Abstract

Our first result is a stochastic sewing lemma with quantitative estimates for mild incremental processes, with which we study SPDEs driven by fractional Brownian motions in a random environment. We obtain uniform $L^p$-bounds. Our second result is a fractional averaging principle admitting non-stationary fast environments. As an application, we prove a fractional averaging principle for SPDEs.

MSC2010: 60G22, 60H10, 37A25.

Keywords: Fractional Brownian motion, fractional averaging, non-stationary fast process.

1 Introduction

2 Pathwise Solutions
   2.1 Preliminaries
   2.2 Mild Hölder Spaces and Mild Sewing Lemma
   2.3 Pathwise Well-Posedness and a Deterministic Stability Lemma

3 Mild Stochastic Sewing Lemma

4 Uniform $L^p$-Estimates
   4.1 Trace-Class Fractional Brownian Motions
   4.2 Basic Properties of Integrals Against Trace-Class fBm
   4.3 $L^p$-Estimates and Stochastic Stability

5 Non-Stationary Averaging
   5.1 Quantitative Ergodic Theorem
   5.2 Examples

6 Application to Slow-Fast Systems

*xue-mei.li@imperial.ac.uk. Research support by the EPSRC (grants nos: EP/S023925/1 and EP/V026100/1).
†j.sieber19@imperial.ac.uk.
1 Introduction

A. This article is written on the occasion of Björn Schmalfuß’ 65th birthday. There had been significant novel developments both on stochastic analysis of and on stochastic dynamics driven by fractional Brownian motions—these topics are at the heart of Björn Schmalfuß’s research. Instead of reporting on recent progresses, this article contains original research which deserves both a careful treatment and which also illustrates recent directions in two-scale random dynamics driven non-Markovian processes. We propose to study stochastic partial differential equations (SPDEs) in a random environment. The driving noise \( B \) is a fractional Brownian motion (fBm) with trace-class covariance operator \( Q \), which takes values in a separable Hilbert space \( K \) and is independent of the fast environment. The Hurst parameter \( H \) of \( B \) is assumed to be greater than \( \frac{1}{2} \). Fractional noise with large Hurst parameter models noise increments with long range dependence (LRD). The LRD phenomenon is prevalent in data from social economical and natural sciences.

The solution of the SPDE takes values in a separable Hilbert space \( \mathcal{H} \) and the environment process takes its value in a complete separable metric space \( \mathcal{Y} \). Let \( f : \mathcal{H} \times \mathcal{Y} \to \mathcal{H}, g : \mathcal{H} \times \mathcal{Y} \to \mathcal{L}(K, \mathcal{H}) \), and \( A : \mathcal{H} \to \mathcal{H} \) a densely defined closed linear operator with \( \mathcal{D}(A) \) generating a bounded analytical semigroup \((S_t)_{t \geq 0}\). We consider

\[
    dX_t = AX_t \, dt + f(X_t, Y_t) \, dt + g(X_t, Y_t) \, dB_t. \tag{1.1}
\]

The integral in the SPDE is interpreted as Riemann-Stieltjes integrals for each fixed sample path \( B(\omega) \). We say that a path \( X : [0, T] \to \mathcal{H} \) is a mild solution to (1.1) if, for any \( t \leq T \),

\[
    X_t(\omega) = S_tX_0(\omega) + \int_0^t S_{t-s}f(X_s(\omega), Y_s(\omega)) \, ds + \int_0^t S_{t-s}g(X_s(\omega), Y_s(\omega)) \, dB_s(\omega) \quad \mathbb{P}\text{-a.s.} \tag{1.2}
\]

Preliminary results on the pathwise calculus are presented in § 2.

B. Our first technical result is a ‘mild’ version of the stochastic sewing lemma. This lemma, presented in Proposition 3.2, generalizes the sewing lemma of Gubinelli, Feyel, and de La Pradelle [21, 13], the stochastic sewing lemma of Lê [31], and the mild sewing lemma of Gubinelli and Tindel [22, Theorem 3.5], see also Gerasimović and Hairer [20]. For the mild stochastic sewing lemma the incremental process takes into account the interaction between the linear motion of \( A \) and the nonlinear part. The mild stochastic sewing lemma is convenient for studying mild solutions of SPDEs, its first application is a uniform \( L^p \)-bounds for the solution of the SPDE in a class of random environments. Stochastic sewing lemma in Banach spaces were recently obtained by Lê [32] and by Athreya, Butkovsky, Lê, and Mytnik [1], albeit both with the usual increment process. After submitting this to the arXiv, we learnt from Lê a nice trick allowing to obtain the sewing part of the lemma from that in [32], while the error estimate does not follow directly from the corresponding result in their latest article [1].

We then turn to the main topic of this article: stochastic evolution equations in a random environment evolving at a faster scale. In a stochastic slow-fast model, the slow variables evolve in the random environment modelled by the fast random motion. The principle of fractional averaging leverages ergodicity of the fast variable to obtain an approximation of the slow variables \((X^\varepsilon_t)_{t \in [0,T]}\) as \( \varepsilon \to 0 \) whose microscopic structure resembles a fractional noise \( B \). Despite of the statistical evidence, such slow-fast stochastic models have been overlooked because of the lack of tools for studying them. Stochastic averaging for noise with LRD has attracted a lot of attention recently, which has led to renovation of both techniques and results. Non-Markovian multi-scale dynamics only very recently gained traction, see e.g. [24, 19, 18, 17, 12, 3, 4, 40, 41, 42]. Pathwise well posedness of the equations and bounds on its solutions are given in § 2; uniform \( L^p \) bounds of solutions in \( \varepsilon \) are obtained with the mild stochastic sewing lemma in the subsequent sections.
C. Our next technical result is an ergodic theorem allowing to obtain a fractional averaging principle for non-stationary and possibly non-Markovian environment. Suppose that the random environment evolves on the *macroscopic* time-scale $\varepsilon^{-1}$ and converges to a stationary distribution $\pi$ with algebraic rate $t^{-\delta}$ and $f, g$ satisfy suitable analytic conditions. One defines

$$\bar{f}(x) \triangleq \int_Y f(x, y)\pi(dy), \quad \bar{g}(x) \triangleq \int_Y g(x, y)\pi(dy).$$

The key to the averaging theorem is to show that $|f(\cdot, Y^\varepsilon_{\cdot}) - \bar{f}|_{-\kappa, \gamma} \to 0$ (and similarly for $g$), in a Hölder norm with negative exponent $-\kappa < 0$ in time. This was shown in [24] for a mixing stationary environment and both $\mathcal{H}, \mathcal{K}$ finite-dimensional Euclidean spaces. In § 5, we obtain an ergodic theorem for non-stationary fast processes allowing also correlation of the increments:

$$\|g_\varepsilon(\cdot, x) - g_\varepsilon(\cdot, z) - (\bar{g}(x) - \bar{g}(z))\|_{C^{-\kappa}(L(\mathcal{K}, \mathcal{H}))} \leq \varepsilon^\frac{\delta}{2},$$

where $g_\varepsilon(t, x) = g(x, Y^\varepsilon_{\cdot})$ and similarly for $f$. See Lemma 5.2 and the subsequent examples. [These are the ergodic conditions needed for fractional averaging, see Definition 5.1 and Theorem 6.3 for details.] The ergodic theorem nicely illustrates the fundamental differences between Markovian and non-Markovian dynamics. We show that the ergodic theorem actually follows from the total variation convergence of the conditioned dynamics, which can often be easily verified for Markovian fast processes. If $Y$ is also the solution to a fractional stochastic equation, its disintegration is usually governed by a different dynamics and no longer solves the original equation, see [33] for a detailed discussion.

D. As an application of the ergodic theorem and the mild stochastic sewing lemma, we obtain a fractional averaging principle for a two-scale SPDE: Suppose that $Y$ has the $C^{\frac{\gamma}{2}}$ regularity of a typical diffusion process and consider the SPDE

$$dX^\varepsilon_t = AX^\varepsilon_t \, dt + f(X^\varepsilon_t, Y^\varepsilon_{\cdot}) \, dt + g(X^\varepsilon_t, Y^\varepsilon_{\cdot}) \, dB_t, \quad X^\varepsilon_0 = X_0. \quad (1.3)$$

As $\varepsilon \to 0$ we expect $X^\varepsilon$ to be well approximated by an *effective* autonomous dynamics $\bar{X}$ obtained through averaging of the coefficients. These kind of problem have a long history in the theory of dynamical systems and stochastic equations; in fact stochastic averaging for Markovian systems came up shortly after stochastic integration was introduced [2, 25]. The typical questions are: What is the nature of the effective dynamics and what is the mode of convergence?

These questions are answered by our fractional averaging principle presented in Theorem 6.3. It is obtained under the ergodic assumptions on $f, g$ stated in paragraph C as well as additional regularity conditions ensuring the well-posedness of the equation and the uniform bounds on the solution, discussed in paragraph B. Examples of fast motions with the given ergodic assumptions include Markov processes with polynomial rate of convergence to their stationary measures, stationary processes with an algebraic mixing rate, and fractional dynamics satisfying a large scale contraction assumption. As $\varepsilon \to 0$, we prove that the pathwise mild solution $X^\varepsilon$ to (1.3) converges in (mild) Hölder norm in probability to the solution of the naïvely averaged, autonomous equation

$$d\bar{X}_t = A\bar{X}_t + \bar{f}(\bar{X}_t) \, dt + \bar{g}(\bar{X}_t) \, dB_t, \quad \bar{X}_0 = X_0. \quad (1.4)$$

E. To the best of our knowledge, this article proves the first genuine fractional averaging theorem for SPDEs. The difference between the usual (Markovian) and the fractional effective dynamics lies in the definition of $\bar{g}$. For the former, $\bar{g}$ is a square root of $g(x, y)g^T(x, y)$, and—in case $\bar{g}$ does not vanish—the convergence is in a weak sense. If $g(x, y)$ does not depend on
have defined the following interpolation spaces

\[ \| \cdot \|_h \]

where \( h \in C(\{0, T\}, \mathbb{K}) \), \( \gamma > \frac{1}{2} \), is a deterministic Hölder continuous path, and obtain pathwise bounds.

2.1 Preliminaries

The aim of the section is to fix notation and conventions, they are essentially standard. Throughout this article \( \mathcal{H} \) is a Hilbert space. Let \( A \) be the generator of a strongly continuous semigroup \((S_t)_{t\geq 0}\). Although multiple norms are needed in the sequel, the symbol \( \| \cdot \| \) will always denote the norm \( \mathcal{H} \) and \( \|S_t\| \) the norm of \( S_t \) regarded as a linear map \( \mathcal{H} \to \mathcal{H} \).

Recall that a strongly continuous semigroup \((S_t)_{t\geq 0}\) on \( \mathcal{H} \) is an analytic semigroup of angle \( \theta \in (0, \frac{\pi}{2}] \) if \((S_t)_{t\geq 0}\) extends to

\[ \Delta_\theta := \{ z \in \mathbb{C} \setminus \{0\} : |\arg z| < \theta \} \]

such that (i) \( S_{z_1+z_2} = S_{z_1}S_{z_2} \) for \( z_1, z_2 \in \Delta_\theta \); (ii) the map \( z \mapsto S_z \) is analytic on \( \Delta_\theta \), and (iii) for any \( \theta' < \theta \) and every \( x \in \mathcal{H} \), the strong continuity \( \lim_{t\to 0} S_{z}x = x \) holds along any sequence in \( \Delta_{\theta'} \). If in addition, \( \sup_{z \in \Delta_{\theta'}} \| S_z \| < \infty \) for every \( \theta < \theta' \) we say \( S(t) \) is a bounded analytic semigroup. We assume this condition on \( S_t \), equivalently \( A \) is sectorial.

Without loss of generality, by shifting \( A \) if needed, we assume that there are constant \( C, \nu > 0 \) such that

\[ \| S_t \| \leq C e^{\nu t} \quad \forall t \geq 0. \quad (2.1) \]

By standard theory [39], this allows us to define the bounded injective operator

\[ (-A)^{-\kappa} \triangleq \frac{1}{\Gamma(\kappa)} \int_0^\infty t^{\kappa-1} S_t \, dt, \quad \kappa > 0. \]

Here \( \Gamma(\kappa) \triangleq \int_0^\infty t^{\kappa-1} e^{-t} \, dt \) is Euler’s gamma function. Then \( \mathcal{H} \) is a pre-Hilbert space with respect to the new norm \( \|(-A)^\kappa \cdot \|_{\mathcal{H}} \) whose completion we denote by \( \mathcal{H}_{-\kappa} \). We also set \( A^0 = I \) and \( (-A)^\kappa = ((-A)^{-\kappa})^{-1} \). In the latter case, \((A)^\kappa \) is a closed densely defined operator with domain \( \text{ran } ((-A)^{-\kappa}) \). If \( \kappa_2 > \kappa_1 > 0 \), then \( D((-A)^{-\kappa_2}) \subset D((-A)^{-\kappa_1}) \). In summary, we have defined the following interpolation spaces

\[ \mathcal{H}_\kappa \triangleq \begin{cases} \mathcal{D}((-A)^\kappa) \cap \mathcal{D}_\kappa, \| \cdot \|_{\mathcal{H}_\kappa} = \|(-A)^\kappa \cdot \|_{\mathcal{H}} \} & \kappa > 0, \\ \mathcal{H}_0, \| \cdot \|_{\mathcal{H}} & \kappa = 0, \\ \mathcal{H}_\kappa^\perp \cap \mathcal{D}_\kappa, \| \cdot \|_{\mathcal{H}_\kappa} = \|(-A)^\kappa \cdot \|_{\mathcal{H}} \} & \kappa < 0, \end{cases} \quad (2.2) \]

\( \text{Acknowledgement. The authors are grateful to Khoa Lê for a very helpful conversation.} \)

2 Pathwise Solutions

In this section we develop the mild solution theory for the evolution equation

\[ dx_t = (Ax_t + f(t, x_t)) \, dt + g(t, x_t) \, dh_t, \]

where \( h \in C([0, T], \mathbb{K}) \), \( \gamma > \frac{1}{2} \), is a deterministic Hölder continuous path, and obtain pathwise bounds.
which are themselves Hilbert spaces. In case $A$ is self-adjoint and negative-definite, then for any $\kappa \in \mathbb{R}$, $A^\kappa$ coincides with the operator obtained by functional calculus.

There are the following analytical bounds. If $\kappa_1 \leq \kappa_2$, then $\mathcal{H}_{\kappa_2} \hookrightarrow \mathcal{H}_{\kappa_1}$ densely and

$$\|(id - S_t)x\|_{\mathcal{H}_{\kappa_1}} \lesssim t^{\kappa_2 - \kappa_1} \|x\|_{\mathcal{H}_{\kappa_2}},$$

uniformly in $t \geq 0$ for $x \in \mathcal{H}_{\kappa_2}$, provided that $\kappa_2 - \kappa_1 \in (0, 1)$. On the other hand, for any $\kappa_1 \leq \kappa_2$, we also have the estimate

$$\|S_tx\|_{\mathcal{H}_{\kappa_2}} \lesssim \|x\|_{\mathcal{H}_{\kappa_1}} t^{\kappa_2 - \kappa_1},$$

uniformly in $t > 0$.

### 2.2 Mild Hölder Spaces and Mild Sewing Lemma

Given a function $f : [0, T] \to \mathcal{H}$, we set $\delta f_{s,t} = f_t - f_s$. We also write

$$\hat{\delta} f_{s,t} = f_t - S_{t-s}f_s$$

for the mild increment of $f$. Then for $s \leq t$, it relates to the usual increment process $\delta f_{s,t} = f_t - f_s$ as follows:

$$\delta f_{s,t} = \hat{\delta} f_{s,t} + (id - S_{t-s})f_s$$

For $\gamma \in (0, 1]$ we define the mild Hölder space

$$\bar{\mathcal{C}}^\gamma = \left\{ f : [0, T] \to \mathcal{H} : \|f\|_{\bar{\mathcal{C}}^\gamma} \triangleq \sup_{s \leq t \in [0, T]} \frac{\|\delta f_{s,t}\|}{|t-s|^\gamma} < \infty \right\}.$$  

The notation $\bar{\mathcal{C}}^\gamma([0, T], \mathcal{H})$ is also used if we want to emphasize the underlying space or time interval.

Note that $\bar{\mathcal{C}}^\gamma([0, T], \mathcal{H}) \hookrightarrow \mathcal{C}([0, T], \mathcal{H})$ embeds continuously. In fact, the continuity of $f$ follows from the strong continuity of the semigroup in combination with (2.5). Let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration with the usual assumptions. For $\alpha \in (0, 1]$ and $p \geq 1$ we define

$$\bar{\mathcal{B}}_{\alpha,p} \triangleq \left\{ Z : [0, T] \times \Omega \to \mathcal{H} : Z \text{ is } (\mathcal{F}_t)_{t \geq 0}\text{-adapted, } \|Z\|_{\bar{\mathcal{B}}_{\alpha,p}} \triangleq \sup_{s \neq t} \frac{\|\delta Z_{s,t}\|_{L^p}}{|t-s|^{\alpha}} < \infty \right\}.$$  

As before, we may also use the longer notation $\bar{\mathcal{B}}_{\alpha,p}([0, T], \mathcal{H})$. Elements of $\bar{\mathcal{B}}_{\alpha,p}$ are $L^p$-bounded: one has the bound $\|Z_t\|_{L^p} \lesssim \|Z_s\|_{L^p} + |t-s|^\alpha \|Z\|_{\bar{\mathcal{B}}_{\alpha,p}}$. We furthermore have the following mild version of the Kolmogorov continuity theorem:

**Proposition 2.1** (Mild Kolomogorov criterion). Let $(Z_t)_{t \in [0, T]}$ be an $\mathcal{H}$-valued stochastic process such that for some $\alpha, p > 0$,

$$\|\delta Z_{s,t}\|_{L^p} \lesssim |t-s|^{\alpha} \quad \forall s, t \in [0, T].$$

Then, for any $\gamma < \alpha - \frac{1}{p}$, $Z \in \bar{\mathcal{C}}^\gamma([0, T], \mathcal{H})$ (up to modification) and $\|Z\|_{\bar{\mathcal{C}}^\gamma, L^p} < \infty$. In other words,

$$\bar{\mathcal{B}}_{\alpha,p} \hookrightarrow L^p\left(\Omega, \bar{\mathcal{C}}^{\alpha - \frac{1}{p}}\right).$$
Proof. The argument closely resembles the classical proof of the Kolmogorov criterion. We shall include it for the reader’s convenience.

