^2 \) on \( X \times Y \). Then it holds that
\[
F(X_t^\alpha,Y_t^\gamma) - F(x_0,y_0) = \int_0^t \langle \nabla_x F(X_s^\alpha,Y_s^\gamma) \rangle \delta X_s^\alpha + \int_0^t \langle \nabla_y F(X_s^\alpha,Y_s^\gamma) \rangle dY_s^\gamma \\
+ \alpha \int_0^t \langle \nabla_x^2 F(X_s^\alpha,Y_s^\gamma) \rangle \cdot \sigma(Y_s^\gamma) \left( \int_0^s \sigma(Y_u^\gamma)(s-u)^{2H-2} du \right) ds \\
+ \frac{1}{2\eta} \int_0^t \langle \nabla_y^2 F(Y_s^\gamma) \rangle : \left( \sigma \tau^T \right)(Y_s^\gamma) ds,
\]
where \( \alpha H := H(2H - 1) \).

**Proof.** This is a straightforward extension of the well-known Itô formula for the divergence integral (see e.g. Remark 4).

**Remark 4.** With \( \mathcal{L} \) as in \( \[ \text{eq:ito} \] \), equation (11) may also be written as
\[
F(X_t^\alpha,Y_t^\gamma) - F(x_0,y_0) = \frac{1}{\eta} \int_0^t \langle \mathcal{L} F \rangle (X_s^\alpha,Y_s^\gamma) ds + \int_0^t \langle \nabla_x F \sigma \rangle (X_s^\alpha,Y_s^\gamma) ds \\
+ \alpha \int_0^t \langle \nabla_x^2 F \rangle (X_s^\alpha,Y_s^\gamma) \cdot \sigma(Y_s^\gamma) \left( \int_0^s \sigma(Y_u^\gamma)(s-u)^{2H-2} du \right) ds \\
+ \sqrt{\epsilon} \int_0^t \langle \nabla_x F \sigma \rangle (X_s^\alpha,Y_s^\gamma) dW_s^H + \frac{1}{\sqrt{\eta}} \int_0^t \langle \nabla_y F \rangle (X_s^\alpha,Y_s^\gamma) dB_s.
\]

We are now ready to state and prove our ergodic theorem.

**Theorem 3.** Assume Conditions \( \[ \text{eq:cond} \] \) and \( \[ \text{eq:cond1} \] \) (\( \alpha, \beta, \gamma \), where \( \alpha > 0, \beta \geq 2, \gamma > 0 \), and \( T \beta \gamma sup_{y \in Y} ||\tau(y)||^2 \leq 2\alpha \)). Let \( h \) be a differentiable function on \( X \times Y \) and suppose that constants \( K, r, q > 0 \) exist for which \( |h(x,y)| \leq K(1 + |x|^\gamma)(1 + |y|^\gamma) \). Suppose further that each derivative of \( h \) up to second order is Hölder continuous in \( y \) uniformly in \( x \), with absolute value growing at most polynomially in \( |y| \) as \( y \to \infty \). For any \( 0 < p < \frac{2\alpha}{T \beta \gamma sup_{y \in Y} ||\tau(y)||^2} \), there is a constant \( \tilde{K} \) such that for \( \epsilon := (\epsilon,\eta) \) sufficiently small,
\[
E \sup_{0 \leq t \leq T} \left| \int_0^t (h(X_s^\alpha,Y_s^\gamma) - \tilde{h}(X_s)) \right|^p \leq \tilde{K} \sqrt{\eta}^p,
\]
where \( \tilde{h}(x) \) is the averaged function \( \int_Y h(x,y) d\mu(y) \).

**Proof.** It is enough to prove the theorem for \( p \geq 2 \). By \( \[ \text{eq:ito} \] \) Theorem 3, the equations
\[
\left\{ \begin{array}{l}
\mathcal{L} \Phi(x,y) = h(x,y) - \tilde{h}(x) \\
\int_Y \Phi(x,y) d\mu(y) = 0
\end{array} \right.
\]
admit a unique solution $Φ$ in the class of functions that grow at most polynomially in $|y|$ as $y \to \infty$. Applying Lemma 6 with $F = Φ$ and rearranging terms gives
\[
\int_0^t \left( h(X_\epsilon^t, Y_\epsilon^\eta) - \bar{h}(X_\epsilon^t) \right) ds = \sqrt{\eta} \left( \sqrt{\eta} \left( Φ(X_\epsilon^t, Y_\epsilon^\eta) - Φ(x_0, y_0) \right) - \sqrt{\eta} \int_0^t (\nabla_x Φ)(X_\epsilon^t, Y_\epsilon^\eta) ds \right. \\
- \epsilon \sqrt{\eta} \alpha_H \int_0^t (\nabla^2 x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \cdot \left( \int_0^s \sigma(Y_\eta^u)(s - u)^{2H-2} du \right) ds \\
- \sqrt{\eta} \int_0^t (\nabla_y Φσ)(X_\epsilon^t, Y_\epsilon^\eta) dB_s - \int_0^t (\nabla_y Φτ)(X_\epsilon^t, Y_\epsilon^\eta) dB_s \right),
\]
where $\alpha_H := H(2H - 1)$; hence, for $ε$ sufficiently small,
\[
E \sup_{0 \leq t \leq T} \left| \int_0^t \left( h(X_\epsilon^t, Y_\epsilon^\eta) - \bar{h}(X_\epsilon^t) \right) ds \right|^p \leq 5^p \sqrt{\eta}^p \left( E \sup_{0 \leq t \leq T} \sqrt{\eta}^p |Φ(X_\epsilon^t, Y_\epsilon^\eta) - Φ(x_0, y_0)|^p + \sqrt{\eta}^p E \int_0^T |(\nabla_x Φ)(X_\epsilon^t, Y_\epsilon^\eta)|^p ds \right. \\
+ (\epsilon \sqrt{\eta})^p \alpha_H E \int_0^T \left| (\nabla^2 x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \cdot \left( \int_0^s \sigma(Y_\eta^u)(s - u)^{2H-2} du \right) \right|^p ds \\
+ \sqrt{\eta}^p E \sup_{0 \leq t \leq T} \left| \int_0^t \left( (\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \delta W_s^H \right) \right|^p + E \sup_{0 \leq t \leq T} \left| \int_0^T (\nabla_y Φτ)(X_\epsilon^t, Y_\epsilon^\eta) dB_s \right|^p).
\]
It remains to show that the expected value terms inside the parentheses are bounded uniformly in $ε$ sufficiently small. Recalling the stochastic representation of $Φ$ in [22,23] and the argument in the proof of [23] Theorem 3], the function $Φ$ itself and all of the derivatives of $Φ$ that appear are continuous in $x$ and $y$ and bounded by expressions of the form $K(1 + |x|^p)(1 + |y|^q)$. Thus, the term $E \sup_{0 \leq t \leq T} \sqrt{\eta}^p |Φ(X_\epsilon^t, Y_\epsilon^\eta) - Φ(x_0, y_0)|^p$ is bounded by Lemma [2] above and [22 Corollary 1]. Meanwhile, the Riemann integral terms are bounded by Lemma [3] (separating the two factors of the product $σ(Y_\eta^u)(s - u)^{2H-2}$ that appears in [12] by, for example, Young’s inequality), and the ordinary Brownian integral term is bounded by the Burkholder-Davis-Gundy inequality and Lemma [3].

