Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise

Aurélien Deya, Fabien Panloup, Samy Tindel

Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise

Aurélien Deya, *Fabien Panloup†, Samy Tindel‡

April 29, 2016

Abstract

We investigate the problem of the rate of convergence to equilibrium for ergodic stochastic differential equations driven by fractional Brownian motion with Hurst parameter $H \in (1/3, 1)$ and multiplicative noise component σ. When σ is constant and for every $H \in (0, 1)$, it was proved in [17] that, under some mean-reverting assumptions, such a process converges to its equilibrium at a rate of order $t^{-\alpha}$ where $\alpha \in (0, 1)$ (depending on H). In [9], this result has been extended to the multiplicative case when $H > 1/2$. In this paper, we obtain these types of results in the rough setting $H \in (1/3, 1/2)$. Once again, we retrieve the rate orders of the additive setting. Our methods also extend the multiplicative results of [9] by deleting the gradient assumption on the noise coefficient σ. The main theorems include some existence and uniqueness results for the invariant distribution.

Keywords: Stochastic Differential Equations; Fractional Brownian Motion; Multiplicative noise; Ergodicity; Rate of convergence to equilibrium; Lyapunov function; Total variation distance.

AMS classification (2010): 60G22, 37A25.

1 Introduction

Convergence to an equilibrium distribution is one of the most natural and most studied problems concerning Markov processes. This holds true in particular for diffusions processes, seen as solutions to stochastic differential equations (SDEs in the sequel) driven by a Brownian motion. More specifically, consider the $\mathbb{R}^d$-valued process $(Y_t)_{t\geqslant 0}$ solving the following SDE:

$$dY_t = b(Y_t)dt + \sigma(Y_t)\, dW_t \tag{1.1}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{M}_{d,d}$ are smooth enough functions, where $\mathbb{M}_{d,d}$ is the set of $d \times d$ real matrices, and where W is a d-dimensional Wiener process. Assume for simplicity that $\sigma(x)$ is invertible for every $x \in \mathbb{R}^d$ and that σ^{-1} is a bounded function.

In the context of equation (1.1), a simple assumption which ensures ergodicity of the process Y is the following reinforcing condition on the drift b (see Hypothesis (H2) below for further details): There exist $C_1, C_2 > 0$ such that for every $v \in \mathbb{R}^d$, one has

$$\langle v, b(v)\rangle \leqslant C_1 - C_2\|v\|^2 \quad . \tag{1.2}$$

Under condition (1.2) (and the non-degeneracy of σ), exponential convergence of the probability law $\mathcal{L}(Y_t)$ to a unique invariant measure μ in total variation is a classical fact, and can be mainly obtained via two different methods:

*Institut Elie Cartan, Université de Lorraine, B.P. 239, 54506 Vandoeuvre-lès-Nancy, Cedex, France. E-mail: aurelien.deya@univ-lorraine.fr

†Institut de Mathématiques de Toulouse, Université Paul Sabatier & INSA Toulouse, 135, av. de Rangueil, F-31077 Toulouse Cedex 4, France. E-mail: fabien.panloup@math.univ-toulouse.fr

‡Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, United States. E-mail: stindel@purdue.edu

(i) Functional inequalities. Starting from Poincaré type inequalities (or further refinements) for the solution of (1.1), and invoking Dirichlet form techniques, exponential and sub-exponential rates of convergence are obtained e.g in [2, 25].

(ii) Lyapunov/Coupling techniques. In these methods (see *e.g.* [8]), the idea is to try to stick some solutions of (1.1) (in an exponential time) with the following strategy: taking advantage of the Lyapunov assumption (1.2) (which can be strongly alleviated in the context of (1.1)) leads to some exponential bounds on the return-time of the (coupled) process into compact subsets of $\mathbb{R}^d$ (or more generally *petite sets*). Then, classical coupling techniques (involving the non-degeneracy of σ) allow to attempt the sticking of the paths when being in the compact subset.

Notice that in the setting of equation (1.1), the convergence analysis relies heavily on the Markov property for Y, or equivalently on the semi-group property for the transition probability. It also hinges on the irreducibility of Y, which can be seen as a non-degeneracy condition on the noisy part of the equation. Finally, notice that the first approach generally leads to sharper exponents but usually requires stronger assumptions.