For each \(n \in \mathbb{N}\) let \(\mathbb{D}_n \triangleq \{ \frac{t}{2^n} : k = 0, \ldots, 2^n \}\) be the dyadic rationals of order \(n\) in \([0,T]\). We also set \(\mathbb{D} \triangleq \bigcup_{n \in \mathbb{N}} \mathbb{D}_n\). Fix \(\gamma < \alpha - \frac{1}{p}\). By Markov’s inequality, we have

\[
\mathbb{P} \left( \max_{k=0, \ldots, 2^n} \left\| \hat{\delta}Z_{\frac{(k-1)T}{2^n}, \frac{kT}{2^n}} \right\| \geq \gamma \right) \leq 2^{p\gamma n} \sum_{k=1}^{2^n} \mathbb{E} \left[ \left\| \hat{\delta}Z_{\frac{(k-1)T}{2^n}, \frac{kT}{2^n}} \right\|^p \right] \leq T^{(n-\gamma)p} 2^{n(p\gamma - p\alpha + 1)}.
\]

The right-hand side is summable in \(n \in \mathbb{N}\), whence

\[
\max_{k=0, \ldots, 2^n} \left\| \hat{\delta}Z_{\frac{(k-1)T}{2^n}, \frac{kT}{2^n}} \right\| \leq C \left( \frac{T}{2^n} \right)^{\gamma} \quad \forall n \in \mathbb{N}\tag{2.9}
\]

for an almost surely finite random variable \(C > 0\) by a Borel-Cantelli argument.

Let \(s, t \in \mathbb{D}\) with \(s < t\). Then we can find an \(n \in \mathbb{N}\) such that \(|t - s| \in [2^{-(n+1)}T, 2^{-n}T]\). Pick increasing sequences \((s_k)_{k \geq n}\) and \((t_k)_{k \geq n}\) converging to \(s\) and \(t\), respectively, such that

1. \(s_k, t_k \in \mathbb{D}_k\) for each \(k \geq n\),
2. \(t_n - s_n = 2^{-n}T\), and
3. \(|s_k + 1 - s_k|, |t_k + 1 - t_k| \in \{0, 2^{-(k+1)}T\} \forall k \geq n\).

By the telescopic expansion of \(Z_t\) using \(Z_t = \lim_{k \to \infty} S_{t-t_k} Z_{t_k}\) and similarly for \(Z_s\), it follows from (2.9) that

\[
\left\| \hat{\delta}Z_{s,t} \right\| \leq \left\| S_{t-t_k} Z_{t_n} - S_{s-s_k} Z_{s_n} \right\| + \sum_{k=n}^{\infty} \left\| S_{s-s_k} \hat{\delta}Z_{s_k,s_k+1} \right\| + \sum_{k=n}^{\infty} \left\| S_{t-t_k} \hat{\delta}Z_{t_k,t_k+1} \right\|
\]

\[
\leq \left\| \hat{\delta}Z_{s_n,t_n} \right\| + \sum_{k=n}^{\infty} \left\| \hat{\delta}Z_{s_k,s_k+1} \right\| + \sum_{k=n}^{\infty} \left\| \hat{\delta}Z_{t_k,t_k+1} \right\|
\]

\[
\leq CT^{\gamma} + 2CT^{\gamma} \sum_{k=n}^{\infty} \frac{1}{2^{(k+1)\gamma}} \leq 3C \left( \frac{T}{2^n} \right)^{\gamma} \leq 3 \cdot 2^\gamma C |t - s|^\gamma.
\]

This gives the asserted Hölder continuity on the set \(\mathbb{D}\) and it is now standard to construct the desired modification of \((Z_t)_{t \in [0,T]}\).

It remains to show that \(\|Z\|_{L^p} < \infty\). As usual, this follows from the Garsia-Rodemich-Rumsey inequality, c.f. Friz-Victoir [15, Theorem A.1]: For each \(\eta > \frac{1}{p}\), we have

\[
\| f(t) - f(s) \|^p \leq |t - s|^{\eta p - 1} \int_s^t \int_s^u \frac{\| f(u) - f(v) \|^p}{|v - u|^{\eta p + 1}} \, du \, dv,
\]

where the prefactor is independent of \(f\). Choosing \(\eta = \gamma + \frac{1}{p}\) and \(f(r) = S_{t-r} Z_r\), we find

\[
\mathbb{E} \left[ \|Z\|^p \right] \leq \mathbb{E} \left[ \sup_{s,t \in [0,T]} \int_s^t \int_s^u \frac{\| S_{t-u} Z_u - S_{s-u} Z_u \|^p}{|v - u|^{\eta p + 2}} \, du \, dv \right]
\]

\[
\leq \int_0^T \int_0^v \mathbb{E} \left[ \| \hat{\delta}Z_{u,v} \|^p \right] \, du \, dv \leq \int_0^T \int_0^v |v - u|^{(\alpha - \gamma)p - 2} \, du \, dv < \infty,
\]

as required. \(\square\)
Fix a deterministic function \( h \in C^\gamma([0,T],\mathcal{K}) \) where \( \gamma > \frac{1}{2} \). Let \( f : [0,T] \times \mathcal{H} \to \mathcal{H} \) and \( g : [0,T] \times \mathcal{H} \to L(\mathcal{K},\mathcal{H}) \). We want to show that under suitable regularity conditions on \( f \) and \( g \), the evolution equation
\[
 dx_t = (Ax_t + f(t,x_t)) \, dt + g(t,x_t) \, dh_t, \quad t \in [0,T], \tag{2.10}
\]
has a unique mild solution. Since any trace-class fBm belongs to \( C^{H-} \) by Lemma 4.3 below, the deterministic results in the sequel carry over to almost sure statements when \( h \) is replaced by \( B \).

We understand the integral \( \int_0^t S_{t-s}g(s,x_s) \, dh_s \), appearing in the formulation of the mild solution to (2.10) as a Young integral. In particular, we shall make frequent use of the next result which was proven in the classical work of Young [43], see also Lyons [36, 35] for a contemporary proof.

**Proposition 2.2 (Young integral).** Let \( f \in C^\alpha([0,T],L(\mathcal{K},\mathcal{H})) \) and \( h \in C^\gamma([0,T],\mathcal{K}) \) with \( \alpha + \gamma > 1 \). Then, for each \( s, t \in [0,T] \) the integral
\[
 \int_s^t f(r) \, dh_r = \lim_{|P| \to 0} \sum_{[u,v] \in P} f(u)(h_v - h_u)
\]
exists as a limit in \( \mathcal{H} \) along an arbitrary sequence of partitions of \([s,t]\) with mesh tending to zero. Moreover, we have the estimate
\[
 \left\| \int_s^t f(r) \, dh_r - f(s)(h_t - h_s) \right\|_{\mathcal{H}} \lesssim \|f\|_{C^\alpha}\|h\|_{C^\gamma}|t-s|^{\alpha+\gamma}, \tag{2.11}
\]
uniform in \( s, t \in [0,T] \).

While the estimate (2.11) has been successfully used to prove pathwise existence results for (random) Young differential equations in finite dimensions, the following computation shows that it is ill-suited for our investigation of mild solutions:
\[
 \left\| \int_s^t S_{t-r}f(r) \, dh_r \right\|_{\mathcal{H}} \lesssim \|h\|_{C^\gamma} \left( \sup_{u \in [s,t]} \|S_{t-u}f(u)\| + \sup_{u < v \in [s,t]} \|S_{t-u}f(v) - S_{t-u}f(u)\| |v-u|\right)^\gamma \lesssim \|h\|_{C^\gamma} (\|f\|_\infty + \|\hat{f}\|_{C^\alpha})|t-s|^{\gamma},
\]
where we used \( \sup_{t \in [T]} \|S_t\| < \infty \). Now to show that \( \|\hat{f}\|_{C^\alpha} < \infty \), we write
\[
 \|\hat{f}\|_{C^\alpha} \leq \|f(t) - f(s)\|_{\mathcal{H}} + \|((id - S_{t-s})f)(s)\|_{\mathcal{H}}.
\]
While the first term can be handled by imposing appropriate assumptions on \( f \), the second term in general only admits an \( O(|t-s|) \)-bound if we go up the ladder of interpolation spaces; viz. \( \|((id - S_{t-s})f)(s)\|_{\mathcal{H}} \lesssim \|f\|_{H^\alpha}|t-s|^{\alpha} \), see (2.3). This is of course notoriously bad for any kind of fixed point argument aimed at showing well-posedness for (2.10). Hence, we resort to a technique introduced by Gubinelli and Tindel in [22]. We give a presentation of their results which is however more in the style of the monograph of Friz and Hairer [14].

Let \( \Delta^2 = \{(s,t) \in [0,T]^2 : s < t\} \) and \( \Delta^3 = \{(r,s,t) \in [0,T]^3 : r \leq s \leq t\} \) be the two- (respectively three-) dimensional simplices. For \( \Xi : \Delta^2 \to \mathcal{H} \) and \( (r,s,t) \in \Delta^3 \) let
\[
 \delta \Xi_{r,s,t} = \Xi_{s,t} - \Xi_{s,t} - S_{t-s} \Xi_{r,s}.
\]
For \( \gamma > 0 \) and \( \mathcal{H} \) a Hilbert space, we set
\[
 \hat{\mathcal{C}}^\gamma_\gamma(\Delta^2, \mathcal{H}) = \left\{ \Xi : \Delta^2 \to \mathcal{H} : \|\Xi\|_{\hat{\mathcal{C}}^\gamma} = \sup_{0 \leq r < s < t \leq T} \|\delta \Xi_{r,s,t}\|_{\mathcal{H}} < \infty \right\}.
\]
In the following, we take the Hilbert space to be \( \mathcal{H}_\kappa \), an interpolation space defined by \( \mathcal{A} \). The following version of Gubinelli’s sewing lemma \([21, 13]\) can be found in Gubinelli \([22, \text{Theorem 3.5}]\), see also Gerasimovics and Hairer \([20]\).

**Proposition 2.3** (Mild Sewing lemma). Let \( \kappa \in \mathbb{R} \) and \( 0 < \alpha \leq 1 < \mu \). Let \( \Xi : \Delta^3 \to \mathcal{H}_\kappa \) satisfy
\[
\sup_{s < t} \frac{\|\Xi_{s,t}\|_{\mathcal{H}_\kappa}}{|t-s|^{\alpha}} < \infty
\]
and suppose that there is a \( \Lambda : \Delta^3 \to \mathcal{H}_\kappa \) such that \( \hat{\delta}\Xi_{r,s,t} = S_{t-s}\Lambda_{r,s,t} \) for all \( (r, s, t) \in \Delta^3 \). Furthermore assume that there is a \( C > 0 \) such that
\[
\|\Lambda_{r,s,t}\|_{\mathcal{H}_\kappa} \leq C|t-r||t-s|^{\mu-1} \quad \forall (r, s, t) \in \Delta^3.
\]
Then the limit
\[
\mathcal{T}\Xi_{s,t} \triangleq \lim_{|P| \to 0} \sum_{[u,v] \in P} S_{t-u}\Xi_{u,v}
\]
exists along any sequence of partitions \( P \) of \([s, t]\) with mesh tending to zero. Moreover, for each \( \varpi \in (0, \mu) \), we have the bound
\[
\|\mathcal{T}\Xi_{s,t} - \Xi_{s,t}\|_{\mathcal{H}_{\kappa+\varpi}} \lesssim C|t-s|^{\mu-\varpi}
\]
for a hidden prefactor depending only on \( \mu \) and \( \varpi \).

An immediate consequence of Proposition 2.3 is the following corollary which allows us to trade time-regularity for interpolation space-regularity:

**Corollary 2.4** (Mild Young bound). Let \( f \in \dot{C}^\alpha([0,T],L(\mathcal{K},\mathcal{H}_\kappa)) \) and \( h \in \dot{C}^\gamma([0,T],\mathcal{K}) \) with \( \alpha + \gamma > 1 \). Then, for each \( \varpi \in (0, \alpha + \gamma) \), the Young integral of Proposition 2.2 satisfies the bound
\[
\left\| \int_0^u S_{u-r} f(r) \, dh_r - S_{u-u} f(u) \, h_{u,v} \right\|_{\mathcal{H}_{\kappa+\varpi}} \lesssim \|h\|_{\dot{C}^\varpi} \|f\|_{\dot{C}^\alpha} |t-s|^{(\alpha+\gamma-\varpi)\wedge \gamma}. \tag{2.13}
\]

*Proof.* The Young integral is given by the limit of the Riemann sum over \( \Xi_{u,v} \triangleq S_{v-u} f(u) h_{u,v} \). Then it is readily checked that \( \hat{\delta}\Xi_{r,u,v} = S_{v-u} \Lambda_{r,u,v} \) for
\[
\Lambda_{r,u,v} = \left( S_{u-r} f(r) - f(u) \right) h_{u,v} = -\hat{\delta} f_{r,u} h_{u,v}.
\]
Moreover, we see that
\[
\|\Lambda_{r,u,v}\|_{\mathcal{H}_\kappa} \leq \|f\|_{\dot{C}^\alpha} \|h\|_{\dot{C}^\gamma} |v-r| |v-u|^{|\alpha+\gamma-1|
\]
for all \( (r, u, v) \in \Delta^3 \). Therefore, the required bound follows from Proposition 2.3. \( \square \)

### 2.3 Pathwise Well-Posedness and a Deterministic Stability Lemma

We now turn to a well-posedness result for \((2.10)\). As a preparation we recall the following:

**Lemma 2.5.** Let \( f : [0,T] \times \mathcal{H} \to \mathcal{H} \) be such that the functions \( \{f(t, \cdot) : t \in [0,T]\} \) and their first order spatial derivatives uniformly Lipschitz continuous. In addition, suppose that \( t \mapsto D_x f(t, x) \) is uniformly \( \alpha \)-Hölder continuous. Then we have
\[
\|f(t, x_1) - f(t, x_2) - f(s, x_3) + f(s, x_4)\| \lesssim \|x_1 - x_2 - x_3 + x_4\| + \|x_1 - x_2\| \left( |t-s|^\alpha + \|x_1 - x_3\| + \|x_2 - x_4\| \right),
\]
for all \( x_1, \ldots, x_4 \in \mathcal{H} \) and \( s, t \in [0,T] \).
Proof. By assumption there is a constant $L > 0$ such that, for all $x, y \in \mathcal{H}$ and all $t \in [0, T]$,

$$
\|f(t, x) - f(t, y)\| \leq L\|x - y\|,
\|D_x f(t, x) - D_x f(t, y)\|_{L(H)} \leq L\|x - y\|,
\|D_x f(s, x) - D_x f(t, x)\|_{L(H)} \leq L|t - s|^{\alpha}.
$$

From the identity

$$
f(t, x_1) - f(t, x_2) - f(s, x_3) + f(s, x_4)
= \int_0^1 D_x f(t, \theta x_1 + (1 - \theta)x_2)(x_1 - x_2) d\theta - \int_0^1 D_x f(s, \theta x_3 + (1 - \theta)x_4)(x_3 - x_4) d\theta
= \int_0^1 D_x f(s, \theta x_3 + (1 - \theta)x_4)(x_1 - x_2 - x_3 + x_4) d\theta
+ \int_0^1 \left[D_x f(t, \theta x_1 + (1 - \theta)x_2) - D_x f(s, \theta x_3 + (1 - \theta)x_4)\right](x_1 - x_2) d\theta,
$$

the claimed estimate follows at once. \hfill \Box

The announced existence and uniqueness result reads as follows:

**Proposition 2.6.** Let $\alpha, \gamma$ be two numbers with $\gamma \in (\frac{1}{2}, 1]$ and $\alpha + \gamma > 1$. Let $\mathfrak{h} : [0, T] \to \mathcal{K}$, $f : [0, T] \times \mathcal{H} \to \mathcal{H}$, and $g : [0, T] \times \mathcal{H} \to L(\mathcal{K}, \mathcal{H})$ satisfy the following conditions:

(i) $\mathfrak{h} \in C^\gamma([0, T], \mathcal{K})$.

(ii) For any $s, t \in [0, T]$ and $x, y \in \mathcal{H}$,

$$
\|f(t, x) - f(s, y)\| \leq C(t - s)^{\alpha} + \|x - y\|. \quad (2.14)
$$

(iii) The functions $g(t, \cdot) : \mathcal{H} \to L(\mathcal{K}, \mathcal{H})$ are Lipschitz continuous, uniformly in $t \in [0, T]$. Moreover,

$$
\sup_{s \in [0, T]} \|g(s, x) - g(s, y)\|_{L(\mathcal{K}, \mathcal{H}^{\gamma - \alpha})} \lesssim \|x - y\|_{\mathcal{H}^{\gamma - \alpha}},
\sup_{x \in \mathcal{H}} \|g(t, x) - g(s, x)\|_{L(\mathcal{K}, \mathcal{H}^{\gamma - \alpha})} \lesssim |t - s|^{\alpha},
\sup_{t \in [0, T]} \|D_x g(t, x) - D_x g(t, y)\|_{L(\mathcal{K}, L(\mathcal{H}^{\gamma - \alpha}))} \lesssim \|x - y\|_{\mathcal{H}^{\gamma - \alpha}},
\sup_{x \in \mathcal{H}} \|D_x g(s, x) - D_x g(t, x)\|_{L(\mathcal{K}, L(\mathcal{H}^{\gamma - \alpha}))} \lesssim |t - s|^{\alpha}.
$$

Then, for any $x_0 \in \mathcal{H}$, the Young equation (2.10) has a unique global mild solution in $C^{\gamma - \alpha}((0, T], \mathcal{H})$. Furthermore, for each $\tilde{\gamma} < \gamma$, there are constants $C > 0$ and $N \in \mathbb{N}$ independent of $x_0$ and $\mathfrak{h}$ such that

$$
\|x\|_{C^{\tilde{\gamma}}([0, T], \mathcal{H})} \leq C(1 + \|x_0\|) \left(1 + \|\mathfrak{h}\|_{C^\gamma([0, T], \mathcal{H})}\right)^N \quad (2.15).
$$

Proof. Fix $\tilde{\gamma} < \gamma$. Without loss of generality, we may assume that $\alpha < \gamma$.