It remains only to bound the stochastic integral term $E \sup_{0 \leq t \leq T} \left| \int_0^t (\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \delta W_s^H \right|^p$. The maximal inequality stated after (2.14) in [12] gives a satisfactory bound. To complete the proof of the theorem, it therefore suffices to verify that the integrand $(\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta)$ is in the appropriate class, i.e., that
\[
E \left( \left| (\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \right|^{pL_{1/H}(0,T)} + |D(\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta)|^{pL_{1/H}(0,T)^2} \right) < \infty,
\]
uniformly in $ε$ sufficiently small.

The first summand, $E \left| (\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \right|^{pL_{1/H}(0,T)}$, is easily handled by Lemma [2] above and [22 Corollary 1], so we proceed to consider the second summand. For this, we have by Jensen’s inequality and then Young’s inequality with conjugate exponents $ℓ$ and $\frac{ℓ}{ℓ-1}$,
\[
E \left| D(\nabla_x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \right|^{pL_{1/H}(0,T)^2} = E \left( \int_0^T \int_0^T \left| (\nabla^2 x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \cdot D_s X_\epsilon^t \right|^p ds dt \right)^{\frac{ℓ}{ℓ-1}} \leq T^{2pH-2} E \int_0^T \left| D_s X_\epsilon^t \right|^p \left| (\nabla^2 x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \right|^p ds dt \\
\leq I + II,
\]
where
\[
I := \frac{T^{2pH-2}}{ℓ} E \int_0^T \int_0^T \left| D_s X_\epsilon^t \right|^p ds dt \\
II := \frac{T^{2pH-2}(ℓ-1)}{ℓ} \int E \int_0^T \left| (\nabla^2 x Φσ)(X_\epsilon^t, Y_\epsilon^\eta) \right|^p \left( \int_0^T dt \right)^{\frac{ℓ-1}{ℓ}} dt.
\]
Choosing ℓ > 1 sufficiently small, the term I is handled by Lemma 5 while the term II is easily handled by Lemma above and Corollary 1. This completes the proof of the theorem.

We are now ready to prove Theorem 1 based on the above ergodic theorem.

Proof of Theorem 1. Given Theorem 3, the argument is as in the proof of Theorem 1.

4. Second-Order Limit

This section is dedicated to proving Theorem 2 which establishes a limit in distribution of the (appropriately rescaled) fluctuations of \(X^\varepsilon\) about its deterministic typical behavior \(\bar{X}\). We denote the fluctuations process by \(\theta^\varepsilon := \frac{1}{\sqrt{\varepsilon}}(X^\varepsilon - \bar{X})\). We then have the following decomposition:

\[
\theta^\varepsilon = I^\varepsilon + II^\varepsilon + III^\varepsilon,
\]

where, for \(0 \leq t \leq T\),

\[
I^\varepsilon_t = \frac{1}{\sqrt{\varepsilon}} \int_0^t (\bar{c}(X^\varepsilon_s) - \bar{c}(\bar{X}_s)) \, ds,
\]

\[
II^\varepsilon_t = \frac{1}{\sqrt{\varepsilon}} \int_0^t (c(X^\varepsilon_s, Y^\varepsilon_s) - \bar{c}(X^\varepsilon_s)) \, ds,
\]

\[
III^\varepsilon_t = \int_0^t \sigma(Y^\varepsilon_s) \delta W_s^H.
\]

Lemma 7. Suppose that \(\eta = \eta(\varepsilon)\) and that \(\lim_{\varepsilon \to 0} \frac{\sqrt{T}}{\varepsilon} =: \lambda \in [0, \infty)\). Assume Conditions 4 and 5 \((\alpha, \beta, \gamma)\), where \(\alpha \geq 0, \beta \geq 2, \gamma \geq 0,\) and \(T \beta \gamma \sup_{y \in Y} \|\tau(y)\|_{y}^{2} < 2a\). For some \(\varepsilon_0 > 0\), one has tightness of the family of distributions on \(C([0, T]; X^\varepsilon)\) (endowed, as usual, with the topology of uniform convergence) associated with the family of processes \(\{\Theta^\varepsilon\}_{\varepsilon < \varepsilon_0}\), where \(\Theta^\varepsilon := (\theta^\varepsilon, I^\varepsilon, II^\varepsilon, III^\varepsilon)\).