Convergence to equilibrium being a relatively well understood phenomenon for equations like (1.1), recent developments in ergodic theory for stochastic equations have focused on deviations from the irreducible Markov setting. The reference [16] handles for instance infinite dimensional situations where only asymptotic couplings of the process (starting from different initial conditions) are available. Let us also mention [18], about a situation where the strong Feller property is fulfilled as $t \to \infty$, due to the degeneracy of the noise.

The current contribution is more directly related to another line of investigation, which aims at handling cases deviating from the fundamental Markov assumption. A general setting for this kind of situation is provided in the landmark of random dynamical systems [1, 5, 13]. However, the type of information one can retrieve with these techniques seldom include rates of convergence to an equilibrium measure. Alternatively, one can also consider differential systems driven by a fractional Brownian motion (fBm) as a canonical example on which non standard Markovian approaches to convergence can be elaborated. This point of view is ours, and is justified by the fact that fBm is widely used in applications (see e.g [14, 24, 21, 22]), and also by the fact that fBm can be seen as one of the simplest processes exhibiting long range dependence.

In this paper, we are thus concerned by the long time behavior of an equation which is similar to (1.1), except for the fact that the noisy input is a fractional Brownian motion. Specifically, we consider the following SDE:

$$dY_t = b(Y_t)dt + \sigma(Y_t)\, dX_t, \tag{1.3}$$

where the coefficients b and σ satisfy the same assumptions as above (in particular relation (1.2)), and where $(X_t)_{t\geqslant 0}$ is a d-dimensional H-fBm with Hurst parameter $H \in (\frac{1}{3}, 1)$. Notice that in the case $H > \frac{1}{2}$ equation (1.3) makes sense owing to Young integration techniques, whereas the case $H \in (1/3, 1/2)$ requires elements of rough paths theory (see Section 2.1).

The study of ergodic properties for fractional SDEs (under the stability assumption (1.2)) has been undertaken by Hairer [17], Hairer and Ohashi [19], and by Hairer and Pillai [20], respectively in the additive noise, multiplicative noise with $H > 1/2$ and multiplicative hypoelliptic noise with $H \in (1/3, 1/2)$. These papers mostly focus on a way to define stationary solutions, and on extending tools of the Markovian theory to the fBm setting. In particular, criteria for uniqueness of the invariant distribution are proved in increasingly demanding settings. Let us also mention the references [3, 4] for some results on approximations of stationary solutions. In all those references, the Markovian formalism is based on the Mandelbrot-Van Ness representation of the fractional Brownian motion, namely:

$$X_t = \alpha_H \int_{-\infty}^{0} (-r)^{H-\frac{1}{2}} \left(dW_{r+t} - dW_r\right), \quad t \geqslant 0, \tag{1.4}$$

where $(W_t)_{t\in\mathbb{R}}$ is a two-sided $\mathbb{R}^d$-valued Brownian motion and α_H is a normalization coefficient depending on H. It is then shown that $(Y_t, (X_{s+t})_{s\leqslant 0})_{t\geqslant 0}$ can be realized through a Feller transformation $(\mathcal{Q}_t)_{t\geqslant 0}$ whose definition is recalled below (see Section 3.1.1). In particular, an initial distribution of the dynamical system (Y, X) is a distribution μ_0 on $\mathbb{R}^d \times \mathcal{W}_-$, where $\mathcal{W}_-$ is an appropriate Hölder space (see Section 3.1.2 for more details). Rephrased in more probabilistic terms, an initial distribution is the distribution of a couple $(Y_0, (X_s)_{s\leqslant 0})$ where $(X_s)_{s\leqslant 0}$ is an $\mathbb{R}^d$-valued fBm on $(-\infty, 0]$. Then, such an initial distribution is called an

invariant distribution if it is invariant by the transformation $\mathcal{Q}_t$ for every $t \geqslant 0$. The uniqueness of such an invariant distribution is investigated in [17, 19, 20].