(i) We first establish the *a priori* bound (2.15) on $\|x\|_{C^{\tilde{\gamma}}}$, showing that the solution (if it exists) cannot blow up in finite time.

(ii) We show that the application

$$
(Ay)_t \triangleq S_t y_0 + \int_0^t S_{t-s} f(s, y_s) \, ds + \int_0^t S_{t-s} g(s, y_s) \, d\mathfrak{h}_s, \quad t \in [0, T],
$$

where

...
is contracting on the complete metric space
\[ V \triangleq \left\{ y : [0, \rho] \to H : y_0 = x_0, \| y \|_{C^5([0, \rho], H)} \leq 1 \right\}, \]
provided we choose \( \rho \in (0, T) \) small enough. Hence, the Banach fixed point theorem implies that the equation (2.10) has a unique mild solution on the interval \([0, \rho]\). Owing to (i) we can iterate this procedure to construct the unique mild solution on all of \([0, T]\).

**Step (i):** Let \( \rho \in (0, T) \) and \( x_0 \in H \). From
\[ x_t = S_t x_0 + \int_0^t S_{t-s} f(s, x_s) \, ds + \int_0^t S_{t-s} g(s, x_s) \, dh_s = S_t x_0 + \delta x_{0,t} \]
and by the uniform boundedness of the semigroup, we have
\[ \| x_t \| \leq \| S_t x_0 \| + \| \delta x_{0,t} \| \leq \| x_0 \| + t \gamma \| x \|_{\hat{C}^\gamma}, \quad \forall t \in [0, \rho]. \] (2.16)
To obtain the required a priori bound we begin with \( x = \mathcal{A} x \) and observe that
\[ \delta(\mathcal{A} x)_{s,t} = \int_s^t S_{t-r} f(r, x_r) \, dr + \int_s^t S_{t-r} g(r, x_r) \, dh_r, \]
Since \( \| f(r, x) \|_H \lesssim 1 + \| x \|_H \) by the Lipschitz continuity of \( f \), we have the trivial bound:
\[ \left\| \int_s^t S_{t-r} f(r, x_r) \, dr \right\|_H \lesssim (1 + \| x \|_\infty) |t - s|. \] (2.17)
Next, we use the Young bound from Corollary 2.4 with \( \varpi = \alpha \) to find
\[ \left\| \int_s^t S_{t-r} g(r, x_r) \, dh_r \right\|_H \lesssim \| b \|_{\hat{C}^\gamma} \left( \| g(\cdot, x_r) \|_\infty + \| g(\cdot, x_r) \|_{\hat{C}^\alpha(L(K, H_\alpha))} \right) |t - s|^\gamma. \] (2.18)
Note that \( \| g(\cdot, x_r) \|_\infty \lesssim 1 + \| x \|_\infty \) by uniform Lipschitz continuity. We split the second term in the bracket on the right hand side of (2.18):
\[ \| \delta g(\cdot, x_r)_{s,t} \|_{L(K, H_\alpha)} \lesssim \| g(t, x_t) - g(s, S_{t-s} x_s) \|_{L(K, H_\alpha)} + \| g(s, S_{t-s} x_s) - S_{t-s} g(s, x_s) \|_{L(K, H_\alpha)}. \]
Firstly, by the uniform Hölder continuity of \( g \) in time and its Lipschitz continuity on \( H_\alpha \),
\[ \| g(t, x_t) - g(s, S_{t-s} x_s) \|_{L(K, H_\alpha)} \lesssim \left( |t - s|^{\alpha} + \| \delta x_{s,t} \|_{H_\alpha} \right) \lesssim \left( 1 + \| x \|_{\hat{C}^\gamma([0, \rho], H)} \right) |t - s|^{\alpha}, \]
where in the last step we used that \( H_\alpha \hookrightarrow H \) and that \( \alpha < \gamma \). Moreover by (2.3) and Lipschitz continuity of \( g(s, \cdot) : H_\alpha \to L(K, H_\alpha) \), we have
\[ \| g(s, S_{t-s} x_s) - S_{t-s} g(s, x_s) \|_{L(K, H_\alpha)} \lesssim \left( |t - s|^{\alpha} + \| \delta x_{s,t} \|_{H_\alpha} \right) \lesssim \left( 1 + \| x \|_{\hat{C}^\gamma([0, \rho], H)} \right) |t - s|^{\alpha}. \]
In summary, since \( \| x \|_\infty \lesssim \| x \|_0 + \| x \|_{\hat{C}^\gamma} \) by (2.16),
\[ \| g(\cdot, x_r) \|_{\hat{C}^\alpha(L(K, H_\alpha))} \lesssim \left( 1 + \| x \|_0 + \| x \|_{\hat{C}^\gamma} \right). \]
Inserting this back into (2.17)–(2.18), we have shown that
\[
\left\| \int_s^t S_{t-r} g(r, x_r) \, dh_r \right\|_{\mathcal{H}} \lesssim \|b\|_{c^\gamma} \left( 1 + \|x_0\| + \|x\|_{\hat{c}^\gamma} \right) |t - s|^{\gamma}.
\]
and
\[
\|x\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})} = \|Ax\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})} \lesssim \left( 1 + \|b\|_{c^\gamma} \right) \left( 1 + \|x_0\| + \|x\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})} \right) \rho^{\gamma - \gamma}, \tag{2.19}
\]
whence, upon choosing \( \rho > 0 \) small enough,
\[
\|x\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})} \leq C \left( 1 + \|x_0\| \right) \left( 1 + \|b\|_{c^\gamma([0, \rho], \mathcal{H})} \right).
\]
This bound can be iterated to obtain
\[
\|x\|_{\hat{c}^\gamma([t, t + \rho]; \mathcal{H})} \leq C \left[ \frac{\rho}{\rho - 1} \right] \left( 1 + \|x_0\| \right) \left( 1 + \|b\|_{c^\gamma([0, \rho], \mathcal{H})} \right) \left[ \frac{\rho}{\rho - 1} \right]^{-1},
\]
where \( [a] \) denotes the integer part of a number \( a \).

Now let \( \gamma \in (0, 1) \) and \( \rho \in (0, 1) \), we show that
\[
\|x\|_{c^\gamma([0, T], \mathcal{H})} \leq \frac{1}{\rho^{1 - \gamma}} \sup_{t \in [0, T - \rho]} \|x\|_{\hat{c}^\gamma([\rho, T + \rho]; \mathcal{H})}. \tag{2.20}
\]
Fix \( 0 \leq s < t \leq T \). Without loss of generality we can assume that \( |t - s| > \rho \). Let \( P \) be a partition of \([s, t] \) with \( |t - s| \rho^{-1} + 1 \) elements and mesh \( |P| \leq \rho \). For concreteness choose for instance \( \{(s + n\rho) \wedge t : n = 0, \ldots, |t - s| \rho^{-1}\} \). Then
\[
\|\delta x_{s,t}\|_{\mathcal{H}} \leq \sum_{[u,v] \in P} \|S_{t-v} \delta x_{u,v}\|_{\mathcal{H}} \lesssim \sum_{[u,v] \in P} \|\delta x_{u,v}\|_{\mathcal{H}} \leq \rho^{\gamma} \left( \frac{|t - s|}{\rho} + 1 \right) \sup_{t \in [0, T - \rho]} \|x\|_{\hat{c}^\gamma([\rho, T + \rho]; \mathcal{H})} \leq 2 \rho^{\gamma} |t - s| \sup_{t \in [0, T - \rho]} \|x\|_{\hat{c}^\gamma([\rho, T + \rho]; \mathcal{H})}.
\]
This finally implies the apriori estimate
\[
\|x\|_{\hat{c}^\gamma([0, T], \mathcal{H})} \lesssim C \left[ \frac{\rho}{\rho - 1} \right] \left( 1 + \|x_0\|_{\mathcal{H}} \right) \left( 1 + \|b\|_{c^\gamma([0, T], \mathcal{H})} \right) \left[ \frac{\rho}{\rho - 1} \right]^{-1}.
\]

**Step (ii):** The invariance of the set \( V \) under the application \( A, \mathcal{A}(V) \subset V \), has already been shown in step (i). In fact, it follows immediately from (2.19) that \( \|Ay\|_{\hat{c}^\gamma \mathcal{H}} \leq 1 \) for \( \rho > 0 \) sufficiently small. We are thus left to show that there is a \( C < 1 \) and \( \rho > 0 \) such that
\[
\|Ay - A\tilde{y}\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})} \leq C \|y - \tilde{y}\|_{\hat{c}^\gamma([0, \rho]; \mathcal{H})}, \quad \forall y, \tilde{y} \in V. \tag{2.21}
\]
By the Lipschitz continuity of \( f \),
\[
\left\| \int_s^t S_{t-r} \left( f(r, y_r) - f(r, \tilde{y}_r) \right) \, dr \right\| \lesssim \|y - \tilde{y}\|_{\infty} |t - s|. \tag{2.22}
\]
Similarly to (2.18), we also find
\[ \left\| \int_s^t S_{t-r} \left( g(r, y_r) - g(r, \bar{y}_r) \right) \, dh_r \right\| \leq \|h\|C^\gamma \left( \left\| g(\cdot, y) - g(\cdot, \bar{y}) \right\|_\infty + \left\| g(\cdot, \cdot) - \hat{g}\left( C^{\alpha(L(\mathcal{K}, \mathcal{H})-\gamma)} \right) \right\| \right) |t-s|^\gamma. \] (2.23)

By assumption, \( \left\| g(\cdot, y) - g(\cdot, \bar{y}) \right\|_\infty \lesssim \|y - \bar{y}\|_\infty \). On the other hand, we can write
\[ \hat{\delta}(g(\cdot, y_r) - g(\cdot, \bar{y}_r))_{s,t} = \left( g(t, y_t) - g(t, \bar{y}_t) - g(s, S_{t-s} y_s) + g(s, S_{t-s} \bar{y}_s) \right) + \left( g(s, S_{t-s} y_s) - g(s, S_{t-s} \bar{y}_s) - g(s, y_s) + g(s, \bar{y}_s) \right) + \left( g(s, y_s) - g(s, \bar{y}_s) - S_{t-s} g(s, y_s) + S_{t-s} g(s, \bar{y}_s) \right) = (I)_{s,t} + (II)_{s,t} + (III)_{s,t}. \]

Next, we apply Lemma 2.5 to the first two terms. Since \( y, \bar{y} \in V \), this gives
\[ \| (I)_{s,t} \|_{L(K, H_{-\alpha})} \lesssim \| \hat{\delta}(y - \bar{y})_{s,t} \|_{H_{-\alpha}} + \| y_t - \bar{y}_t \|_{H_{-\alpha}} \left( |t-s|^\alpha + \| \hat{\delta} y_{s,t} \|_{H_{-\alpha}} + \| \hat{\delta} \bar{y}_{s,t} \|_{H_{-\alpha}} \right) \lesssim \left( \| y - \bar{y} \|_{C^\gamma} + \| y_t - \bar{y}_t \|_{H_{-\alpha}} \right) |t-s|^\alpha, \]
so that \( \| (I)_{s,t} \|_{C^\alpha(L(K, H_{-\alpha}))} \lesssim (1 + \| x_0 \|) \| y - \bar{y} \|_{C^\gamma(H)} |t-s|^\alpha \). Similarly, we find
\[ \| (II)_{s,t} \|_{L(K, H_{-\alpha})} \lesssim \| (id - S_{t-s})(y_s - \bar{y}_s) \|_{H_{-\alpha}} + \| S_{t-s}(y_s - \bar{y}_s) \|_{H_{-\alpha}} \left( \| (id - S_{t-s}) y_s \|_{H_{-\alpha}} + \| (id - S_{t-s}) \bar{y}_s \|_{H_{-\alpha}} \right) \lesssim \left( \| y - \bar{y} \|_{H_{-\alpha}} + \| y_t - \bar{y}_t \|_{H_{-\alpha}} \right) |t-s|^\alpha \lesssim (1 + \| x_0 \|) \| y - \bar{y} \|_{C^\gamma(H)} |t-s|^\alpha, \]
where we made use of (2.3). Finally, we observe that
\[ \| (III)_{s,t} \|_{L(K, H_{-\alpha})} = \| (id - S_{t-s}) \left( g(s, y_s) - g(s, \bar{y}_s) \right) \|_{L(K, H_{-\alpha})} \lesssim \left( \| g(s, y_s) - g(s, \bar{y}_s) \|_{L(K, H)} \right) |t-s|^\alpha \lesssim (1 + \| x_0 \|) \| y - \bar{y} \|_{C^\gamma} |t-s|^\alpha. \]
Combining these bounds, we see that
\[ \| g(\cdot, y_r) - g(\cdot, \bar{y}_r) \|_{C^\alpha([0, t], H_{-\alpha})} \lesssim (1 + \| x_0 \|) \| y - \bar{y} \|_{C^\gamma}, \]
whence we conclude
\[ \| A y - A \bar{y} \|_{C^\gamma} \lesssim \rho^{\gamma-\gamma} \| y - \bar{y} \|_{C^\gamma} \] (2.24)
from (2.22)–(2.23). This establishes (2.21) upon potentially further decreasing \( \rho > 0 \) and the proof is complete.

Finally, the following Grönwall-type stability result will be used in the proof of the averaging principle in §6:

**Lemma 2.7** (Mild residue lemma). Let \( \gamma \in \left( \frac{1}{\alpha}, 1 \right], \alpha \in (1 - \gamma, \gamma) \). Suppose that \( f, g, \) and \( h \) satisfy the conditions of Proposition 2.6. Then, for any \( y, \bar{y} \in \hat{C}^\alpha([0, T], \mathcal{H}) \), the equations
\[
x_t = y_t + \int_0^t S_{t-s}f(s, x_s) \, ds + \int_0^t S_{t-s}g(s, x_s) \, dh_s, \\
\bar{x}_t = \bar{y}_t + \int_0^t S_{t-s}f(s, \bar{x}_s) \, ds + \int_0^t S_{t-s}g(s, \bar{x}_s) \, dh_s
\]
have a unique solution in $\mathcal{C}^\alpha([0, T], \mathcal{H})$. If in addition $y_0 = \bar{y}_0$, we have the stability estimate
\[
\|x - \bar{x}\|_{\mathcal{C}^\alpha([0, T], \mathcal{H})} \leq C\|y - \bar{y}\|_{\mathcal{C}^\alpha([0, T], \mathcal{H})}
\] (2.25)
where the constant $C$ depends only on $f$, $g$, $\|b\|_{C^\gamma}$, $\|y_0\|$, $\|y\|_{\mathcal{C}^\alpha}$, $\|\bar{y}\|_{\mathcal{C}^\alpha}$, and $T$.

**Proof.** The well-posedness of the equations for $x$ and $\bar{x}$ follows as in Proposition 2.6. On a small interval $[0, \rho]$ this is the consequence of the contraction property of $A$. By the a priori bound (2.15) of solutions, the required estimate holds.

**Remark 2.8.** A more refined analysis shows that in (2.25) one could choose
\[
C = D \exp \left( D\|b\|_{C^\gamma}^\frac{1}{\alpha} + D\|y\|_{\mathcal{C}^\alpha}^\frac{1}{\alpha} + D\|\bar{y}\|_{\mathcal{C}^\alpha}^\frac{1}{\alpha} \right)
\]
for a constant $D > 0$ depending only on $f$, $g$, $\|y_0\|$, and the terminal time $T > 0$; see [24, Lemma 2.2] for details. To prove convergence in probability, it is however enough to have $C < \infty$ with probability 1, see § 6 below.

### 3 Mild Stochastic Sewing Lemma

Our main endeavor in the next section is to make sense of the integral
\[
\int_0^T g(t, X_t) \, dB_t
\]
and to derive strong $L^p$-estimates on its Hölder norm. This is based on a variant of Lê’s stochastic sewing lemma [31] adapted to the mild calculus used in this article, which we present in the sequel.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space. Fix a terminal time $T > 0$ and let $S^p(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space, denote the set of adapted, continuous two-parameter stochastic processes on the simplex with finite $p$th moments; in symbols:
\[
S^p(\mathcal{H}) \triangleq \{ \Sigma : \Delta^2 \times \Omega \rightarrow \mathcal{H} : (s, t) \mapsto \Sigma_{s,t} \in L^p(\Omega, \mathcal{F}_t; \mathcal{H}) \text{ is continuous} \}.
\]
Given $\eta, \tilde{\eta} > 0$, we define the spaces
\[
H^p_{\tilde{\eta}}(\mathcal{H}) \triangleq \left\{ \Xi \in S^p(\mathcal{H}) : \|\Xi\|_{H^p_{\tilde{\eta}}} \triangleq \sup_{0 \leq s < t < T} \frac{\|\Xi_{s,t}\|_{L^p}}{|s - t|^\eta} < \infty \right\},
\]
\[
\tilde{H}^p_{\tilde{\eta}}(\mathcal{H}) \triangleq \left\{ \Xi \in S^p(\mathcal{H}) : \|\Xi\|_{\tilde{H}^p_{\tilde{\eta}}} \triangleq \sup_{0 \leq s < u < t < T} \frac{\|\Xi_{s,u} \delta_{s,u,t}\|_{L^p}}{|s - t|^\eta} < \infty \right\},
\]
where we recall that
\[
\tilde{\delta}_{s,u,t} = \Xi_{s,t} - \Xi_{u,t} - S_{t-u} \Xi_{s,u}.
\]
In the sequel we take the Hilbert space to be $\mathcal{H}_N$ and occasionally abbreviate $L^p(\mathcal{H}_N)$ to $L^p$. As in Lê’s work, our proof of the mild stochastic sewing lemma (Proposition 3.2 below) is based on the following inequality: Let $(Z_i)_{i \in \mathbb{N}}$ be an $\mathcal{H}_N$-valued, discrete-time stochastic process adapted to a filtration $(\mathcal{F}_i)_{i \in \mathbb{N}}$, and let $d_i = Z_i - \mathbb{E}[Z_i \mid \mathcal{F}_{i-1}]$ be the sequence of martingale differences. Then
\[
\left\| \sum_{i=1}^n Z_i \right\|_{L^p} \leq \left\| \sum_{i=1}^n \mathbb{E}[Z_i \mid \mathcal{F}_{i-1}] \right\|_{L^p} + \left\| \sum_{i=1}^n d_i \right\|_{L^p} \tag{3.1}
\]

\[
\lesssim \sum_{i=1}^n \left\| \mathbb{E}[Z_i \mid \mathcal{F}_{i-1}] \right\|_{L^p} + \left\| \sum_{i=1}^n d_i \right\|_{L^{p/2}} \frac{1}{2} \left( \sum_{i=1}^n \left\| Z_i \right\|^2_{L^p} \right)^{\frac{1}{2}}.
\]
In the second line we used Minkowski’s and Burkholder’s inequalities. The last step follows from another application of Minkowski’s inequality. Burkholder’s inequality on Hilbert spaces can be proven by the classical square function approach of Burkholder [6]. In particular, the prefactors, which may depend on \( p \geq 1 \), are independent of \( n \in \mathbb{N} \).