Proof. As in Theorem 7.3 in [3], a family \(\{\Pi^\varepsilon\}_{\varepsilon < \varepsilon_0}\) represents a tight family of distributions if and only if for all \(\zeta > 0\),

\[
\exists N \in \mathbb{N}: \sup_{0 \leq \varepsilon < \varepsilon_0} P \left[ \sup_{0 \leq t \leq T} |\Pi^\varepsilon_t| \geq N \right] \leq \zeta
\]

and

\[
\forall M \in \mathbb{N}, \lim_{\varepsilon \to 0} \sup_{0 \leq \varepsilon < \varepsilon_0} P \left[ \sup_{0 \leq t_1, t_2 \leq T, |t_1 - t_2| < \rho} |\Pi^\varepsilon_{t_1} - \Pi^\varepsilon_{t_2}| \geq \zeta, \sup_{0 \leq t \leq T} |\Pi^\varepsilon_t| \leq M \right].
\]

Applying the triangle inequality in conjunction with this characterization, it is enough to show that \(\varepsilon_0 > 0\) may be chosen so that each family \(\{I^\varepsilon\}_{\varepsilon < \varepsilon_0}, \{II^\varepsilon\}_{\varepsilon < \varepsilon_0}, \{III^\varepsilon\}_{\varepsilon < \varepsilon_0}\) represents a tight family of distributions.

Let us first consider separately the family \(\{I^\varepsilon\}\),

\[
I^\varepsilon_t = \int_0^t (\nabla x \bar{c})(\bar{X}_s) \cdot \theta^\varepsilon_s \, ds + \int_0^t [(\nabla x \bar{c})(X^\varepsilon_s) - (\nabla x \bar{c})(\bar{X}_s)] \cdot \theta^\varepsilon_s \, ds
\]

\[
= \int_0^t (\nabla x \bar{c})(\bar{X}_s) \cdot \theta^\varepsilon_s \, ds + \mathcal{R}^\varepsilon_{t, t},
\]

where \(X^\varepsilon_t\) is an appropriately-chosen point on the segment connecting \(X^\varepsilon_s\) with \(\bar{X}_s\).

By Theorem 1, one may choose an \(\varepsilon_0 > 0\) for which \(\sup_{0 \leq t \leq T} |\theta^\varepsilon_t|\) is bounded in probability uniformly in \(0 < \varepsilon < \varepsilon_0\). The criteria \(13\) and \(\ref{eq:14}\) are then obviously satisfied with \(\left\{ t \mapsto \int_0^t (\nabla x \bar{c})(\bar{X}_s) \cdot \theta^\varepsilon_s \, ds \right\}_{0 < \varepsilon < \varepsilon_0}\) in the role of \(\{\Pi^\varepsilon\}_{0 < \varepsilon < \varepsilon_0}\), whence it follows that \(\left\{ t \mapsto \int_0^t (\nabla x \bar{c})(\bar{X}_s) \cdot \theta^\varepsilon_s \, ds \right\}_{0 < \varepsilon < \varepsilon_0}\) is tight. Meanwhile,

\[
\mathcal{R}^\varepsilon_{t, t} := \int_0^t [(\nabla x \bar{c})(X^\varepsilon_s) - (\nabla x \bar{c})(\bar{X}_s)] \cdot \theta^\varepsilon_s \, ds
\]
vanishes in probability uniformly in \( t \in [0, T] \) as \( \varepsilon \to 0 \). In order to see this, note first that
\[
\int_0^t [(\nabla_\varepsilon \bar{c})(X_{\varepsilon}^s) - (\nabla_\varepsilon \bar{c})(X_s)] \, ds
\]
vanishes in probability by compactness of \([0, T] \), continuity of \( \nabla_\varepsilon \bar{c} \), and Theorem \( \text{IV} \) and secondly that \( \sup_{0 \leq t \leq T} |\theta_t^\varepsilon| \) is bounded in probability by Theorem \( \text{IV} \). It is easy to deduce that \( \{R_{I,t}^\varepsilon\}_{0 < \varepsilon < \varepsilon_0} \) is tight. It follows then that \( \{I^\varepsilon\}_{0 < \varepsilon < \varepsilon_0} \) is tight.

Let us now consider separately the family \( \{II^\varepsilon\}_\varepsilon \). As mentioned before, by Theorem 3 in \( \text{[23]} \), the equations
\[
\begin{align*}
\mathcal{L} \Phi(x, y) &= -(c(x, y) - \bar{c}(x)) \\
\int_Y \Phi(x, y) d\mu(y) &= 0
\end{align*}
\]
admit a unique solution \( \Phi \) in the class of functions that grow at most polynomially in \( |y| \) as \( y \to \infty \). Applying Lemma \( \text{I} \) with \( F = \Phi \) and rearranging terms gives
\[
II_t^\varepsilon := \frac{1}{\sqrt{\varepsilon}} \int_0^t \left( c(X_{\varepsilon}^s, Y_\varepsilon^s) - \bar{c}(X_{\varepsilon}^s) \right) ds
\]

\[
= \frac{\sqrt{\eta}}{\sqrt{\varepsilon}} \left( \sqrt{\eta} \left( \Phi(x_0, y_0) - \Phi(X_{\varepsilon}^0, Y_\varepsilon^0) \right) \right) + \frac{\sqrt{\eta}}{\sqrt{\varepsilon}} \int_0^t (\nabla_\varepsilon \Phi)(X_{\varepsilon}^s, Y_\varepsilon^s) ds
\]

\[
+ \varepsilon \sqrt{\eta} \int_0^t (\nabla_\varepsilon \Phi)(X_{\varepsilon}^s, Y_\varepsilon^s) \cdot \left( \int_0^s \sigma(Y_u^\varepsilon)(s - u)^{2H-2} du \right) ds
\]

\[
+ \frac{\sqrt{\eta}}{\sqrt{\varepsilon}} \int_0^t (\nabla_\varepsilon \Phi)(X_{\varepsilon}^s, Y_\varepsilon^s) \delta W_s^H + \int_0^t (\nabla_\varepsilon \Phi)(X_{\varepsilon}^s, Y_\varepsilon^s) dB_s
\]

(16)

where \( \alpha_H := H(2H - 1) \).