Before stating the mild version of the sewing lemma, we first make some basic error estimates.

**Lemma 3.1.** Let \( P = \{ t_0, \ldots, t_n \} \) be a partition of \([s, t] \subset [0, T] \) with \( t_0 = s \) and \( t_n = t \). Let \( \kappa \in \mathbb{R} \). Let \( p \geq 2 \), \( \eta > \frac{1}{2} \), and \( \bar{\eta} > 1 \). Suppose that \( \Xi \in H^p_\eta(\mathcal{H}_\kappa) \cap \dot{H}^p_{\bar{\eta}}(\mathcal{H}_\kappa) \). We have that

\[
\left\| \mathbb{E} \left[ \sum_{i=1}^{n-1} S_{t_{n-i+1}} \Xi_{t_i, t_{i+1}} - \Xi_{t_0, t_n} \right] \right\|_{L^p(\mathcal{H}_\kappa)} \lesssim |t-s|^{\bar{\eta}}. \tag{3.2}
\]

Furthermore,

\[
\left\| \sum_{i=1}^{n-1} S_{t_{n-i+1}} \Xi_{t_i, t_{i+1}} - \Xi_{t_0, t_n} \right\|_{L^p(\mathcal{H}_\kappa)} \lesssim |t-s|^{\eta}. \tag{3.3}
\]

**Proof.** Let \( \Delta_k = \frac{i \Delta k}{2M} \) and consider sequences of points \( r^k_i = u^k_i \leq v^k_i \leq w^k_i \) in the interval \([t_0 + i \Delta k, t_0 + (i + 1) \Delta k]\), \( i = 0, \ldots, 2^k - 1 \). Let

\[ R^k_i = -S_{t_{n-i+1}}(w^k_i - v^k_i) \delta \Xi_{t_i, t_{i+1}} + \delta \Xi_{t_i, t_{i+1}} + w^k_i - v^k_i). \]

Applying the assumption \( \Xi \in H^p_\eta(\mathcal{H}_\kappa) \cap \dot{H}^p_{\bar{\eta}}(\mathcal{H}_\kappa) \) to each term in the definition of \( R^k_i \), we immediately have the trivial estimate

\[ \left\| \mathbb{E} \left[ R^k_i \mid \mathcal{F}_{r^k_i} \right] \right\|_{L^p} \lesssim (\Delta_k)^{\bar{\eta}}. \]

Furthermore,

\[ \left\| R^k_i - \mathbb{E} \left[ R^k_i \mid \mathcal{F}_{r^k_i} \right] \right\|_{L^p} \lesssim \| R^k_i \|_{L^p} \lesssim (\Delta_k)^{\eta}. \]

We claim that we can calibrate \( r^k_i \leq u^k_i \leq v^k_i \leq w^k_i \) in such a way that

\[ \sum_{i=1}^{n-1} S_{t_{n-i+1}} \Xi_{t_i, t_{i+1}} - \Xi_{t_0, t_n} = \sum_{k=0}^{2^k-1} \sum_{i=0}^{2^k-1} R^k_i, \tag{3.4} \]

which then immediately leads to (3.2):

\[ \left\| \mathbb{E} \left[ \sum_{i=1}^{n-1} S_{t_{n-i+1}} \Xi_{t_i, t_{i+1}} - \Xi_{t_0, t_n} \right] \right\|_{L^p} \lesssim \sum_{k=0}^{\infty} \left\| \mathbb{E} \left[ R^k_i \mid \mathcal{F}_{r^k_i} \right] \right\|_{L^p} \lesssim |t-s|^{\eta}. \]

Similarly, the estimate (3.1) yields

\[ \left\| \sum_{i=0}^{2^k-1} R^k_i \right\|_{L^p} \lesssim \sum_{i=0}^{2^k-1} \left\| \mathbb{E} \left[ R^k_i \mid \mathcal{F}_{r^k_i} \right] \right\|_{L^p} + \left( \sum_{i=0}^{2^k-1} \left\| R^k_i - \mathbb{E} \left[ R^k_i \mid \mathcal{F}_{r^k_i} \right] \right\|_{L^p} \right)^{\frac{1}{2}} \]

\[ \lesssim 2^{k-2\bar{\eta}} |t-s|^{\eta} \leq 2^{k-2\bar{\eta}} |t-s|^{\eta} \lesssim \frac{|t-s|^{\bar{\eta}}}{2^{k([\eta](\eta \wedge (\eta - \frac{\bar{\eta}}{2}))}\]}

and (3.3) follows at once.

To conclude the proof, it remains to show (3.4) holds. Let \( Q = \{ s_0, \ldots, s_m \} \) be also a partition of \([s, t] \). We set

\[ \mathcal{M}(Q) = \begin{cases} \sum_{i=0}^{m-1} S_{t_{n-s_{i+1}}} \Xi_{s_i, s_{i+1}} - S_{t_{n-s_m}} \Xi_{s_0, s_m}, & \text{if } \#Q > 1 \\ 0, & \text{otherwise.} \end{cases} \tag{3.5} \]
We also need the following dyadic sub-partitions of $P$: Let $k \in \mathbb{N}_0$ and $P_k^i$ those points from $P$ falling into the $i$th sub-interval of the dyadic partition of level $k$:

$$P_k^i \triangleq P \cap \begin{cases} [i\Delta_k, (i+1)\Delta_k), & i = 0, \ldots, 2^k - 2, \\ (2^k - 1)\Delta_k, t_n], & i = 2^k - 1. \end{cases}$$

Note that each $P_k^i$ contains at most one point for $\ell \in \mathbb{N}$ sufficiently large.

We define the random variables $\tilde{R}_k^i$ by

$$\tilde{R}_k^i \triangleq M(P_k^i) - M(P_{2i}^{k+1}) - M(P_{2i+1}^{k+1}). \tag{3.6}$$

Since $P_k^i = P_{2i}^{k+1} \cup P_{2i+1}^{k+1}$, $\tilde{R}_k^i$ vanishes unless both $P_{2i}^{k+1}$ and $P_{2i+1}^{k+1}$ are non-empty. Since $P_0^0 = P$, it holds that

$$\sum_{i=1}^{n-1} S_{t_{u}-t_{i+1}} \Xi_{t_{i},t_{i+1}} - \Xi_{t_{0},t_{n}} = M(P_0^0).$$

Also, rearranging the definition (3.6) we have

$$M(P_0^0) = M(P_0^0) + M(P_1^0) + \tilde{R}_0^0.$$ 

Iterating this identity shows that, for any $\ell \in \mathbb{N}$,

$$\sum_{i=1}^{n-1} S_{t_{u}-t_{i+1}} \Xi_{t_{i},t_{i+1}} - \Xi_{t_{0},t_{n}} = \sum_{i=0}^{2\ell-1} M(P_i^0) + \sum_{k=0}^{\ell-1} \sum_{i=0}^{2^k-1} \tilde{R}_k^i.$$

Recall that, for $\ell$ sufficiently large, we have $M(P_\ell^0) = 0$.

Finally, if either of $P_{2i}^{k+1}$ and $P_{2i+1}^{k+1}$ is empty, then $\tilde{R}_k^i = 0$ so that we can choose

$$r_k^i = u_k^i = v_k^i = w_k^i = i\Delta_k.$$

Otherwise, we can pick

$$r_k^i \triangleq \min(P_{2i}^{k+1}), \quad u_k^i \triangleq \max(P_{2i}^{k+1}), \quad v_k^i \triangleq \min(P_{2i+1}^{k+1}), \quad w_k^i \triangleq \max(P_{2i+1}^{k+1}).$$

In fact since $P_i^k = P_{2i}^{k+1} \cup P_{2i+1}^{k+1}$, for the numbers $r_k^i, u_k^i, v_k^i, w_k^i$ constructed earlier we have that

$$\tilde{R}_k^i = S_{t_{u}-t_{v}} \Xi_{u_k^i,v_k^i} + S_{t_{v}-t_{u}} \Xi_{v_k^i,u_k^i} + S_{t_{u}-t_{v}} \Xi_{v_k^i,u_k^i} - S_{t_{v}-t_{u}} \Xi_{u_k^i,v_k^i} = -S_{t_{u}-t_{v}} \left( \delta \Xi_{u_k^i,v_k^i} + \delta \Xi_{u_k^i,v_k^i} + \delta \Xi_{v_k^i,u_k^i} \right).$$

We define $R_k^i$ to be $\tilde{R}_k^i$ to conclude the construction of (3.4) and the proof of the proposition. 

Next, we present the main result of this section:

**Proposition 3.2** (Mild stochastic sewing lemma). Let $\kappa \in \mathbb{R}$. Let $p \geq 2$, $\eta > \frac{1}{2}$, and $\bar{\eta} > 1$. Suppose that $\Xi \in H_\eta^p(\mathcal{H}_\kappa) \cap \bar{H}_{\bar{\eta}}^p(\mathcal{H}_\kappa)$. Then,

(i) for every $0 \leq s \leq t \leq T$, the limit

$$I\Xi_{s,t} \triangleq \lim_{|P| \to 0} \sum_{[u,v] \in P} S_{t-u} \Xi_{u,v} \tag{3.7}$$

exists in $L^p(\Omega; \mathcal{H}_\kappa)$ along any sequence of partitions $P$ of $[s,t]$ with mesh tending to zero and the process $I\Xi$ is additive in the following sense:

$$S_{t-s} I\Xi_{r,s} + I\Xi_{s,t} = I\Xi_{r,t} \quad \forall (r,s,t) \in \Delta^3.$$

Moreover $I\Xi$ vanishes if $\|\Xi_{s,t}\|_{L^p(\mathcal{H}_\kappa)} \lesssim |t-s|^\bar{\eta}$. 

15
(ii) If, in addition, for any \((r, s, t) \in \Delta^3\), there exists \(\Lambda\) such that
\[
\delta \Xi_{r,s,t} = S_{t-s}\Lambda_{r,s,t}
\]
and such that
\[
\|\Lambda_{r,s,t}\|_{L^p(\mathcal{H}_n)} \leq C|t - r||t - s|^{\tilde{q} - 1}, \tag{3.8}
\]
then, for any \(\varpi \in [0, \tilde{q})\),
\[
\|\mathcal{I}\Xi_{s,t} - \Xi_{s,t}\|_{L^p(\mathcal{H}_{n+\varpi})} \lesssim C|t - s|^{	ilde{q} - \varpi} \tag{3.9}
\]
for all \(0 \leq s \leq t \leq T\).

**Proof.** **Step 1: Existence.** We first show the existence of the limit (3.7) in \(L^p\). Let \(P = \{t_i\}_{i=1}^n\) and \(P' = \{t_i\}_{i=1}^m\) be two partitions of \([s, t]\) and \(Q = P \cup P' = \{s_i\}_{i=1}^m\) their mutual refinement. Denote
\[
\Xi^P = \sum_{[u,v] \in P} S_{t-u} \Xi_{u,v}
\]
and similar for the other partitions. Define
\[
Z_i = \sum_{j:s_j \in [t_i, t_{i+1}]} S_{t-s_{j+1}} \Xi_{s_j, s_{j+1}} - S_{t-t_{i+1}} \Xi_{t_i, t_{i+1}}
\]
so that
\[
\Xi^Q - \Xi^P = \sum_{i=1}^{n-1} Z_i.
\]
The inequality (3.1) and Lemma 3.1 applied to each sub-interval \([t_i, t_{i+1}]\), therefore show that
\[
\|\Xi^Q - \Xi^P\|_{L^p(\mathcal{H}_n)} \lesssim \sum_{i=1}^{n-1} \|\mathbb{E}[Z_i | \mathcal{F}_{t_i}]\|_{L^p} + \left(\sum_{i=1}^{n-1} \|Z_i\|_{L^p}\right)^{\frac{1}{2}}
\lesssim \sum_{i=1}^{n-1} |t_{i+1} - t_i|^{\tilde{q}} + \left(\sum_{i=1}^{n-1} |t_{i+1} - t_i|^{2\tilde{q}}\right)^{\frac{1}{2}} \lesssim |P|^\rho,
\]
where \(\rho \triangleq (\tilde{q} - 1) \wedge (\eta - \frac{1}{2}) > 0\). Similarly we find \(\|\Xi^Q - \Xi^{P'}\|_{L^p(\mathcal{H}_n)} \lesssim |P'|^\rho\), whence
\[
\|\Xi^P - \Xi^{P'}\|_{L^p(\mathcal{H}_n)} \lesssim (|P| \vee |P'|)^\rho.
\]
Consequently, the sequence \((\Xi^{P_n})_n)\) is Cauchy in \(L^p\) along any sequence of partitions \(P_n\) of \([s, t]\) with mesh \(|P_n| \to 0\), and the existence of \(\mathcal{I}\Xi_{s,t}\) follows. Note that the asserted ‘additivity’ \(S_{t-u}\Xi_{s,u} + \Xi^P_{t,s} = \Xi^P_{s,t}\) holds for each partition \(P\). It passes immediately to the limit \(\mathcal{I}\Xi\) from this construction.

**Step 2: Uniqueness.** Suppose that \(\|\Xi_{s,t}\|_{L^p(\mathcal{H}_n)} \lesssim |t - s|^{\tilde{q}}\). We show that then \(\mathcal{I}\Xi = 0\). This is in fact an immediate consequence of the a priori \(L^p\)-bound:
\[
\|\mathcal{I}\Xi_{s,t}\|_{L^p(\mathcal{H}_n)} \lesssim \lim_{|P| \to 0} \sum_{[u,v] \in P} \|\Xi_{u,v}\|_{L^p(\mathcal{H}_n)} \lesssim \lim_{|P| \to 0} |P|^{\tilde{q} - 1} = 0.
\]
Step 3: The bound (3.9). We consider an approximation of $I\Xi_{s,t}$ along the dyadic rationals in the interval $[s, t]$. Recall our notation $D_n = \{ s + k\frac{t-s}{2^n} : k = 0, \ldots, 2^n \}$. Then we have

$$D_{n+1}^{s,t} = \sum_{[u,v] \in D_{n+1}} S_{t-s} \Xi_{u,v}$$

$$= D_n^{s,t} - \sum_{[u,w] \in D_n} (S_{t-w} \Xi_{u,w} - S_{t-w} \Xi_{w,u})$$

$$= D_n^{s,t} - \sum_{[u,w] \in D_n} S_{t-w} \hat{\Xi}_{u,w}.$$

Hence, we find

$$\left\| D_n^{s,t+1} - D_n^{s,t} \right\|_{L^p(\mathcal{H}_{s,t})} = \left\| \sum_{[u,w] \in D_n} S_{t-w} \hat{\Xi}_{u,w} \right\|_{L^p(\mathcal{H}_{s,t})} = \left\| \sum_{[u,w] \in D_n} S_{t-w} \hat{\Xi}_{u,w} \right\|_{L^p(\mathcal{H}_{s,t})}.$$

Applying the smoothing property of $(S_t)_{t \geq 0}$, $\left\| S_t x \right\|_{\mathcal{H}_{s,t}} \lesssim t^{-\varpi} \left\| x \right\|_{\mathcal{H}}$ (see (2.4)), we have

$$\left\| D_n^{s,t+1} - D_n^{s,t} \right\|_{L^p(\mathcal{H}_{s,t})} \lesssim \sum_{[u,w] \in D_n} \left( t - \frac{w + u}{2} \right)^{-\varpi} \left\| \Lambda_{u,w} \right\|_{L^p(\mathcal{H})} \lesssim \frac{C(t-s)^{\bar{\eta}}}{2n^{(n+1)(\bar{\eta}-1)}} \sum_{[u,w] \in D_n} \left( t - \frac{w + u}{2} \right)^{-\varpi} \lesssim \frac{C(t-s)^{\bar{\eta}}}{2n^{(n+1)(\bar{\eta}-1)}} \sum_{[u,w] \in D_n} \left( t - \frac{w + u}{2} \right)^{-\varpi}.$$

Since $t - \frac{w + u}{2} \geq \frac{t-s}{2^n}$, for each $\epsilon \in ((\underline{\varpi}-1, \bar{\eta} - 1)$,

$$\sup_{n \in \mathbb{N}} \frac{t-s}{2^n} \frac{(t-s)}{2^{n+1}} \sum_{[u,w] \in D_n} \left( t - \frac{w + u}{2} \right)^{-\varpi} \leq \sup_{n \in \mathbb{N}} \frac{t-s}{2^n} \sum_{[u,w] \in D_n} \left( t - \frac{w + u}{2} \right)^{-\varpi + \epsilon}$$

$$\leq \int_s^t (t-r)^{-\varpi + \epsilon} \, dr \lesssim (t-s)^{1-\varpi + \epsilon},$$

where we used the fact that the Riemann sums increase monotonically to the integral. In particular, we see that

$$\left\| D_n^{t,s} - D_n^{t,s} \right\|_{L^p(\mathcal{H}_{t,s})} \lesssim \frac{C(t-s)^{\bar{\eta}-\varpi}}{2^{n(1)(\bar{\eta}-1-\epsilon)}}.$$