The first summand, \( \frac{\sqrt{\eta}}{\sqrt{\varepsilon}} \int_0^t (\nabla_\varepsilon \Phi)(X_{\varepsilon}^s, Y_\varepsilon^s) dB_s \), converges in distribution to \( \lambda \int_0^t \Phi(x_0, y_0) d\tilde{B}_s \), where \( \Phi := ((\nabla_\varepsilon \Phi)(\nabla_\varepsilon \Phi)^T)^{1/2} \) and \( \tilde{B} \) is a standard Brownian motion. Meanwhile, by arguments as in the proof of Theorem \( \text{[23]} \) \( R_{II,t}^\varepsilon \) vanishes in probability uniformly in \( t \in [0, T] \) as \( \varepsilon \to 0 \). It is easy to deduce that \( \{II^\varepsilon\}_{0 < \varepsilon < \varepsilon_0} \) is tight.

As for the family \( \{III^\varepsilon\}_\varepsilon \), we have
\[
III_t^\varepsilon = \int_0^t \bar{\sigma} \delta W_s^H + \int_0^t (\sigma(Y_\varepsilon^s) - \bar{\sigma}) \delta W_s^H
\]

(17)

We claim that \( R_{III,t}^\varepsilon \) vanishes in probability uniformly in \( t \in [0, T] \) as \( \varepsilon \to 0 \), or more properly, as \( \eta \to 0 \). As in the proof of Lemma \( \text{[9]} \) let \( \tilde{B} \) be a standard Brownian motion, in this case assumed to be independent of \( W^H \), and let \( \tilde{Y} \) denote the solution of the stochastic differential equation
\[
\begin{align*}
d\tilde{Y}_t &= f(\tilde{Y}_t) dt + \tau(\tilde{Y}_t) d\tilde{B}_t \\
\tilde{Y}_0 &= y_0.
\end{align*}
\]

Since \( \{Y_\varepsilon^t\}_{0 \leq t \leq T} \) has the same law as \( \{\tilde{Y}_{t/\eta}^\varepsilon\}_{0 \leq t \leq T} \) and both are independent of \( W^H \), for the purposes of this argument, one may work with either process. Thus, if we had assumed \( \sigma \) to be uniformly bounded and \( Y_\varepsilon^T \) to begin at time \( t = 0 \) in stationarity, then Theorem 4.15 in \( \text{[14]} \) would apply directly to establish the claim. Although we do not make these assumptions, our conditions nevertheless suffice to recover the desired convergence in probability, as we now proceed to explain.

The proof of \( \text{[14]} \), Theorem 4.15] relies on \( \text{[13]} \) Lemma 4.10], in which the crucial statements are based on certain decay rates for the associated Markov semigroup. We will verify the same bounds in our framework. By \( \text{[22]} \) Proposition 1], the uniform nondegeneracy of \( \tau \tau^T \) in Condition \( \text{III} \) together with the dissipativity of \( f \) in Condition \( \text{II} \) or \( \text{III} \) allow us to conclude not only that \( \tilde{Y} \) has a unique invariant measure with finite
moments of all orders, but also that for any initial condition \( y_0 \in \mathcal{Y} \), each moment of \( \tilde{Y}_t \) may be bounded uniformly in \( t \geq 0 \). Thus, polynomial bounds on \( \sigma \) are enough to obtain uniform bounds on moments of the diffusion coefficient. Moreover, denoting the invariant measure of \( Y \) by \( \mu \) and the distribution of \( \tilde{Y}_t \) by \( \mu_{t,\eta} \), Condition \( \ref{unif_convergence} \) implies convergence as \( t \to \infty \) of the total variation distance \( \text{var}(\mu_t - \mu) \) (see for example \Ref{27}). Taking all of this together, one has then

\[
|E(\sigma(Y_t^\eta) - \tilde{\sigma})| = |E(\sigma(\tilde{Y}_{t/\eta}) - \tilde{\sigma})| \\
= \left| \int_{\mathcal{Y}} \sigma(y)(\mu_{t/\eta}^{y_0} - \mu)(dy) \right| \\
\leq \left( \int_{\mathcal{Y}} |\sigma(y)|^p(\mu_{t/\eta}^{y_0} + \mu)(dy) \right)^{1/p} \left( \int_{\mathcal{Y}} |\mu_{t/\eta}^{y_0} - \mu|(dy) \right)^{1/r} \\
\leq \left( \int_{\mathcal{Y}} \lambda(1 + |y|^p)(\mu_{t/\eta}^{y_0} + \mu)(dy) \right)^{1/p} \left( \text{var}(\mu_{t/\eta}^{y_0} - \mu) \right)^{1/r} \\
\leq C_1 e^{-C_2 \frac{t}{\eta}},
\]

where \( C_1, C_2 \) are finite positive constants that depend neither on \( \eta \) nor on \( t \).

Therefore, in light of this exponential decay, the arguments of \Ref{14} carry over to our setting, and we conclude as desired that \( \mathcal{R}_{\eta,II,t} = \int_0^\delta (\sigma(Y_t^\eta) - \tilde{\sigma})dWH_t \) vanishes in probability uniformly in \( t \in [0, T] \) as \( \eta \to 0 \). Details are omitted due to the similarity of the argument. It is then easy to deduce that \( \{III_t^\eta \}_{0 < \epsilon < 0} \) is tight.

We are now ready to prove Theorem \Ref{24}

**Proof of Theorem \Ref{24}** It suffices to show that any sequence of values of \( \epsilon \) tending to 0 admits a subsequence along which the \( \theta^\epsilon \) converge in distribution to the law of \( \theta \). Let us therefore consider now an arbitrary sequence \( \{\epsilon_n\}_{n=1}^\infty \) tending to 0. By Lemma \Ref{22} passing to a subsequence \( \{\epsilon_n\}_{n=1}^\infty \), we may suppose that \( \{(\theta^{\epsilon_n}, I^{\epsilon_n}, II^{\epsilon_n}, III^{\epsilon_n})\}_{n=1}^\infty \) is convergent in distribution. By the Skorohod representation theorem, there is a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) supporting stochastic processes \( \{(\tilde{\theta}^{\epsilon_k}, \tilde{I}^{\epsilon_k}, \tilde{II}^{\epsilon_k}, \tilde{III}^{\epsilon_k})\}_{k=1}^\infty \) equal in distribution to \( \{(\theta^{\epsilon_n}, I^{\epsilon_n}, II^{\epsilon_n}, III^{\epsilon_n})\}_{k=1}^\infty \) as well as a limiting stochastic process \( (\tilde{\theta}, \tilde{I}, \tilde{II}, \tilde{III}) \) to which the former converge almost surely as \( k \) tends to infinity.

In light of the decompositions \Ref{16}, \Ref{14}, \Ref{14} and the limits identified in the proof of Lemma \Ref{22}, \( \tilde{\theta} \) has the distribution of \( \theta \), which completes the proof of the theorem.

\[ \square \]

5. AN EXTENSION OF THE MODEL

We now consider an extension of the model. Consider

\[
\begin{cases}
    dX_t^\epsilon &= \frac{\epsilon}{\sqrt{\eta}} b(X_t^\epsilon, Y_t^\epsilon)dt + c(X_t^\epsilon, Y_t^\epsilon)dt + \sqrt{\epsilon} \sigma(Y_t^\epsilon) dW_t^H \\
    dY_t^\epsilon &= \frac{1}{\eta} f(Y_t^\epsilon)dt + \frac{1}{\sqrt{\eta}} g(Y_t^\epsilon)dt + \frac{1}{\sqrt{\eta}} \tau(Y_t^\epsilon) dB_t \\
    X_0^\epsilon &= x_0 \in \mathcal{X}, \quad Y_0^\epsilon = y_0 \in \mathcal{Y}.
\end{cases}
\]

As in Theorem \Ref{24} we suppose that \( \eta = \eta(\epsilon) \) and that \( \lim_{\epsilon \to 0} \frac{\sqrt{\epsilon}}{\sqrt{\eta}} =: \lambda \in [0, \infty) \). Notice that if \( \lambda = 0 \) then the term \( \frac{\sqrt{\epsilon}}{\sqrt{\eta}} b(X_t^\epsilon, Y_t^\epsilon) \) is asymptotically singular. Accordingly, we distinguish two possibilities:

1. \( \lambda = 0 \), the ‘first regime’ or ‘homogenization regime’
2. \( \lambda \in (0, \infty) \), the ‘second regime’ or ‘averaging regime.’

The extended model \Ref{16} is particularly relevant when, for example, a fast intermediate scale forms part of the slow component. In part of the literature this is referred to as the homogenization regime (see for example \Ref{23} or \Ref{30} for related examples in the framework of perturbation by standard Brownian motion rather than fractional Brownian motion). The scaling in front of the term corresponding to the coefficient \( g \) is that which results in a nontrivial limiting contribution in the event that additional intermediate fast scales form part of the dominant fast component.
We introduce in Condition 4 our growth and regularity conditions for the new coefficients in the extended model.

**Condition 4.**

- $b$ satisfies the same smoothness and growth conditions as $c$.
- In the first regime, $b(x, y) = b(y)$ is a function of the fast variable only and not the slow, and $b$ and its derivatives grow at most polynomially.
- In the first regime, $g$ satisfies the same conditions as $c$ does in terms of the $y$-dependence; in the second regime, $g$ satisfies the same conditions as $f$.

We have as before a basic condition of recurrence type on the fast component, yielding ergodic behavior.

**Condition 5.**

$$
\lim_{|y| \to \infty} y \cdot (f + \lambda g)(y) = -\infty; \quad (19)
$$

As before we shall in fact assume a stronger recurrence condition for our main results.

**Condition 6.**

- Condition 6-($\alpha, \beta$): there is a neighborhood $\Lambda$ of $\lambda$ in $[0, \infty)$ such that one has
  $$
  \sup_{\lambda \in \Lambda} y \cdot (f + \lambda g)(y) + \alpha |y|^{\beta} + \frac{1}{2} (\beta - 2 + d - m) \sup_{\tilde{y} \in \mathcal{Y}} \tau(\tilde{y}) \leq 0
  $$
  for $|y|$ sufficiently large.
- Condition 6-($\alpha, \beta, \gamma$): Condition 6-($\alpha, \beta$) holds and, moreover, one has, in the first regime, $||\nabla_x c(x, y)|| \leq \gamma |y|^{\beta}$ for $|y|$ sufficiently large, and in the second regime, perhaps for a smaller neighborhood $\Lambda$,
  $$
  \sup_{\lambda \in \Lambda} ||\nabla_x b(x, y) + \nabla_x c(x, y)|| \leq \gamma |y|^{\beta}
  $$
  for $|y|$ sufficiently large.

One has the limiting infinitesimal generator

$$
\mathcal{L} := (f + \lambda g) \cdot \nabla_y + \frac{1}{2} (\tau \tau^T) : \nabla^2_y
$$

for the rescaled fast dynamics. Conditions 4, 5, and 6 are enough to guarantee that one has on $\mathcal{Y}$ a unique invariant measure $\mu$ corresponding to the operator $\mathcal{L}$ in equation (20), as discussed for example in [22] and [27].

In the first regime, a standard centering condition tempers the asymptotic singularity of the term $\sqrt{\epsilon} \sqrt{\eta} b(X_\epsilon, Y_\epsilon) = \sqrt{\epsilon} \sqrt{\eta} b(Y_\epsilon)$ (recall that in this regime, we assume that $b$ is a function of the fast variable only and not of the slow one).

**Condition 7.**

$$
\int_{\mathcal{Y}} b(y) d\mu(y) = 0.
$$

The above conditions are sufficient to derive a first-order limit for the slow process $X_\epsilon$ in the context of the extended model [18]. In order to obtain a second-order limit, we assume that the convergence of $\sqrt{\epsilon} \sqrt{\eta}$ to $\lambda$ takes place at a particular rate. Precisely, we assume that $\lim_{\epsilon \to 0} \frac{1}{\sqrt{\epsilon}} (\sqrt{\epsilon} \sqrt{\eta} - \lambda) =: \kappa \in \mathbb{R}$.