Since $\bar{\eta} - 1 - \epsilon > 0$, this gives

$$\left\| I\Xi_{s,t} \right\|_{L^p(\mathcal{H}_{s,t})} \leq \sum_{n=0}^{\infty} \left\| D_n^{t,s} - D_n^{t,s} \right\|_{L^p(\mathcal{H}_{t,s})} \lesssim C(t-s)^{\bar{\eta}-\varpi},$$

as required. \hfill \Box

Remark 3.3. After putting the article on the arXiv, we learnt from Khoa Lê a nice trick to deduce part (i) of Proposition 3.2 directly from the stochastic sewing lemma in [32], where the usual
incremental process is used. To see this let us fix a terminal time $t > 0$ and for $u < v < w \leq t$, and set

$$A_{u,v} = S_{t-u} \Xi_{u,v}. $$

Then

$$\delta A_{u,v,w} = S_{t-w}(\Xi_{u,v,w} - \Xi_{v,w,w} - S_{t-v} \Xi_{u,v,w}) = S_{t-w} \delta \Xi_{u,v,w}. $$

Naturally, $\Xi \in H^p_\beta(\mathcal{H}_n) \cap H^p_{\bar{\beta}}(\mathcal{H}_n)$ implies that

$$\sup_{0 \leq s,t \leq T} \|A_{u,v}\|_{L^p} < \infty, \quad \|E[\delta A_{u,v,w}]\|_{L^p} < \infty,$$

and [32, Theorem A] allows to conclude that $I\Xi_{s,t} = \lim \sum_{[u,v]} S_{t-u} \Xi_{u,v}$. The same trick does not seem to deduce the error bound in Proposition 3.2 from the corresponding estimate in [1, Theorem 4.1]. The conditions in this section appear to be of more deterministic nature. The ‘seesaw’ allowing us to trade interpolation space regularity for time regularity is crucial for the application in the sequel. Let $s < u < w \leq t$. In [1, Theorem 4.1] two conditions are imposed, the first of which is: $\|E[\delta A_{s,u,w}]\|_{L^p} \lesssim u^{-\alpha_1} |t-u|^{-\beta_1} |w-s|^{1+\epsilon_1}$ where $\alpha_1, \beta_1 \in (0, 1)$. The second is on the conditional norm of the $L^p$ norm of $\delta A_{s,u,w}$ with constants $\alpha_2, \beta_2$ restricted to $(0, \frac{1}{2})$—note no analogous condition is needed here. Then the authors obtained an estimate of the order $|t-s|^{-1}$. In Equation (3.9) the estimate is of the order $|t-s|^{\bar{\eta}-\bar{\epsilon}}$ with $\bar{\omega} \in [0, \bar{\eta})$ and $\bar{\eta} > 1$. In addition, given the restriction $\alpha_1, \beta_1 \in (0, 1)$, it does not seem possible to deduce the aforementioned condition on $\|E[\delta A_{s,u,w}]\|_{L^p}$ from our bound $\|\Lambda_r,\|_{L^p(\mathcal{H}_n)} \leq C |w-s| |w-u|^{\bar{\eta}-1}$. In fact, from the inequality

$$\|\delta A_{s,u,w}\|_{L^p(\mathcal{H}_n)} \lesssim \|S_{t-w} \Xi_{s,u,v,w}\|_{L^p(\mathcal{H}_n)} \lesssim \|t-w|^{-\alpha} \|\delta \Xi_{s,u,w}\|_{L^p(\mathcal{H}_n)} \lesssim \|t-u|^{-\alpha} \|\delta \Xi_{s,u,w}\|_{L^p(\mathcal{H}_n)},$$

one has only $\|\delta A_{s,u,w}\|_{L^p(\mathcal{H}_n)} \lesssim |t-w|^{-\bar{\omega}} |w-s| |w-u|^{\bar{\eta}-1} \lesssim |t-w|^{-\bar{\omega}} |w-s|^\bar{\eta}$. One can attempt to use $\alpha_1 = 0, \beta_1 = \bar{\omega}$ and $\epsilon_1 = \bar{\eta} - 1$. However $\bar{\omega}$ may fall out of the permissible range of $\beta_1$.

4 Uniform $L^p$-Estimates

The aim of this section is to leverage the mild stochastic sewing lemma (Proposition 3.2) to obtain strong $L^p$-bounds on the SPDE

$$dx_t = (Ax_t + f(t, x_t)) \, dt + g(t, x_t) \, dB_t, \quad (4.1)$$

where $B$ is a trace-class fBm, see § 4.1 for details. We emphasize that these bounds do not follow from the pathwise approach taken in § 2.3.

4.1 Trace-Class Fractional Brownian Motions

A one-dimensional fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$ is the centered Gaussian process $(\beta_t)_{t \geq 0}$ with covariance

$$E[\beta_s \beta_t] = \frac{1}{2} \left( t^{2H} + s^{2H} - |t-s|^{2H} \right) \quad \forall s, t \geq 0.$$  

For $H = \frac{1}{2}$, this is the standard Wiener process.

Let $Q \in L(\mathcal{K})$ be a symmetric, non-negative trace-class operator on a separable Hilbert space $\mathcal{K}$. Recall a trace-class operator is a compact. Let $\{\lambda_n\}_{n \in \mathbb{N}} \subset \mathbb{R}_+$ be its discrete family of eigenvalues, counted with multiplicity. The associated normalized eigenvectors $Q e_n = \lambda_n e_n$ form an orthonormal basis of $\mathcal{K}$. We refer the readers to the classical book by Da Prato and Zabczyk [9]. Albeit the interest there lies in Markovian systems, we follow their terminology.
Definition 4.1. An $K$-valued centred Gaussian process $(B_t)_{t \geq 0}$ is called a trace-class fractional Brownian motion (fBm) with covariance $Q$ and Hurst parameter $H$ if

$$
\mathbb{E}[(B_s, x)(B_t, y)] = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}) \langle Qx, y \rangle \quad \forall s, t \geq 0 \quad \forall x, y \in K.
$$

It is also called a $Q$-fBm.

Remark 4.2. In the study of stochastic partial differential equations driven by trace-class fBms, it is often assumed that $\sqrt{Q}$ is trace-class. See e.g. Garrido-Atienza-Lu-Schmalfuß [16], Pei-Xu-Bai [40], and Nascimento-Ohashi [10]. We emphasize that we do not impose this stronger condition.

Set $\beta^n_t = 1/\sqrt{\lambda^n} (B_t, e_n)$. Then, $\{\beta^n\}$ are independent one-dimensional fBm’s with parameter $H$. Conversely, for any orthonormal basis $\{e_n\}_{n \in \mathbb{N}} \subset K$ and any sequence $\{\beta^n\}_{n \in \mathbb{N}}$ of i.i.d. one-dimensional fBm’s, the series

$$
B_t = \sum_{n=1}^{\infty} \sqrt{\lambda^n} \beta^n_t e_n, \quad t \geq 0,
$$

(4.2)

which converges almost surely and in $L^p$ ($p \geq 1$), is a $Q$-fBm:

Lemma 4.3. A trace-class fBm $(B_t)_{t \in [0,T]}$ with Hurst parameter $H$ takes values in $C^{H-}([0,T],K)$ with probability one.

Let $\{\beta^n_t\}$ be i.i.d. one-dimensional centred Gaussian processes with continuous sample paths with finite variance with $t \in [0,T]$. Then, for any sequence $(\lambda_n) \in \ell^1(\mathbb{R}_+)$ and any orthonormal system $(e_n)_{n \in \mathbb{N}} \subset K$, the series $\sum_{n=1}^{\infty} \sqrt{\lambda^n} \beta^n_t e_n$ converges in $C([0,T],K)$ $\mathbb{P}$-almost surely.

The first statement is a straightforward application of the Kolmogorov continuity theorem, while the second follows from the convergence of $\sup_{t \leq T} \sum_{k=1}^{n} \sqrt{\lambda^n} \beta^n_t e_k$ for any $T > 0$ in $L^2$. This in turn is due to the fact that $\| \sum_{n=1}^{\infty} \sqrt{\lambda^n} \beta^n_t e_n \|_{L^2} \leq \mathbb{E} \left[ \| \beta^n \|_2 \right] \sum_n \lambda_n < \infty$.

4.2 Basic Properties of Integrals Against Trace-Class fBm

Our analysis in the sequel is underpinned by the following Mandelbrot-van Ness representation of the one-dimensional fBm with Hurst parameter $H \in (0,1)$ [37]:

$$
\beta_t = \alpha_H \left( \int_{-\infty}^{0} (t-u)^{H-\frac{1}{2}} \, dW_u - \int_{0}^{t} (t-u)^{H-\frac{1}{2}} \, dW_u \right). \quad (4.3)
$$

Here $\alpha_H > 0$ is an explicitly known normalization constant. By virtue of (4.3), we have the following locally independent decomposition of the one-dimensional fBm increment [23, 33]. For $t, h \geq 0$,

$$
\beta_{t+h} - \beta_t = \alpha_H \int_{-\infty}^{t} (t+h-u)^{H-\frac{3}{2}} \, dW_u + \alpha_H \int_{t}^{t+h} (t+h-u)^{H-\frac{3}{2}} \, dW_u + \beta^h_h + \beta^0_h.
$$

(4.4)

If $t = 0$, we shall write $\beta^h_h \triangleq \beta^0_h$ and similarly $\beta^h_h \triangleq \beta^0_h$. For $h > 0$, there is no singularity in the integrand of the first term, whence an integration by parts formula turns it into a Riemann integral. This is exploited in the next lemma:

Lemma 4.4. For each $t \geq 0$, the mapping $(0, \infty) \ni h \mapsto \beta^t_h$ is smooth with probability one and we have

$$
\frac{d}{dh} \beta^t_h = \alpha_H \left( H - \frac{1}{2} \right) \int_{-\infty}^{t} (t+h-u)^{H-\frac{3}{2}} \, dW_u, \quad (4.5)
$$

and

$$
\frac{d^2}{dh^2} \beta^t_h = \alpha_H \left( H - \frac{1}{2} \right) \left( H - \frac{3}{2} \right) \int_{-\infty}^{t} (t+h-u)^{H-\frac{3}{2}} \, dW_u \quad (4.6)
$$

for each $h > 0$. 


Proof. We only demonstrate the computation of the first derivative as the second one is similar. Fix \( h \in (0, \infty) \) to be \( h \geq 0 \). Let us begin with an integration by parts:

\[
\left[\frac{d}{dh} \lambda_{h} \phi_{h}' \right]_{t} = \left( H - \frac{1}{2} \right) \int_{-\infty}^{t} (t + h - u)^{H-\frac{3}{2}} (W_{u} - W_{t}) du.
\]

Note that there is an almost surely finite random variable \( C > 0 \) such that

\[
|W_{u} - W_{t}| \leq \begin{cases} C|u - t|, & u \in [t - 1, t], \\ C(|u|^{\frac{1}{2}} + 1), & u \in (-\infty, t - 1). \end{cases}
\]

In particular, the boundary term \((4.7)\) vanishes. Moreover, we may differentiate under the integral sign to obtain

\[
\alpha_{h}^{-1} \frac{d}{dh} \lambda_{h} = \alpha_{h}^{-1} \frac{d}{dh} \lambda_{h} \phi_{h}' = \left( H - \frac{1}{2} \right) \int_{0}^{t} (t + h - u)^{H-\frac{3}{2}} (W_{u} - W_{t} - \frac{1}{2}(W_{u} - W_{t})) du.
\]

Again, the boundary term vanishes and \((4.5)\) follows at once. \(\Box\)

The increment decomposition \((4.4)\) gives rise to a similar decomposition of the trace-class fBm \((4.2)\):

\[
B_{t+h} - B_{t} = \sum_{n=1}^{\infty} \sqrt{\lambda_{n}} \beta_{h}^{n,t} e_{n} + \sum_{n=1}^{\infty} \sqrt{\alpha_{n}} \beta_{h}^{n,t} e_{n} \triangleq B_{h}^{t} + \tilde{B}_{h}^{t}.
\]

Once again, these series converge almost surely and in \(L^{p}(\Omega; K)\) for any \( p \geq 1 \).

Lemma 4.5. Let \( t \geq 0 \). The mapping \( h \in (0, \infty) \mapsto \tilde{B}_{h}^{t} \) is almost surely smooth. Moreover, for each \( p \geq 1 \), we have that

\[
\|	ilde{B}_{h}^{t}\|_{L^{p}} \lesssim \frac{1}{h^{1-\frac{1}{2}}}, \quad \|	ilde{B}_{h}^{t}\|_{L^{p}} \lesssim \frac{1}{h^{2-\frac{1}{2}}},
\]

uniformly in \( t \geq 0 \) and \( h \geq 0 \).

Proof. Fix \( t \geq 0 \). Let \( h \geq 0 \). We claim that

\[
\tilde{B}_{h}^{t} = \sum_{n=1}^{\infty} \sqrt{\lambda_{n}} \beta_{h}^{n,t} e_{n}.
\]

To see this, for any \( n \in \mathbb{N} \), let us abbreviate \( \Theta_{h}^{N} = \sum_{n=1}^{\infty} \sqrt{\lambda_{n}} \beta_{h}^{n,t} e_{n} \). By Lemma 4.4, we know that \( \frac{d}{dh} \Theta_{h}^{N} = \sum_{n=1}^{N} \sqrt{\lambda_{n}} \beta_{h}^{n,t} e_{n} \) for each \( h > 0 \). Thanks to Lemma 4.3, we see that the partial sums \( \Theta_{h}^{N} \) and \( \frac{d}{dh} \Theta_{h}^{N} \) converge \( \mathbb{P} \)-a.s. uniformly on \((0, \infty)\) to \( \tilde{B}^{t} \) and \( \tilde{B}^{t} \), respectively. In particular, \( h \mapsto \tilde{B}_{h}^{t} \) is almost surely differentiable and \( \frac{d}{dh} \tilde{B}_{h}^{t} = \tilde{B}_{h}^{t} \) for each \( h > 0 \).

For the \( L^{p} \)-bound, we note \( \tilde{B}_{h}^{t} \) is centered Gaussian with variance \( \sum_{n} \lambda_{n} \mathbb{E}[\tilde{B}_{h}^{n,t}]^{2} \). By \((4.5)\), we find

\[
\mathbb{E}[\tilde{B}_{h}^{n,t}]^{2} \lesssim \int_{-\infty}^{0} (h - u)^{2H-3} du \lesssim h^{2H-2}
\]

and the bound on \( \|	ilde{B}_{h}^{t}\|_{L^{p}} \) follows by Gaussianity. The second derivative can be handled similarly using \((4.6)\). \(\Box\)
Let \((\mathcal{F}_t^B)_{t \geq 0}\) be the filtration generated by the trace-class fBm. Let \(g : [s, t] \to L(\mathcal{K}, \mathcal{H})\) be \(\mathcal{F}_s\)-measurable and write \(g_n \triangleq g e_n\) with the orthonormal basis in the series representation of \(B\) (4.2). We define a mixed Wiener-Young integral against \(B\) (which is a priori different than the Young integral):

\[
\int_s^t g(r) \, dB_r \triangleq \int_s^t g(r) \, \tilde{B}^s_{t-r} + \int_t^s g(r) \, dB^s_{r-s}
\]

\[
= \sum_{n=1}^{\infty} \sqrt{\lambda_n} \left( \int_s^t g_n(r) \, d\tilde{\beta}^{n,s}_{t-r} + \int_t^s g_n(r) \, \tilde{\beta}^{n,s}_{r-s} \, dr \right),
\]

where the first integral is the Wiener integral against the Gaussian process \((\tilde{\beta}^{n,s})_{t \geq 0}\). Set

\[
\mathbb{R}_s(r, s) = \mathbb{E} \left[ \tilde{\beta}^1_r \tilde{\beta}^1_s \right].
\]

**Lemma 4.6.** The Wiener integral \(\int_s^t g(r) \, dB^s_{t-r}\) is a centered Gaussian process with covariance operator \(Q_g\) given by

\[
\langle x, Q_g y \rangle \triangleq \sum_{n=1}^{\infty} \lambda_n \int_s^t \int_s^t \langle g_n(v), x \rangle \langle g_n(u), y \rangle \frac{\partial^2}{\partial u \partial v} \mathbb{R}(u - s, v - s) \, du \, dv, \quad \forall x, y \in \mathcal{H}.
\]

**Proof.** This follows from \(\int_s^t g(r) \, dB_r = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \int_s^t g_n(r) \, d\tilde{\beta}^{n,s}_{t-r}\) and the Itô isometry [29],

\[
\mathbb{E} \left[ \left( \int_s^t \langle g_n(r), x \rangle \, d\tilde{\beta}^{n,s}_{t-r} \right) \left( \int_t^s \langle g_n(r), y \rangle \, d\tilde{\beta}^{n,s}_{t-r} \right) \right] = \int_s^t \int_s^t \langle g_n(u), x \rangle \langle g_n(v), y \rangle \frac{\partial^2}{\partial u \partial v} \mathbb{R}(u - s, v - s) \, du \, dv.
\]

Let \(|G|_{C^{-\delta}} = \sup_{0 \leq s \leq t \leq T} |s - t|^{|\alpha|+1} \| \int_s^t g(r) \, dr \|_{L^\infty}\) denote the Hölder norm of a function \(G : [0, T] \to \mathcal{E}\) of negative exponent \(-\delta\) where \(\mathcal{E}\) is a Banach space.