We now sketch how to extend the results of the paper to the extended model. In the first regime, we must carefully consider the limiting contribution of the asymptotically-singular term $\sqrt{\epsilon} \sqrt{\eta} b(X_\epsilon, Y_\epsilon) = \sqrt{\epsilon} \sqrt{\eta} b(Y_\epsilon)$ to the dynamics of the slow process $X_\epsilon$ (recall that in this regime, we assume that $b$ is a function of the fast variable only and not of the slow one). It turns out that under Condition 7, the limiting contribution may be captured in terms of the solution of a certain Poisson equation. By Theorem 3 in [23], the equations

$$
\begin{cases}
\mathcal{L} \Psi(y) = -b(y)
\quad \\
\int_{\mathcal{Y}} \Psi(y) d\mu(y) = 0
\end{cases}
$$
admit a unique solution $\Psi$ in the class of functions that grow at most polynomially in $|y|$ as $y \to \infty$.

In the first regime, we will need the auxiliary drift coefficient

$$\varphi_1(x, y) := (\nabla_y \Psi \cdot g)(y) + c(x, y), \tag{21}$$

where $\Psi$ is as in (21). To play the same role in the second regime, we will need the auxiliary drift coefficient

$$\varphi_2(x, y) := (\lambda^{-1} b + c)(x, y). \tag{22}$$

Finally, note that in the Itô formula (Lemma 1), when one considers the extended model, two additional terms, $\frac{\sqrt{\tau}}{\sqrt{\eta}} \int_0^t \int \nabla_y Fb)(X^*_s, Y^*_s) ds$ and $\frac{\sqrt{\tau}}{\sqrt{\eta}} \int_0^t \int \nabla_y Fg)(X^*_s, Y^*_s) ds$, appear on the right hand side.

We are now ready to state our asymptotic theorems for the extended model.

**Theorem 4.** Suppose that $\eta = \eta(\epsilon)$ and that $\lim_{\epsilon \to 0} \sqrt{\epsilon} =: \lambda \in [0, \infty)$. Let $\star \in \{1, 2\}$ indicate, respectively, the first or second regime. Assume Conditions 1, 4, and 6- (22) and (23), where $\alpha \geq 0$, $\beta \geq 2$, $\gamma \geq 0$, and $\lambda \beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2 < 2\alpha$; in the first regime, assume also Condition 4. For any $0 < p < \frac{2n}{\lambda \beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2}$, there is a constant $K$ such that for $\epsilon := (\epsilon, \eta)$ sufficiently small,

$$E \sup_{0 \leq t \leq T} |X^\epsilon_t - \bar{X}^\epsilon_t|^p \leq K \left( \sqrt{\epsilon}^p + \sqrt{\eta}^p \right),$$

where $\bar{X}^\epsilon$ is the (deterministic) solution of the integral equation

$$\bar{X}^\epsilon_s = x_0 + \int_0^t \varphi_\star(\bar{X}^\epsilon_s) ds,$$

where $\varphi_\star$ is obtained, depending on the regime, by averaging (21) or (22) with respect to the invariant measure $\mu$.

**Proof.** The proof is almost exactly the same as that of Theorem 1 except that in establishing the analogue of Lemma 1, in the first regime we must now consider carefully the asymptotically-singular term. Letting $\Psi$ be as in (21), applying the Itô lemma with $F = \Psi$, and rearranging terms, we obtain

$$\frac{\sqrt{\epsilon}}{\sqrt{\eta}} \int_0^t b(Y^\epsilon_s) ds = \sqrt{\epsilon} \eta \left( \Psi(y_0) - \Psi(Y^\epsilon_t) \right) + \int_0^t (\nabla_y \Psi g)(Y^\epsilon_s) ds + \sqrt{\epsilon} \int_0^t (\nabla_y \Psi \tau)(Y^\epsilon_s) dB_s$$

$$=: \int_0^t (\nabla_y \Psi g)(Y^\epsilon_s) ds + R^\epsilon_t,$$

where $\alpha_H := H(2H - 1)$. Here, $R^\epsilon$ vanishes. Thus, the proof may proceed as before with $\varphi_1$ in place of $c$.

To study the distribution of the fluctuations we must quantify more precisely the difference between the true drift and the approximate drift, as was done in formulating the Poisson equation (16) in the proof of Theorem 1. By Theorem 3 in [21], with $\star \in \{1, 2\}$ indicating the regime, the equations

$$\mathcal{L} \Phi_\star(x, y) = - (\varphi_\star(x, y) - \varphi_\star(x))$$

admit a unique solution $\Phi_\star$ in the class of functions that grow at most polynomially in $|y|$ as $y \to \infty$.

**Theorem 5.** Suppose that $\eta = \eta(\epsilon)$ and that $\lim_{\epsilon \to 0} \sqrt{\epsilon} =: \lambda \in [0, \infty)$. Suppose moreover that $\lim_{\epsilon \to 0} \frac{\sqrt{\epsilon}}{\sqrt{\eta}} \left( \frac{\sqrt{\epsilon}}{\sqrt{\eta}} - \lambda \right) =: \kappa \in \mathbb{R}$. Assume Conditions 1, 4, and 12, where $\alpha \geq 0$, $\beta \geq 2$, $\gamma \geq 0$, and $\lambda \beta \gamma \sup_{y \in \mathcal{Y}} ||\tau(y)||^2 < 2\alpha$; in the first regime, assume also Condition 4. With $\Psi$ and $\Phi_\star$, respectively, as in (21) and (24), set $\Sigma_\Psi := ((\nabla_y \Phi_\Psi)(\nabla_y \Phi_\Psi)^T)^{1/2}$ and $\Sigma_{\Phi_\star} := ((\nabla_y \Phi_\Psi)(\nabla_y \Phi_\Psi)^T)^{1/2}$.

In the first regime, the family of processes $\{\theta^\epsilon\}_\epsilon$ converges in distribution as $\epsilon \to 0$ to the law of the solution $\theta_1$ of the mixed SDE

$$d\theta_1, t = \int_0^t (\nabla_x \varphi_1)(\bar{X}_{1, s}) \cdot \theta_1 ds + \kappa \int_0^t \nabla_y \Phi_1 \cdot g(\bar{X}_{1, s}) ds$$

$$+ \int_0^t \Sigma_\Psi(\bar{X}_{1, s}) d\tilde{B}_s + \int_0^t \tilde{\sigma} d\tilde{W}^H_s,$$
where $\tilde{W}^H$ is a fractional Brownian motion with Hurst index $H$ and $\tilde{B}$ is a standard Brownian motion independent of $W^H$. We point out that under our assumptions in this regime we have in fact $\nabla_x \tilde{\varphi}_1 = \nabla_x \epsilon$.