**Lemma 4.7.** Let \(0 \leq s \leq t \leq T\) and \(\delta \in [0, 1)\). Let \(g \in C^{-\delta}([s, t], L(\mathcal{K}, \mathcal{H}))\) be \(\mathcal{F}_s\)-measurable. Then, for each \(1 \leq p < q\),

\[
\left\| \int_s^t g(r) \, dB_r \right\|_{L^p} \lesssim \|g|_{C^{-\delta}}\|_{L^q} |t - s|^{H-\delta}
\]

with a prefactor uniform in \(s, t \in [0, T]\).

**Proof.** Let us first assume that \(g \in C^{-\delta}\) is smooth. We prove the asserted \(L^p\)-bound for both integrals in (4.8) individually. Let us begin with the smooth part: An integration by parts shows that

\[
\int_s^t g(r) \, dB^s_{t-r} = \left( \int_s^t g(r) \, dr \right) \tilde{B}^s_{t-s} - \int_s^t \left( \int_r^t g(u) \, du \right) \tilde{B}^s_{r-s} \, dr.
\]

In particular, using Lemma 4.5 we obtain

\[
\left\| \int_s^t g(r) \, dB^s_{t-r} \right\|_{L^p} \lesssim \|g|_{C^{-\delta}}\|_{L^q} \left( |t - s|^{H-\delta} + \int_s^t |r - s|^{\delta+H-1} \, dr \right) \lesssim \|g|_{C^{-\delta}}\|_{L^q} |t - s|^{H-\delta}.
\]

For the rough part, we simply notice that the Wiener integral satisfies

\[
\int_s^t g(r) \, dB^s_{t-r} = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \int_s^t g_n(r) \, d\tilde{\beta}^{n,s}_{t-r}.
\]
Since \((\hat{B}^n_t)_{t \in [0, t-s]}\) is independent of the stochastic processes \(\{g(r)\}_{r \in (s, t)}\), the integral is a Gaussian variable. It suffices to estimate the \(L^2\)-norm. By Lemma 4.6 we obtain

\[
\left\| \int_s^t g(r) \, d\hat{B}^n_{r-s} \right\|_{L^2}^2 = \sum_{n=1}^{\infty} \lambda_n \mathbb{E} \left[ \int_s^t \int_s^t \langle e_n, g(v)g^*(u)e_n \rangle \frac{\partial^2}{\partial u \partial v} \mathbb{R}(u-s, v-s) \, dv \, du \right].
\]

It follows from [24, Lemma A.1], c.f. also Lemma 3.4 there,

\[
\int_s^t \int_s^t \langle e_n, g(v)g^*(u)e_n \rangle \frac{\partial^2}{\partial u \partial v} \mathbb{R}(u-s, v-s) \, dv \, du \lesssim |t-s|^{2(\alpha-\delta)} |g_n g^*_n|_{C^{\alpha-\delta}}.
\]

Notice that \(|g_n^*|_{C^{\alpha-\delta}} \leq |g|_{C^\alpha}\), whence

\[
\left\| \int_s^t g(r) \, d\hat{B}^n_{r-s} \right\|_{L^2}^2 \lesssim |t-s|^{2(\alpha-\delta)} \left\| g \right\|_{C^{\alpha-\delta}}^2 \sum_{n=1}^{\infty} \lambda_n \lesssim |t-s|^{2(\alpha-\delta)} \left\| g \right\|_{C^{\alpha-\delta}}^2,
\]

completing the proof. \(\square\)

### 4.3 \(L^p\)-Estimates and Stochastic Stability

Let \((B_t)_{t \geq 0}\) be a trace-class fBm with Hurst parameter \(H > \frac{1}{2}\), driving the SPDE (4.1), and \((\mathcal{F}_t^B)_{t \geq 0}\) is its filtration. Let \(\mathcal{E}\) and \(\mathcal{K}\) be separable Hilbert spaces and let

\[
g : \Omega \times [0, T] \times \mathcal{E} \to L(\mathcal{K}, \mathcal{E})
\]

be a random field with values in the space of linear operators \(L(\mathcal{K}, \mathcal{E})\), independent of \(\mathcal{F}_t^B\). Set \(\mathcal{F}_t^g = \sigma(g(s, x), s \leq t, x \in \mathcal{E})\). For \(\gamma \in (0, 1], \delta \in (0, 1)\) and \(p \geq 1\) we define the ‘norm-like’ quantity

\[
\left\| g \right\|_{A^{\gamma, \delta}_p} \triangleq \sup_{s \in T} \left( \sup_{X \in \mathcal{F}_s^g \vee \mathcal{F}_s^g} \left\| g(\cdot, X) \right\|_{C^{\alpha-\delta}(s,T)} \right)_{L^p} + \sup_{X, Z \in \mathcal{F}_s^g \vee \mathcal{F}_s^g} \left\| \left\| g(\cdot, X) - g(\cdot, Z) \right\|_{C^{\alpha-\delta}(s,T)} \right\|_{L^p}.
\]

(4.9)

Since it depends on the filtration generated by \(g\), the triangle inequality may not hold. In the sequel, we only use the triangle inequality coming from the \(C^{\alpha-\delta}(s,T)\) norm. Eventually, we take the Hilbert space \(\mathcal{E} = H^{\kappa+\alpha}\), the random field \((t, x) \mapsto g(x, Y_t^x)\), and \(\mathcal{F}_t^g = \mathcal{F}_t^Y\).

Let \(X\) be the solution to the SPDE (1.3) interpreted in the pathwise mild sense of § 2. Our aim is to derive \(L^p\)-estimates on the integral

\[
\int_0^T g(t, X_t) \, dB_t.
\]

We have the following consequence of Proposition 3.2 in which we usually choose \(\kappa = -\alpha\). Recall the definition of \(\bar{B}_{\alpha, p}([0, T], \mathcal{H})\) from (2.8).

**Proposition 4.8.** Let \(p \geq 2, \delta \in (0, 1), \kappa \in \mathbb{R}, \gamma \in (0, 1),\) and \(\alpha \in \left(\frac{\gamma}{2}, 1\right)\). Let \(\eta = H - \delta\) and \(\bar{\eta} = \eta + \gamma\alpha\). Let \(X \in \bar{B}_{\alpha, p}([0, T], H^{\kappa+\alpha})\) with \(X_0 \in L^p(H^{\kappa+\alpha})\). Suppose that

\[
g : \Omega \times [0, T] \times H^{\kappa+\alpha} \to L(K, H^{\kappa+\alpha})
\]

is independent of \(\mathcal{F}_t^B\) and \(\left\| g \right\|_{A^{\gamma, \delta}_p} < \infty\).
(i) If \( \eta > \frac{1}{2} \) and \( \bar{\eta} > 1 \), then the mixed Wiener-Young integral

\[
\int_s^t S_{t-r}g(r, X_r) \, dB_r
\]

defined in (4.8) is given by the \( L^p \)-limit of Riemann sums along an arbitrary sequence of partitions of \([s, t]\) with mesh tending to zero.

(ii) Moreover, it satisfies the estimate

\[
\left\| \int_s^t S_{t-r}g(r, X_r) \, dB_r \right\|_{L^p(\mathcal{H}_{\alpha+n})} \lesssim \|g\|_{\mathcal{A}_{\gamma}^{p/\gamma}(L(\mathcal{K}, \mathcal{H}_{\alpha+n}))} \left( 1 + \|X\|_{\mathcal{B}_{\alpha,p}}^\gamma \right) |t - s|^{\bar{\eta} - \alpha}.
\]

(4.10)

(iii) If, in addition, \( g \in C^{(1-H)+\gamma} \), then this integral coincides with the infinite-dimensional Young integral \( \int_s^t S_{t-r}g(r, X_r) \, dB_r \), see Proposition 2.2.

Proof. Let

\[
\Xi_{s,t} \triangleq \int_s^t S_{t-s}g(r, X_s) \, dB_r, \quad 0 \leq s \leq t \leq T.
\]

We verify the requirements of Proposition 3.2. By Lemma 4.7,

\[
\|\Xi_{s,t}\|_{L^p(\mathcal{H}_n)} \lesssim \|S_{t-s}g(\cdot, X_s)|_{C^{-\delta}(s,t, L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} |t - s|^{H - \delta}.
\]

Since \( \|S_{t-s}g(\cdot, X_s)|_{C^{-\delta}(s,t, L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} \lesssim \|g\|_{\mathcal{A}_{\gamma}^{p/\gamma}(L(\mathcal{K}, \mathcal{H}_n))} \) uniformly in \( s, t \in [0, T] \), we conclude

\[
\|\Xi_{s,t}\|_{L^p(\mathcal{H}_n)} \lesssim \|g\|_{\mathcal{A}_{\gamma}^{p/\gamma}(L(\mathcal{K}, \mathcal{H}_n))} |t - s|^{\eta}.
\]

(4.11)

Let \( \Lambda_{r,s,t} = \int_s^t \left( S_{s-r}g(u, X_r) - g(u, X_r) \right) \, dB_u \). For the second order bound, we first notice that

\[
\delta \Xi_{r,s,t} = -S_{t-s} \int_s^t \left( g(u, X_s) - S_{s-r}g(u, X_r) \right) \, dB_u = S_{t-s} \Lambda_{r,s,t}.
\]

Lemma 4.7 yields

\[
\left\| \int_s^t \left( g(u, X_s) - S_{s-r}g(u, X_r) \right) \, dB_u \right\|_{L^p(\mathcal{H}_n)} \lesssim \|g(\cdot, X_s) - S_{s-r}g(\cdot, X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} |t - s|^{\eta}.
\]

By the triangle inequality, we find

\[
\|g(\cdot, X_s) - S_{s-r}g(\cdot, X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} \leq \|g(\cdot, X_s) - g(\cdot, S_{s-r}X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} + \|g(\cdot, S_{s-r}X_r) - g(\cdot, X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}}
\]

\[
+ \|((id - S_{s-r})g(\cdot, X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}}.
\]

For the first term, we have the bound

\[
\|g(\cdot, X_s) - g(\cdot, S_{s-r}X_r)|_{C^{-\delta}(L(\mathcal{K}, \mathcal{H}_n))}\|_{L^{p/\gamma}} \lesssim \|g \|_{\mathcal{A}_{\gamma}^{p/\gamma}(L(\mathcal{K}, \mathcal{H}_{\alpha+n}))} \|X_s - S_{s-r}X_r\|_{\mathcal{H}_{\alpha+n}} \|X\|_{\mathcal{B}_{\alpha,p}([0, T], \mathcal{H}_{\alpha+n})} |s - r|^{\bar{\eta} \alpha}.
\]

23
Also,
\[
\| g(\cdot, S_{s-r}X_r) - g(\cdot, X_r) \|_{C^\alpha(L(H,K))} \leq \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \| X_r - S_{s-r} X_r \|_{H^{\alpha+\alpha}}^{\alpha} \| X_r \|_{H^{\alpha}}^{\alpha} \sim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \| X_r \|_{H^{\alpha}} \| s - r \|^{\alpha}
\]

The third term is bounded as below:
\[
\| (\text{id} - S_{s-r}) g(\cdot, X_r) \|_{C^\alpha(L(H,K))} \sim \| s - r \|^{\alpha} \| g(\cdot, X_r) \|_{C^\alpha(L(H,K)_{H^{\alpha+\alpha}})} \| X_r \|_{H^{\alpha}} \| s - r \|^{\alpha}
\]

Our bounds show that
\[
\| \Delta_{r,s,t} \|_{L^p(H^2)} \lesssim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \left( \| X_r \|_{L^p(H^{\alpha+\alpha})} + \| X \|_{B_{0,p}(H^\alpha)} \right) |t - s|^\gamma |s - r|^\alpha
\]
\[
\lesssim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \left( 1 + \| X \|_{B_{0,p}(H^\alpha)} \right) |t - s|^\gamma |s - r|^\gamma.
\]

Here, we used that \( \| X_r \|_{L^p(H^{\alpha+\alpha})} \lesssim 1 + \| X \|_{B_{0,p}(H^{\alpha+\alpha})} \) and \( \gamma \alpha < 1 \) as well as \( \eta > \tilde{\eta} - 1 \). In particular, we have also shown that
\[
\| \delta \Xi_{r,s,t} \|_{L^p(H^2)} \lesssim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \left( |t - s|^\gamma |s - r|^\gamma, \right.
\]
whence \( \Xi \in H_{L^p(H^2)} \cap H_{B_{0,p}(H^\alpha)} \). The mild stochastic sewing lemma (Proposition 3.2) applies. Hence, (4.11) yields
\[
\left\| \int_s^t S_{t-r} g(r, X_r) \, dB_r \right\|_{L^p(H^{\alpha+\alpha})} \lesssim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \left( |t - s|^\gamma |s - r|^\gamma, \right.
\]
\[
\left. \| X \|_{B_{0,p}(H^\alpha)} \right| |t - s|^{\gamma} + \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \left| t - s \right|^\gamma.
\]

Since \( \tilde{\eta} - \alpha \leq \eta \), the estimate (4.10) follows.

Let \( g \in C^{1-H+\epsilon,\gamma} \) for some \( \epsilon > 0 \). It remains to show that \( \int_s^t S_{t-r} g(r, X_r) \, dB_r \) coincides with the infinite-dimensional Young integral of Proposition 2.2. But this is an easy consequence of the uniqueness statement of the mild stochastic sewing lemma and Lemma 4.7:
\[
\| \Xi_{u,v} - S_{v-u} g(u, X_u) B_{u,v} \|_{L^p(H^{\alpha+\alpha})} \leq \left\| \int_u^v \left( g(r, X_u) - g(u, X_u) \right) \, dB_r \right\|_{L^p(H^{\alpha+\alpha})} \lesssim \| g \|_{A^{\beta,\gamma}_{\kappa,\delta}(L(H,K)_{H^{\alpha+\alpha}})} \| |v - u|^H
\]
\[
\lesssim |v - u|^\rho,
\]
for some \( \rho > 1 \).

**Lemma 4.9.** Let \( X \in B_{0,p}[0,T] \) and \( f : \Omega \times [0,T] \times H^{\alpha+\alpha} \rightarrow H^{\alpha+\alpha} \) is independent of \( F^B \) with \( \| f \|_{A^{\beta,\gamma}_{\kappa,\delta}} < \infty \), where \( \kappa \in \mathbb{R} \), \( \delta \in (0,1) \), and \( \gamma(0,1) \). If \( \delta < \gamma \alpha \), then
\[
\left\| \int_s^t S_{t-r} f(r, X_r) \, dr \right\|_{L^p(H^{\alpha+\alpha})} \lesssim \| f \|_{A^{\beta,\gamma}_{\kappa,\delta}(H^{\alpha+\alpha})} \left( 1 + \| X \|_{B_{0,p}(H^\alpha)} \right) \left| t - s \right|^{1-\delta-(1-\gamma)\alpha}(4.12)
\]
for each \( 0 \leq s \leq t \leq T \).
Proof. To bound the left-hand side, we employ Proposition 3.2 with

\[ \Xi_{s,t} \triangleq S_{t-s} \int_s^t f(r, X_s) \, dr. \]

Clearly, for any \( \gamma > 0 \), we have

\[ \| \Xi_{s,t} \|_{L^p(\mathcal{H})} \lesssim \| f \|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{L}(\mathcal{K},H_{\alpha+p}))} |t-s|^{1-\delta}. \]

Since

\[ \delta \Xi_{r,s,t} = -S_{t-s} \int_s^t f(u, X_s) - S_{s-r} f(u, X_r) \, du, \]

we have \( \Lambda_{r,s,t} = \int_s^t (f(u, X_s) - S_{s-r} f(u, X_r)) \, du \). The rest of the proof follows closely that for Proposition 4.8: By definition, we have

\[ \left\| \int_s^t (f(u, X_s) - S_{s-r} f(u, X_r)) \, du \right\|_{L^p(\mathcal{H})} \lesssim \left\| f(\cdot, X_s) - S_{s-r} f(\cdot, X_r) \right\|_{C^{-\delta}(\mathcal{H})} \|t-s\|^{1-\delta}. \]

This leads to the estimate

\[ \left\| \Lambda_{r,s,t} \right\|_{L^p(\mathcal{H})} \lesssim \left\| f \right\|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{H}_{\alpha+p})} \left( 1 + \|X\|^\gamma_{\tilde{B}_{r,\gamma}(\mathcal{H}_{\alpha+p})} \right) |t-s|^{\gamma - \delta} |t-r|, \]

The mild stochastic sewing lemma (Proposition 3.2) gives

\[ \left\| \int_s^t S_{t-r} f(r, X_r) \, dr \right\|_{L^p(\mathcal{H}_{\alpha+p})} \lesssim \left\| f \right\|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{H}_{\alpha+p})} |t-s|^{-\delta} + \left\| f \right\|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{H}_{\alpha+p})} \left( 1 + \|X\|^\gamma_{\tilde{B}_{r,\gamma}(\mathcal{H}_{\alpha+p})} \right) |t-s|^{-\delta + \alpha \gamma - \alpha}, \]

as required.

These estimates apply to the solution of the slow SPDE (1.3). We record this in the following proposition:

**Proposition 4.10.** Let \( T > 0 \), \( H > \frac{1}{2} \), and \( p \geq 2 \). Assume the conditions of Proposition 2.6. Then, for every initial condition \( X_0 \in \mathcal{H} \) and every \( T > 0 \), there exists a solution to \( dX_t = (A + f(t, X_t)) \, dt + g(t, X_t) \, dB_t \) in \( \tilde{B}_{\alpha,p}([0,T],\mathcal{H}) \).

This is an immediate consequence of the deterministic a priori bound (2.15) and the fact that, for each \( \gamma < H \), \( \| B \|_{C^\gamma([0,T],\mathcal{K})} \) has moments of all order by virtue of Fernique’s theorem. It is a good place to note that SPDEs driven by fBMs have been widely studied, from the perspective of rough paths, white noise analysis, Caputo-type derivatives, Malliavin calculus, and random dynamical systems, see [30, 16, 7, 11, 38, 26, 8, 28, 27, 5]. But our aim is to obtain uniform bounds in terms of the rather weak \( \| \cdot \|_{\mathcal{A}_{\delta,\gamma}^p} \) -norm of the coefficients. As an application, we prove the convergence of the slow motion \( X^\varepsilon \) (1.3) to the effective equation (1.4).