In the second regime, the family of processes $\{\theta_t\}$ converges in distribution as $\epsilon \to 0$ to the law of the solution $\theta_2$ of the mixed SDE

$$
d\theta_{2,t} = \int_0^t (\nabla_x \tilde{\varphi}_2)(\tilde{X}_{2,s}) \cdot \theta_{2,s} ds + \kappa \int_0^t \sqrt{\chi_2} \tilde{\varphi}_2 \cdot \tilde{g}(\tilde{X}_{2,s}) ds - \frac{\kappa}{\chi_2} \int_0^t \tilde{b}(\tilde{X}_{2,s}) ds + \lambda \int_0^t \Sigma_{\Phi_2}(\tilde{X}_{2,s}) d\tilde{B}_s + \int_0^t \tilde{\sigma} d\tilde{W}_s^H,
$$

where $\tilde{W}^H$ is a fractional Brownian motion with Hurst index $H$ and $\tilde{B}$ is a standard Brownian motion independent of $W^H$.

**Proof.** Given Theorem 4 and, in particular, the representation (23), the proof is nearly identical to that of Theorem 2

\[ \square \]

**Appendix A. Preliminaries**

### A.1. Fractional Brownian motion.

A fractional Brownian motion (fBm) is a centered Gaussian process $W^H = \{W^H_t: t \geq 0\}$ characterized by its covariance function, given by

$$R_H(t,s) = E(W^H_t W^H_s) = \frac{1}{2} \left( s^{2H} + t^{2H} - |t-s|^{2H} \right).$$

The parameter $H \in (0,1)$ is usually referred to as the Hurst parameter or Hurst index, and controls the regularity of the paths of the process, which are Hölder continuous of order $H - \varepsilon$, for any $\varepsilon > 0$.

Note that if $H = \frac{1}{2}$, $R^H_x(t,s) = t \wedge s$, so that $W^H$ is a standard Brownian motion, so that its increments are independent. If $H \neq \frac{1}{2}$, the increments of the fractional Brownian motion are still stationary, but no longer independent. In particular, when $H > \frac{1}{2}$, the process exhibits long-range dependence. Furthermore, when $H \neq \frac{1}{2}$, the fractional Brownian motion is not a semimartingale (so that the usual Itô calculus is not available in this framework).

Another noteworthy property of fractional Brownian motion is that it is self-similar in the sense that, for any constant $a > 0$, the processes $\{W^H_t\}_{t \geq 0}$ and $\{a^{-H} W^H_{at}\}$ have the same distribution.

For more details about fractional Brownian motion, we refer the reader to the monographs \[20\] \[21\] \[3\].

The self-similarity and long memory properties of the fractional Brownian motion make it an interesting and suitable input noise in many models in various fields such as financial time series, hydrology, and telecommunications. However, in order to develop interesting models based on fractional Brownian motion, one needs a stochastic calculus with respect to the fBm, which will make use of the stochastic calculus of variations, or Malliavin calculus, introduced in the next subsection.

### A.2. Elements of Malliavin calculus.

We gather here the main tools of Malliavin calculus needed in this paper. For a complete treatment of this topic, we refer the reader to \[21\]. Fix a time interval $[0,T]$ and consider a fractional Brownian motion $W^H = \{W^H_t\}_{t \geq 0}$ with Hurst index $H \in (\frac{1}{2},1)$. Denote by $\mathcal{E}$ the set of step functions on $[0,T]$, and let $\mathcal{H}$ be the Hilbert space defined as the completion of $\mathcal{E}$ with respect to the inner product

$$\langle \chi_{[0,t]}, \chi_{[0,s]} \rangle_{\mathcal{H}} := R_H(t,s).$$

It can be shown, using fractional calculus, that the inner product in $\mathcal{H}$ has the explicit representation (see \[4\])

$$\langle \varphi, \psi \rangle_{\mathcal{H}} = \alpha_H \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T \int_0^T |r-u|^{2H-2} \varphi(r) \psi(u) du dr,$$

for any $\varphi, \psi \in \mathcal{H}$, and where $\alpha_H := H(2H-1)$. Whenever $H > \frac{1}{2}$, as is the case here, the elements of the Hilbert space $\mathcal{H}$ may not be functions, but distributions of negative order (see \[25\] \[26\]). One can make use of the inner product representation (25) in
We can now introduce the divergence operator (or divergence integral) \(H\)-operator. By definition, an \(H\)-valued random variable is the closure of the derivative operator with domain given by the set of \(H\)-valued random variables of the form

\[
\varphi(r) = \int_0^r |\varphi(u)| |r-u|^{2H-2} \psi(u) du dr < \infty.
\]

Note that we have the following chain of continuous embeddings:

\[
L^{\frac{1}{2}}([0,T]) \subset |\mathcal{F}| \subset \mathcal{F}.
\]

Denote by \(\mathcal{S}\) the set of smooth cylindrical random variables of the form \(F = f(W^H(\varphi_1), \ldots, W^H(\varphi_n))\), \(n \geq 1\), \(\varphi_i \in \mathcal{F}\), \(1 \leq i \leq n\), and \(f \in C^\infty_b(\mathbb{R}^n)\) (it and all its partial derivatives are bounded).

The Malliavin derivative of such a smooth cylindrical random variable \(F\) is defined as the \(\mathcal{F}\)-valued random variable given by

\[
DF = \sum_{i=1}^n \frac{\partial f}{\partial x_i} (W^H(\varphi_1), \ldots, W^H(\varphi_n)) \varphi_i.
\]

The derivative operator \(D\) is a closable operator from \(L^2(\Omega)\) into \(L^2(\Omega; \mathcal{F})\), and we continue to denote by \(D\) the closure of the derivative operator with domain given by the closure \(\mathbb{D}^{1,2}\) of the set \(\mathcal{S}\) with respect to the Sobolev-type norm

\[
\| F \|_{1,2}^2 := E(F^2) + E\left( \| DF \|_{\mathcal{F}}^2 \right).
\]

Note that one can also replace the \(L^2(\Omega)\) norms in the above definition by \(L^p(\Omega)\) norms for any \(p > 1\) to obtain the Sobolev spaces \(\mathbb{D}^{1,p}\) instead.