**Proposition 4.11.** Let \( \gamma \in (0,1) \), \( \delta \in (0, H - \frac{1}{2}) \), and \( \alpha < \frac{H - \delta}{\gamma} \) satisfying \( H - \delta + \alpha \gamma > 1 \). Let \( f : \Omega \times [0,T] \times \mathcal{H} \rightarrow \mathcal{H} \) and \( g : \Omega \times [0,T] \times \mathcal{H} \rightarrow L(\mathcal{K}, \mathcal{H}) \) be independent of \( \mathcal{F}^B \) with

\[ \| f \|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{H})} < \infty, \quad \| g \|_{\mathcal{A}_{\delta,\gamma}^p(\mathcal{L}(\mathcal{K},\mathcal{H}))} < \infty. \]
Then the mild solution to the equation
\[ dX_t = AX_t dt + f(t, X_t) dt + g(t, X_t) dB_t \]
satisfies the a priori bound
\[ \|X\|_{\tilde{B}_{\alpha,p}(\mathcal{H})} \lesssim \left( 1 + \|X\|_{\tilde{B}_{\alpha,p}(\mathcal{H})}^2 \right) \left( \|f\|_{\mathcal{A}^{\alpha}_{\delta-\gamma}(\mathcal{H})} + \|g\|_{\mathcal{A}^{\gamma}_{\delta-\gamma}(L(\mathcal{K}, \mathcal{H}))} \right) |t-s|^{H-\delta-\alpha(2-\gamma)}. \] (4.13)

**Proof.** We may apply the \( L^p \)-estimates obtained earlier in Proposition 4.8 and in Lemma 4.9 to the integrals in the identity
\[ \delta X_{s,t} = \int_s^t S_{t-r} f(r, X_r) dr + \int_s^t S_{t-r} g(r, X_r) dB_r. \]
In (4.10) and (4.12) take \( \kappa = -\alpha \), then
\[ \|\delta X_{s,t}\|_{L^p} \lesssim \left( 1 + \|X\|_{\tilde{B}_{\alpha,p}(\mathcal{H})}^2 \right) \left( \|f\|_{\mathcal{A}^{\alpha}_{\delta-\gamma}(\mathcal{H})} |t-s|^{1-\delta+\alpha\gamma-\alpha} + \|g\|_{\mathcal{A}^{\gamma}_{\delta-\gamma}(L(\mathcal{K}, \mathcal{H}))} |t-s|^{H-\delta+\alpha\gamma-\alpha} \right) \]
completing the proof. \( \square \)

Proposition 4.8 and Lemma 4.9 yield the following probabilistic stability result for the SPDE (4.1):

**Corollary 4.12.** Let \( H > \frac{1}{2} \). Let \( \delta \geq 0, \gamma < 1, \) and \( \alpha > 0 \). Suppose that
\[ H - \delta > \frac{1}{2}, \quad \alpha < \frac{H - \delta}{2 - \gamma}, \quad H - \delta + \alpha\gamma > 1. \]

Let \( f^n, f : \Omega \times [0, T] \times \mathcal{H} \to \mathcal{H} \) and \( g^n, g : \Omega \times [0, T] \times \mathcal{H} \to L(\mathcal{K}, \mathcal{H}) \) be independent of \( \mathcal{F}_t \) and satisfy the conditions of Proposition 2.6. Then the SPDEs
\[ \begin{align*}
   dX^n_t &= (A + f^n(t, X^n_t)) dt + g^n(t, X^n_t) dB_t, \\
   dX_t &= (A + f(t, X_t)) dt + g(t, X_t) dB_t
\end{align*} \] (4.14) (4.15)
have a unique pathwise mild solution, see (1.2). Moreover, if \( X^n_0 = X_0 \) for each \( n \in \mathbb{N} \) and
\[ \|f^n - f\|_{\mathcal{A}^{\alpha}_{\delta-\gamma}(\mathcal{H})} \to 0, \quad \|g^n - g\|_{\mathcal{A}^{\gamma}_{\delta-\gamma}(L(\mathcal{K}, \mathcal{H}))} \to 0, \]
then \( X^n \to \tilde{X} \) in \( \tilde{C}^{(\alpha-\frac{1}{2})}([0, T], \mathcal{H}) \) in probability as \( n \to \infty. \)

**Proof.** First note that the equations (4.14)–(4.15) are well-posed in \( \tilde{C}^{H-\gamma}([0, T], \mathcal{H}) \) by Proposition 2.6. Proposition 4.10 shows that \( X^n \in \tilde{B}_{\alpha,p}([0, T], \mathcal{H}) \) and Proposition 4.11 furnishes the estimate
\[ \|X^n\|_{\tilde{B}_{\alpha,p}(\mathcal{H})} \lesssim \left( 1 + \|X^n\|_{\tilde{B}_{\alpha,p}(\mathcal{H})}^2 \right) \left( \|f^n\|_{\mathcal{A}^{\alpha}_{\delta-\gamma}(\mathcal{H})} + \|g^n\|_{\mathcal{A}^{\gamma}_{\delta-\gamma}(L(\mathcal{K}, \mathcal{H}))} \right) |t-s|^{H-\delta-(2-\gamma)\alpha}. \]
Since the exponent \( H - \delta - (2 - \gamma)\alpha > 0 \) by assumption, we can iterate this bound to see that \( \sup_n \|X^n\|_{\tilde{B}_{\alpha,p}(\mathcal{H})} < \infty. \) Next, we apply Lemma 2.7 with \( y^n_t = X_0 \) and
\[ \bar{y}_t^n \triangleq X_0 + \int_0^t S_{t-r} (f^n(r, X^n_r) - f(r, X^n_r)) dr + \int_0^t S_{t-r} (g^n(r, X^n_r) - g(r, X^n_r)) dB_r, \]
so that

\[
X^n_t = y^n_t + \int_0^t S_{t-r} f(r, X^n_r) \, dr + \int_0^t S_{t-r} g(r, X^n_r) \, dB_r, \\
X_t = y^n_t + \int_0^t S_{t-r} f(r, X_r) \, dr + \int_0^t S_{t-r} g(r, X_r) \, dB_r.
\]

Proposition 4.11 shows that

\[
\|y^n - \bar{y}^n\|_{\mathcal{B}_{a,p}(\mathcal{H})} \lesssim \left(1 + \|X^n\|_{\mathcal{H}}^\gamma\right) \left(\|f^n - f\|_{\mathcal{A}_{\delta,\gamma}(\mathcal{H})} + \|g^n - g\|_{\mathcal{A}_{\delta,\gamma}(L(\mathcal{K},\mathcal{H}))}\right),
\]

whence the claim follows by (2.25) and the mild Kolmogorov continuity theorem (Proposition 2.1).

\[\square\]

5 Non-Stationary Averaging

Let \(\mu, \nu \in \mathcal{P}(\mathcal{X})\). The total variation distance between \(\mu\) and \(\nu\) is defined by

\[
\|\mu - \nu\|_{TV} \triangleq \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)| = \min_{(X,Y) \in \mathcal{C}(\mu,\nu)} \mathbb{P}(X \neq Y),
\]

where \(\mathcal{C}(\mu,\nu) \triangleq \{\rho \in \mathcal{P}(\mathcal{X} \times \mathcal{X}) : [\rho]_1 = \mu, [\rho]_2 = \nu\}\) is the set of couplings.

5.1 Quantitative Ergodic Theorem

In order to prove Theorem 6.3, we present a strengthened version of [24, Lemma 3.14]. The proof there strongly relies on Kolmogorov's continuity theorem. To the best of our knowledge, such a result is not known for random fields indexed by elements of a general infinite-dimensional Hilbert space. Here we follow a different strategy, which also allows us to extend the result to non-stationary fast processes.

Let \(Y\) be a stochastic process on a complete, separable metric space \(\mathcal{Y}\) and \(\pi \in \mathcal{P}(\mathcal{Y})\) be a probability measure. Let \(f : \mathcal{H} \times \mathcal{Y} \to \mathcal{H}\). Let

\[
F_e(t, x) = f(x, Y_t) \quad \text{and} \quad \bar{f}(x) \triangleq \int_{\mathcal{Y}} f(x, y)\pi(dy).
\]

We also set \(\mathcal{F}^Y_t = \sigma(Y_s, s \leq t)\).

**Definition 5.1.** We say that the ergodic condition with rate \(a > 0\) and exponent \(\delta \in (0, 1)\) holds for \(f\) (w.r.t. \(Y\) and \(\pi\)) if for any \(\mathcal{H}\)-valued random variables \(X, Y\),

\[
\sup_{s \leq T} \sup_{X, Z \in \mathcal{F}^Y_s \vee \mathcal{F}^Y_t} \frac{\|F_e(\cdot, X) - F_e(\cdot, Z) - (\bar{f}(X) - \bar{f}(Z))\|_{\mathcal{L}^{1-\delta}(\mathcal{F}^Y_t)}}{\|X - Z\|_{\mathcal{L}^p}} \lesssim \varepsilon^a,
\]

\[
\sup_{s \leq T} \sup_{X \in \mathcal{F}^Y_s \vee \mathcal{F}^Y_t} \frac{\|F_e(\cdot, X) - \bar{f}(X)\|_{\mathcal{L}^{1-\delta}(\mathcal{F}^Y_t)}}{\|X\|_{\mathcal{L}^p}} \lesssim \varepsilon^a.
\]

The constants are allowed to depend on \(f\), but not on \(\varepsilon > 0\). In particular,

\[
\|F_e - \bar{f}\|_{\mathcal{A}_{\delta,\gamma}} \lesssim \varepsilon^a
\]

for the norm (4.9).
We let \( \mathcal{L}(Y_t \mid Y_s) \) denote the conditional law of \( Y_t \) given \( Y_s \). Let \( B_R \) denote the closed ball in \( \mathcal{H} \) of radius \( R \) centered at 0. For simplicity, we formulate the next lemma only for Markovian fast processes. However, we shall show in § 5.2.1 that its conclusion also holds for some non-Markovian process \( Y \).

**Lemma 5.2.** Let \((Y_t)_{t \geq 0}\) be a Markov process on \( \mathcal{Y} \). Let \( \pi \) be a probability measure on \( \mathcal{Y} \). Assume that there is a number \( \rho \in (0,1) \) such that, for all \( 0 \leq s \leq t \),

\[
\mathbb{E} \left[ \left\| \mathcal{L}(Y_t \mid Y_s) - \pi \right\|_{TV} \right] \lesssim \frac{1}{(t-s)^\rho}.
\]  

(5.1)

Let \( f : \mathcal{H} \times \mathcal{Y} \to \mathcal{H} \) and \( g : \mathcal{H} \times \mathcal{Y} \to L(\mathcal{K}, \mathcal{H}) \) be bounded measurable functions for which there is an \( L > 0 \) such that for all \( x, z \in \mathcal{H} \) and all \( y \in \mathcal{Y} \):

\[
\left\| f(x,y) - f(z,y) \right\|_{\mathcal{H}} \leq L \| x - z \|_{\mathcal{H}}, \quad \left\| g(x,y) - g(z,y) \right\|_{L(\mathcal{K}, \mathcal{H})} \leq L \| x - z \|_{\mathcal{H}}.
\]  

(5.2)

Fix \( \delta \in (0,1) \) and \( \gamma \in (0,1) \). Then, for every \( p > \frac{1+\rho}{\rho} \) and exponent \( \delta \) holds for both \( f \) and \( g \).

**Proof.** We only prove the statement for \( g \); for \( f \) one proceeds *mutatis mutandis*. Since \( \bar{g} \) is bounded and Lipschitz continuous, we can subtract it from \( g \) so that we may assume \( \bar{g} \equiv 0 \) without any loss of generality. Let \( e \in \mathcal{K} \) be a unit vector.

Let \( s \in [0,T] \) and fix random variables \( X, Z \in \mathcal{F}_s^B \lor \mathcal{F}_s^Y \). For any \( x, z \in \mathcal{H} \), we abbreviate

\[
\Theta^e_{s,t}(x,z) = \int_s^t \left( G(e(u,x))e - G(e(u,z))e \right) du,
\]

which is a random variable with values in \( \mathcal{H} \), and observe the trivial bound

\[
\left\| \Theta^e_{s,t}(x,z) \right\|_{L^\infty(\Omega)}^{p-2} = \left\| \int_s^t \left( G(e(u,x))e - G(e(u,z))e \right) du \right\|_{L^\infty(\Omega)}^{p-2} \lesssim \| x - z \|^{p-2} |t-s|^{p-2}.
\]

By Markovity, \( Y^x_{s,z} \) is conditionally independent of \( \mathcal{F}_s^B \lor \mathcal{F}_s^Y \) given \( Y_{s,z} \), so we have

\[
\mathbb{E} \left[ \left\| \int_s^t \left( G(e(u,x))e - G(e(u,z))e \right) du \right\|_{\mathcal{H}}^p \right] = \mathbb{E} \left[ \mathbb{E} \left[ \left\| \Theta^e_{s,t}(x,z) \right\|_{\mathcal{H}}^p \mid Y_{s,z} \right] \right]_{x=z=} \mathbb{E} \left[ \left\| \Theta^e_{s,t}(x,z) \right\|_{L^\infty(\Omega)}^{p-2} \right]_{x=z=}.
\]

(5.3)

The first factor can be expanded as

\[
\mathbb{E} \left[ \left\| \Theta^e_{s,t}(x,z) \right\|_{\mathcal{H}}^2 \mid Y_{s,z} \right] = \mathbb{E} \left[ \left\| \int_s^t \left( G(e(u,x))e - G(e(u,z))e \right) du \right\|_{\mathcal{H}}^2 \right]_{Y_{s,z}^z} = 2 \int_s^s \int_s^t \mathbb{E} \left[ \left\langle G(x_{s,t})e - G(z_{s,t})e, \mathbb{E} \left[ G(x_{s,t})e - G(z_{s,t})e \mid Y_{s,z} \right] \right\rangle \right]_{X,z=} \, du \, dv.
\]

We notice that, by (5.1)–(5.2) and the assumption \( \bar{g} \equiv 0 \),

\[
\left\| \mathbb{E} \left[ G(x_{s,t})e - G(z_{s,t})e \mid Y_{s,z} \right] \right\| \lesssim \| x - z \| \left( \frac{\epsilon}{v-u} \right)^\rho.
\]

\[
\lesssim \| x - z \| \left( \frac{\epsilon}{v-u} \right)^\rho.
\]

28
Hence by Cauchy-Schwarz,

\[
\mathbb{E} \left[ \left\| \mathfrak{G}^\varepsilon_{s,t}(x,z) \right\|_{L^p}^2 \right] \lesssim \|x-z\|^2 \int_s^t \int_s^v \left( \frac{\varepsilon}{v-u} \right)^p \, du \, dv \lesssim \varepsilon^p \|x-z\|^2 |t-s|^{2-p}.
\]

in view of (5.2). Inserting this back into (5.3), we have shown that

\[
\left\| \int_s^t G_\varepsilon(u,X)e - G_\varepsilon(u,Z)e \, du \right\|_{L^p} \lesssim \varepsilon^p |t-s|^{1-p} \|X - Z\|_{L^p}.
\]

Since this bound is uniform in both \(0 \leq s < t \leq T\) and \(e \in \{x \in K : \|x\|_K = 1\}\), Kolmogorov’s continuity theorem implies

\[
\left\| G_\varepsilon(\cdot,X) - G_\varepsilon(\cdot,Z) - (\bar{g}(X) - \bar{g}(Z)) \right\|_{C^{-\delta}(L(K,H))} \lesssim \varepsilon^p \|X - Z\|_{L^p},
\]

provided that we choose \(p \geq \frac{1+\rho}{\delta}\).

We also have

\[
\left\| \mathbb{E} \left[ g(x,Y^\varepsilon_{s,t}) \mid Y^\varepsilon_s \right] \right\| \lesssim \|g(x,\cdot)e\|_{L^p} \|\mathcal{L}(Y^\varepsilon_s) - \pi\|_{TV} \lesssim \|g\|_\infty \left( \frac{\varepsilon}{v-u} \right)^p.
\]

Hence replacing \(G_\varepsilon(u,x)e - G_\varepsilon(u,z)e\) by \(G_\varepsilon(u,x)e\) in the definition of \(\mathfrak{G}^\varepsilon_{s,t}(x,z)\) leads to

\[
\left\| G_\varepsilon(\cdot,X) - \bar{g}(X) \right\|_{C^{-\delta}(L(K,H))} \lesssim \varepsilon^p.
\]

This completes the proof. \(\square\)

### 5.2 Examples

A class of fast process falling in the regime of Lemma 5.2 are solutions of S(P)DEs driven by another fractional Brownian motion. Establishing the convergence of the conditional law is much more delicate in this case and there not many results known. For a (time-inhomogeneous) Markov process with transition kernel \((P_{s,t})_{0 \leq s \leq t}\), the assumption (5.1) is implied by \(\|P_{s,t}(y,\cdot) - \pi\|_{TV} \leq \frac{1}{C(\varepsilon)}\) for all \(0 \leq s \leq t\) and \(y \in \mathcal{Y}\), provided that \(\sup_{s \geq 0} \mathbb{E}[C(Y_s)] < \infty\). Such a bound is well known to hold for solutions of SDEs satisfying strong Hörmander’s conditions on compact manifolds. There are of course many examples of random dynamical systems with rate of convergence to invariant measure at least polynomial and for which the above bound holds, we shall focus on the less exploited non-Markovian dynamics.

#### 5.2.1 A Non-Markovian Example

The following *quenched ergodic theorem* furnishes non-Markovian examples falling in the regime of the averaging principle in the next section: Consider the SDE

\[
dY_t = b(Y_t) \, dt + \sigma \, dB_t,
\]

where \(\sigma \in L(\mathbb{R}^n, \mathbb{R}^n)\) is invertible and \(\hat{B}\) is an \(n\)-dimensional fBm independent of \(B\).

To make sense of an invariant measure for the equation (5.4), we course cannot resort to the notion of an invariant measure for a Markov process. A natural replacement can be found in the theory of random dynamical systems. However, the latter invariant measures are not guaranteed to be physically meaningful since the solution may ‘look’ into the future. Hence, Hairer [23] introduced the notion of a ‘physical invariant measure’ for which a Markov process is defined by
including the history of the process—instead of the whole two-sided noise process as used in the theory of random dynamical system. The history space \( \mathcal{H}_H \) is a separable Banach space defined as the closure of the space

\[
C_0^\infty ((-\infty,0], \mathbb{R}^n) \triangleq \{ w : (-\infty,0] \to \mathbb{R}^n \text{ smooth with compact support and } w(0) = 0 \},
\]

under the norm \( \| w \|_H \triangleq \sup_{s,t \leq 0} \frac{|w(t) - w(s)|}{(1 + |t| + |s|)^\frac{1}{2}} \). This space supports the Wiener measure. The construction of the invariant measure then comes down to constructing a family of Feller transition probabilities \( (Q_t)_{t \in [0,T]} \) on \( \mathbb{R}^n \times \mathcal{H}_H \). Then let \( (Y_t^T, Z_t^T)_{t \in [0,T]} \) be a Markov process on \( \mathbb{R}^n \times \mathcal{H}_H \) with transition probabilities \( (Q_t) \) and started from an invariant measure whose projection to \( \mathcal{H}_H \) is the Wiener measure. One can then take \( T \to \infty \) to obtain a weak limit \( (Y_t, Z_t)_{t \geq 0} \). The first marginal of this Feller process solves (5.4) and is stationary. Its time 0 marginal \( \pi \in \mathcal{P}(\mathbb{R}^n) \) is used in the next proposition:

**Proposition 5.3.** Suppose that \( b : \mathbb{R}^n \to \mathbb{R}^n \) is globally Lipschitz continuous and there are \( \lambda, \kappa, R > 0 \) such that

\[
(b(x) - b(y), x - y) \leq \begin{cases} -\kappa|x - y|^2, & |x|, |y| > R, \\ \lambda|x - y|^2, & \text{otherwise.} \end{cases}
\]

Then there is a \( \Lambda = \Lambda(\kappa, R) > 0 \) such that, provided \( \lambda \leq \Lambda \), there exists a physical invariant probability measure \( \pi \in \mathcal{P}(\mathbb{R}^n) \) to (5.4) and, for any initial condition \( Y_0 \in L^1 \), the conclusion of Lemma 5.2 holds:

Let \( f : \mathcal{H} \times \mathbb{R}^n \to \mathcal{H} \) and \( g : \mathcal{H} \times \mathbb{R}^n \to L(K, \mathcal{H}) \) be bounded measurable functions for which there is an \( L > 0 \) such that for all \( x, z \in \mathcal{H} \) and all \( y \in \mathbb{R}^n \):

\[
\| f(x, y) - f(z, y) \|_\mathcal{H} \leq L \| x - z \|_\mathcal{H}, \quad \| g(x, y) - g(z, y) \|_{L(K,\mathcal{H})} \leq L \| x - z \|_\mathcal{H}.
\]

Fix \( \delta \in (0,1) \) and \( \gamma \in (0,1) \). Then, for every \( p > \frac{1 + 2}{\kappa} \vee 2 \), there is a \( c > 0 \) such that the ergodic condition with exponential rate \( e^{-\xi t} \) and exponent \( \delta \) holds for both \( f \) and \( g \).

**Proof.** The only difference to Lemma 5.2 lies in the conditioning step (5.3): Let \( \Phi_{s,t}(y, \hat{B}) \) denote the flow to (5.4) interpreted in the pathwise sense. We can decompose \( \hat{B} = \hat{B}_s^\pm + \hat{B}_t^\pm \) as in (4.4). Since \( \hat{B}_t^\pm \) is independent of \( \mathcal{F}_s^B \vee \mathcal{F}_t^\mathcal{B}_s \), (5.3) becomes

\[
\mathbb{E} \left[ \left\| \int_s^t \left( G_{e}(u, X) e - G_{e}(u, Z)e \right) \, du \right\|_{\mathcal{H}}^p \right] \leq \mathbb{E} \left[ \left\| \Phi_{s,t}^e(x, z) \right\|_{\mathcal{H}}^p \right] \left\| \Phi_{s,t}^\mathcal{B}(x, z) \right\|_{L^\infty(\Omega)}^{p-2} e^{-c(t-s)},
\]

whence we can conclude the proof along the same lines as in Lemma 5.2.

**5.2.2 The Stationary Process Case**

Let us draw a connection of our results to [24]. There, the authors considered a *finite-dimensional* SDE and proved a similar averaging principle for time \( t \) mixing, stationary fast processes. This is to say that Rosenblatt’s mixing coefficient

\[
\alpha(t) = \alpha(Y_0, Y_t) = \sup_{A \in \sigma(Y_0)} \left| \mathbb{P} (A \cap B) - \mathbb{P} (A) \mathbb{P} (B) \right|
\]

is used in the next proposition:
vanishes with an algebraic rate as \( t \to \infty \). Let \( \pi \in \mathcal{P}(\mathcal{Y}) \) be the stationary measure of \( Y \). If 
\[
\| \mathcal{L}(Y_t \mid Y_0 = y) - \pi \|_{TV} \leq C(y)\beta(t) \quad \text{for } \pi\text{-a.e. } y \in \mathcal{Y},
\]
then we have that \( \alpha(t) \leq \| C \|_{L^1(\mathcal{Y}, \pi)} \beta(t) \). Indeed, fix \( A, B \in \mathcal{B}(\mathcal{Y}) \) and \( t \geq 0 \). On disintegration we have
\[
|\mathbb{P}(Y_0 \in A, Y_t \in B) - \mathbb{P}(Y_0 \in A)\mathbb{P}(Y_t \in B)| \leq \int_A \left(C(Y_t \mid Y_0 = y)(B) - \pi(B)\right) \pi(dy) \leq \beta(t) \int_{\mathcal{Y}} C(y) \pi(dy) = \| C \|_{L^1(\mathcal{Y}, \pi)} \beta(t).
\]

We emphasize that we cannot cover such a general class of fast processes. This is due to the failure of Kolmogorov’s continuity theorem for random fields indexed by a general infinite-dimensional Hilbert space, which made the more challenging conditioning procedure necessary. In particular, the identity (5.3) may not hold for general \( Y \). However, the classes of fast processes presented in § 5.2.1 cover a very rich selection of practically relevant examples.

6 Application to Slow-Fast Systems

Recall that \( (Y_t)_{t \geq 0} \) is a sample continuous process on a separable complete metric space \((\mathcal{Y}, d_\mathcal{Y})\), which is independent of the fBm \( B \). We assume that the ergodic condition introduced in Definition 5.1 holds. Examples of \( Y \) satisfying this assumption are given in the last section.

We can now prove a fractional averaging principle for the two-scale dynamics (1.3). To this end, let us first state the regularity assumption on the coefficients:

**Condition 6.1.** The vector fields \( f : \mathcal{H} \times \mathcal{Y} \to \mathcal{H}, g : \mathcal{H} \times \mathcal{Y} \to L(\mathcal{K}, \mathcal{H}) \) are globally Lipschitz continuous. Moreover, for each \( \alpha < 1 - H \),
\[
\sup_{y \in \mathcal{Y}} \| g(x, y) - g(z, y) \|_{L(\mathcal{K}, \mathcal{H}^{-\alpha})} \lesssim \| x - z \|_{\mathcal{H}^{-\alpha}} \quad \forall x, z \in \mathcal{H},
\]
\[
\sup_{y \in \mathcal{Y}} \| D_x g(x, y) - D_x g(z, y) \|_{L(\mathcal{K}, L(\mathcal{H}^{-\alpha}))} \lesssim \| x - z \|_{\mathcal{H}^{-\alpha}} \quad \forall x, z \in \mathcal{H},
\]
\[
\sup_{x \in \mathcal{H}} \| D_x g(x, y) - D_x g(x, \bar{y}) \|_{L(\mathcal{K}, L(\mathcal{H}^{-\alpha}))} \lesssim d_{\mathcal{Y}}(y, \bar{y}) \quad \forall y, \bar{y} \in \mathcal{Y},
\]
\[
\sup_{x \in \mathcal{H}} \| g(x, y) - g(x, \bar{y}) \|_{L(\mathcal{K}, \mathcal{H}^{-\alpha})} \lesssim d_{\mathcal{Y}}(y, \bar{y}) \quad \forall y, \bar{y} \in \mathcal{Y}.
\]

**Remark 6.2.** Let \( f, g \) satisfy Condition 6.1 and define
\[
f_\varepsilon(t, x) = f(x, Y_\varepsilon), \quad g_\varepsilon(t, x) = g(x, Y_\varepsilon).
\]
Then \( f_\varepsilon, g_\varepsilon \) satisfies the conditions of Proposition 2.6 and therefore for each \( \varepsilon > 0 \), the SPDE Equation (1.3) has a unique solution in \( \hat{B}_{\alpha, p}([0, T], \mathcal{H}) \) for any \( T > 0 \), any \( p \geq 2 \), any \( \alpha < H \), and any \( X_0 \in \mathcal{H} \), see Proposition 4.10.

**Theorem 6.3.** Let \( Y \) be a stochastic process with almost surely locally Hölder continuous sample paths of order greater than \( 1 - H \). Suppose that either \( Y \) falls in the regime of Lemma 5.2 or Proposition 5.3. If \( f \) and \( g \) satisfy Condition 6.1 and \( X_0 \in \bigcap_{p \geq 1} L^p(\Omega, \mathcal{H}) \), then the following hold for any \( \alpha < H \):
\begin{enumerate}
\item [(i)] As \( \varepsilon \to 0 \), the unique solution \( X^\varepsilon \in \hat{C}^\alpha([0, T], \mathcal{H}) \) to (1.3) converges in probability in \( \hat{C}^\alpha([0, T], \mathcal{H}) \).
\item [(ii)] The limit is the unique pathwise solution to (1.4).
\end{enumerate}
Proof. Fix $\alpha \in \left(\frac{1}{2}, H\right)$. Let $R > 0$ and let $\varrho_R \in C_0^\infty(\mathcal{H})$ be a smooth bump function with value 1 on $\bar{B}_R$ and which vanishes on $\mathcal{H} \setminus \bar{B}_{2R}$. Since the map $x \mapsto \|x\|^2$ is $C^\infty$ in the Fréchet sense, such a function can be constructed. We also define the stopping time
\[
\tau^R_R = \inf\{t \geq 0 : \|XX_t\| \vee \|\bar{X}\| > R\}.
\]
Set $f_R(t, x) = \varrho_R(x)f(t, x)$ and similar for $g_R$. Notice that $f_R$ and $g_R$ are bounded by the Lipschitz assumption. Let
\[
dX_t^R = (AX_t^R + f_R(X_t^R, Y_t^R)) \, dt + g_R(X_t^R, Y_t^R) \, dB_t
\]
and
\[
d\bar{X}_t^R = (A\bar{X}_t^R + \bar{f}_R(\bar{X}_t^R)) \, dt + \bar{g}_R(\bar{X}_t^R) \, dB_t.
\]
Note that both of these equations are well posed in $\hat{C}^\alpha([0, T], \mathcal{H})$ by Proposition 2.6.

Fix $\delta > 0$ and $\gamma < 1$ such that $H - \delta + \alpha \gamma > 1$ and $\alpha < \frac{H - \delta}{2 - \gamma}$. Let $c_1 \in (0, 1)$ be given. Since $\|\tilde{f}_{0,t}\| \geq ||f_t|| - ||S_t f_0||$, we can estimate
\[
\mathbb{P}\left(\|X^\varepsilon\|_{C([0, \tau^R_R], \mathcal{H})} \geq R\right) \leq \mathbb{P}\left(\|X^\varepsilon\|_{\hat{C}^\alpha([0, \tau^R_R], \mathcal{H})} \geq \frac{R - \|S_t X_0\|_{\mathcal{H}}}{T^\alpha}\right)
\leq \mathbb{P}\left(\|X^\varepsilon\|_{\hat{C}^\alpha([0, \tau^R_R], \mathcal{H})} \geq \frac{R - M}{T^\alpha}\right) + \mathbb{P}\left(\|S_t X_0\|_{\mathcal{H}} \geq M\right)
\leq \mathbb{P}\left(\|X^\varepsilon, R - \bar{X}^R\|_{\hat{C}^\alpha([0, T], \mathcal{H})} \geq c_1\right) + 2\mathbb{P}\left(\|\bar{X}\|_{\hat{C}^\alpha([0, T], \mathcal{H})} \geq \frac{R - M}{T^\alpha} - c_1\right)
+ \mathbb{P}\left(\|S_t X_0\|_{\mathcal{H}} \geq M\right).
\]

Hence, we have
\[
\mathbb{P}(\tau^R_R < T) \leq \mathbb{P}\left(\|X^\varepsilon\|_{C([0, \tau^R_R], \mathcal{H})} \geq R\right) \leq \mathbb{P}\left(\|X^\varepsilon\|_{C([0, \tau^R_R], \mathcal{H})} \geq R\right)
\leq \mathbb{P}\left(\|X^\varepsilon, R - \bar{X}^R\|_{\hat{C}^\alpha([0, T], \mathcal{H})} \geq c_1\right) + 2\mathbb{P}\left(\|\bar{X}\|_{\hat{C}^\alpha([0, T], \mathcal{H})} \geq \frac{R - M}{T^\alpha} - c_1\right)
+ 2\mathbb{P}\left(\|S_t X_0\|_{\mathcal{H}} \geq M\right).
\]

Let $c_2 > 0$. Since $X_0 < \infty$ with probability 1, we can first choose $M_0 > 0$ such that
\[
\mathbb{P}\left(\|S_t X_0\|_{\mathcal{H}} \geq M_0\right) \leq \frac{C_2}{8},
\]
Since $\bar{f}, \bar{g}$ satisfies conditions of Proposition 2.6, Proposition 4.11 applies yielding: For each $p \geq 1$, $\bar{X} \in L^p(\Omega, \hat{C}^\alpha([0, T], \mathcal{H})) \subset \hat{B}_{\alpha,p}$. Hence, pick $R_0 > 0$ such that
\[
\mathbb{P}\left(\|\bar{X}\|_{\hat{C}^\alpha([0, T], \mathcal{H})} \geq \frac{R_0 - M_0}{T^\alpha} - c_1\right) \leq \frac{C_2}{8}.
\]

Then we have that
\[
\mathbb{P}\left(\|X^\varepsilon - \bar{X}\|_{\hat{C}^\alpha([0, T], \mathcal{H})} > c_1\right) \leq \mathbb{P}(\tau^R_R < T) + \mathbb{P}\left(\|X^\varepsilon, R_0 - \bar{X}^R\|_{\hat{C}^\alpha([0, T], \mathcal{H})} > c_1\right). \quad (6.1)
\]
Using that supp $(f_{R_0}(t, \cdot)) \subset \bar{B}_{R_0}$ for each $t \in [0, T]$, we have
\[
\|f_{R_0} - \bar{f}_{R_0}\|_{A^p_{\delta, \gamma}} \leq \|f_{R_0} - \bar{f}_{R_0}\|_{A^p_{\delta, 1}}
\]
32
and similar for $g_{R_0}$. By Lemma 5.2 (resp. Proposition 5.3), it holds that

$$
\| f_{R_0} - \tilde{f}_{R_0} \|_{A^p_{\delta,1}} \leq \varepsilon^p, \quad \| g_{R_0} - \tilde{g}_{R_0} \|_{A^p_{\delta,1}} \leq \varepsilon^p.
$$

Hence by Corollary 4.12 on the stability of solutions, we can choose $\varepsilon_0 > 0$ so that, for any $\varepsilon < \varepsilon_0$,

$$
P \left( \| X^{\varepsilon, R_0} - \tilde{X}^{R_0} \|_{C^0([0,T],\mathcal{H})} > c_1 \right) \leq \frac{c_2}{4}.
$$

Coming back to (6.1), we have shown that, for any $\varepsilon \in (0, \varepsilon_0)$,

$$
P \left( \| X^{\varepsilon} - \tilde{X} \|_{C^0([0,T],\mathcal{H})} > c_1 \right) \leq 2P \left( \| X^{\varepsilon, R_0} - \tilde{X}^{R_0} \|_{C^0([0,T],\mathcal{H})} > c_1 \right) + \frac{c_2}{2} \leq c_2
$$
as required. \hfill \Box

For slow-fast SPDEs driven by Wiener processes where $g(x, y) = g(y)$, it is known that $X^{\varepsilon} \rightarrow \tilde{X}$ in $L^p(\Omega, C([0,T],\mathcal{H}))$ [34]. This is a common feature with SDEs. To obtain such strong convergence for Markovian systems, the requirement $g(x, y) = g(x)$ is necessary even for SDEs as the following very simple example shows:

**Example 6.4.** Let $\tilde{W}$ and $W$ be independent one-dimensional Wiener processes. Consider $dX_t^\varepsilon = \cos(Y_t^\varepsilon) d\tilde{W}_t$ and $dY_t = -Y_t dt + \sqrt{2} dW_t$ with $Y_0 \sim N(0, 1)$ and $X_0^\varepsilon = X_0$. Let $\eta \triangleq \sqrt{\int_{\mathbb{R}} \cos^2(y) \pi_0(dy)} = \sqrt{\frac{1}{2} (1 + e^{-2})}$. Then $X^{\varepsilon} \Rightarrow \tilde{X} = \eta \tilde{W}$ weakly in $C([0,T],\mathbb{R})$, but

$$
\| X_t^\varepsilon - \tilde{X}_t \|_{L^2}^2 = \int_0^t \mathbb{E} \left[ (\cos(Y_s^\varepsilon) - \eta)^2 \right] ds = \left( 1 + \frac{1 - \sqrt{2(e + e^3)}}{e^2} \right) t \not\rightarrow 0.
$$

In our case, the convergence is in probability, even though $g$ depend on both the slow and the fast variable, this is a feature of fractional averaging.

**References**


[34] Wei Liu, Michael Röckner, Xiaobin Sun, and Yingchao Xie. Strong averaging principle for slow-fast stochastic partial differential equations with locally monotone coefficients, 2019.