We can now introduce the divergence operator (or divergence integral) \(\delta\) as the adjoint of the derivative operator. By definition, an \(\mathcal{F}\)-valued random variable \(u \in L^2(\Omega; \mathcal{F})\) is in the domain of \(\delta\), denoted by \(\text{dom } \delta\), if

\[
E \left( \langle DF, u \rangle_{\mathcal{F}} \right) \leq c_u \| F \|_{L^2(\Omega)}
\]

for any \(F \in \mathbb{D}^{1,2}\). In this case, \(\delta(u)\) is defined by duality as the element of \(L^2(\Omega)\) such that

\[
E (F \delta(u)) = E \left( \langle DF, u \rangle_{\mathcal{F}} \right)
\]

for any \(F \in \mathbb{D}^{1,2}\). Furthermore, it holds that \(\mathbb{D}^{1,2}(\mathcal{F}) \subset \text{dom } \delta\), and for any \(u \in \mathbb{D}^{1,2}(\mathcal{F})\),

\[
E \left( \delta(u)^2 \right) = E \left( \| u \|_{\mathcal{F}}^2 \right) + E \left( \langle Du, (Du)^* \rangle_{\mathcal{F} \otimes \mathcal{F}} \right)
\]

where \((Du)^*\) is the adjoint of \(Du\) in the Hilbert space \(\mathcal{F} \otimes \mathcal{F}\).

### A.3. Stochastic integration with respect to fractional Brownian motion.

There are two main methods to define stochastic integrals with respect to the fractional Brownian motion. We refer the reader to the monograph [3] which gathers most of the available approaches to stochastic integration with respect to fractional Brownian motion.

The first approach is based on the stochastic calculus of variations, or Malliavin calculus, and was introduced in [4]. It is based on the divergence operator introduced in Subsection A.2 and stochastic integrals in this context will be referred to as divergence integrals.

The second approach, called the pathwise approach, makes use of the Hölder regularity of the paths of fBm which, in the case where \(H > \frac{1}{2}\), exhibit enough regularity to define integration in the sense of Zähle [32] or [28] (see also [31] for integration based on Hölder regularity). As noted above, this approach is restricted to the case \(H > \frac{1}{2}\) and stochastic integrals in this framework will be referred to as pathwise integrals.

**Remark 5.** Note that even if the divergence approach allows to define stochastic integration for any \(H \in (0,1)\), we restrict ourselves to the case \(H > \frac{1}{2}\) so that both approaches can coexist and relations between the different definitions can be exploited.
A.3.1. *Divergence integration.* As the adjoint of the Malliavin derivative operator, the divergence operator \( \delta \) can be interpreted as a stochastic integral, especially given the fact, in the standard Brownian motion case \((H = \frac{1}{2})\), the divergence of an adapted, Itô-integrable process coincides with the Itô integral of said process. We can hence define, for any element \( u \in \text{dom} \delta \), the indefinite integral

\[
\int_0^t u_s \delta W_s^H := \delta(u\chi_{[0,t]}),
\]

which we refer to as the divergence integral of \( u \). Note that, by construction, the divergence integral is centered.

Define \( L^{1,p}_{H} \) as the subset of \( D^{1,p}(\mathcal{F}) \) such that

\[
E\left( \left\| u \right\|_{L^p[0,T]}^p + \left\| Du \right\|_{L^p((0,T]^2)}^p \right) < \infty.
\]

If \( u \in L^{1,p}_{H} \) with \( pH > 1 \) and

\[
\int_0^T |E(u_s)|^p ds + \int_0^T E\left( \int_0^T |D_s u_r|^p ds \right)^{pH} dr < \infty,
\]

then one has the maximal inequality

\[
E\left( \sup_{t \in [0,T]} \left| \int_0^t u_s \delta W_s^H \right|^p \right) \leq C \left[ \int_0^T |E(u_s)|^p ds + \int_0^T E\left( \int_0^T |D_s u_r|^p ds \right)^{pH} dr \right],
\]

where the constant \( C > 0 \) only depends on \( H \) and \( T \) (see [1] for more details on the maximal inequality).

A.3.2. *Pathwise integration.* The version of pathwise integration we present here appeared in [2 8], where it is referred to as symmetric stochastic integration.

Let \( u = \{u_t : t \in [0,T]\} \) be a stochastic process in \( D^{1,2}(\mathcal{F}) \). Assume that

\[
E\left( \left\| u \right\|_{\mathcal{F}}^2 + \left\| Du \right\|_{\mathcal{F} \otimes \mathcal{F}}^2 \right) < \infty
\]

and

\[
\int_0^T \int_0^T |D_s u_r| |t - s|^{2H-2} dsdt < \infty \ a.s.
\]

Then the symmetric integral

\[
\int_0^T u_t dW_t^H,
\]

defined as the limit in probability as \( \varepsilon \) tends to zero of

\[
\frac{1}{2\varepsilon} \int_0^T u_s \left( W_{s+\varepsilon}^H \wedge T - W_{s-\varepsilon}^H \vee 0 \right) ds,
\]

exists and we have, for any \( t \in [0,T] \), a relation between the pathwise and divergence integrals, given by

\[
\int_0^t u_s dW_s^H = \int_0^t u_s \delta W_s^H + \alpha_H \int_0^t \int_0^T D_r u_s |s - r|^{2H-2} drds.
\]

Note that whereas the divergence integral is centered, the pathwise integral is usually not centered. Under specific circumstances, it can however be centered, which the following remark addresses.

**Remark 6.** Note that if the integrand \( u \) in the pathwise integral is independent of the fractional Brownian motion \( W^H \), then \( Du = 0 \), so that we have

\[
\int_0^t u_s dW_s^H = \int_0^t u_s \delta W_s^H,
\]

which makes the two approaches coincide and makes the pathwise integral centered.