Let us now go back to our original question concerning the rate of convergence to equilibrium, which is obviously a natural problem when uniqueness holds for the invariant distribution. This problem has been first considered in [17], for equation (1.3) with an additive noise. In this context it is shown that the law of Y_t converges in total variation to the stationary regime, with a rate upper-bounded by $C_\varepsilon t^{-(\alpha-\varepsilon)}$ for any $\varepsilon > 0$, where

$$\alpha = \begin{cases} \frac{1}{8} & \text{if } H \in (\frac{1}{4}, 1) \backslash \{\frac{1}{2}\} \\ H(1-2H) & \text{if } H \in (0, \frac{1}{4}]. \end{cases} \tag{1.5}$$

The upper bound above is believed to be non-optimal, though its sub-exponential character can be interpreted as an effect of the non-Markovianity of the fBm X. Referring to our previous discussion on methods to achieve rates of convergence, functional inequalities tools are ruled out in the fBm setting, due to the absence of a real semi-group related to equation (1.3). The method chosen in [17] is thus based on coupling of solutions starting from different initial conditions. More specifically, the problem is reduced to a coupling between two paths starting from some initial conditions μ_0 and μ, where the second one denotes an invariant distribution of $(\mathcal{Q}_t)_{t\geqslant 0}$. The main step consists (classically) in finding a stopping time τ_∞ such that $(Y^{\mu_0}_{t+\tau_\infty})_{t\geqslant 0} = (Y^{\mu}_{t+\tau_\infty})_{t\geqslant 0}$. The rate of convergence in total variation is then obtained by means of an accurate bound on $\mathbb{P}(\tau_\infty > t)$, $t \geqslant 0$.

Within the general framework recalled above, the next challenge consists in extending the rate (1.5) to multiplicative noises. This has been achieved in Fontbona and Panloup [9], where the order of convergence (1.5) is obtained in the case $H > \frac{1}{2}$, with the additional assumption that the diffusion component σ is invertible and satisfies the following gradient type assumption: its inverse σ^{-1} is a Jacobian matrix. Our paper has thus to be seen as an improvement of [9] in two different directions:

(i) We get rid of the gradient type hypothesis assumed in [9], which extends the scope of application of our result.

(ii) We treat the case of an irregular fBm, with Hurst parameter $H \in (1/3, 1/2)$, which means that equation (1.3) has to be understood in the rough paths sense. Our main goal (see Theorem 2.5 for a precise statement) is then to obtain the rate of convergence (1.5) under those general conditions on σ and in the rough case.

Let us specify it right now: the techniques displayed in this paper can cover both the case $H \in (1/3, 1/2)$ (as mentionned in point *(ii)* above) and the case $H > 1/2$ (thus extending the results of [9] beyond the gradient type assumption, as reported in point *(i)*). This being said, for the sake of conciseness, we shall only express our analysis within the rough setting, that is when $H \in (1/3, 1/2)$, and therefore leave to the reader the details of the extension to the (simpler) Young situation $H > 1/2$ (see Remark 2.2 for a few additional comments on this topic).

In order to achieve our claimed rate of convergence, we shall implement the coupling strategy alluded to above. Let us briefly recall how this coupling strategy is divided in 3 steps. As a preliminary step, one waits that the two paths (starting respectively from μ and μ_0) get close. This is ensured by the reinforcing condition (1.2). Then, at each trial, the coupling attempt is divided in two steps. First, one tries in Step 1 to cluster the positions on an interval of length 1. Then, in Step 2, one tries to ensure that the paths stay clustered until $+\infty$. Actually, oppositely to the Markovian case where the paths stay naturally together after a clustering (by putting the same noise on each coordinate), the main difficulty here is that, due to the memory, staying together is costly. In other words, this property can be guaranteed only with the help of a non trivial coupling of the noises. One thus talks of *asymptotic coupling*. If one of the two previous steps fails, a new attempt will be made after a (long) waiting time which is called Step 3. During this step one waits again for the paths to get close, but one also expects the memory of the coupling cost to vanish sufficiently.

In our general rough setting with non constant coefficient σ, the implementation of the coupling strategy requires some non trivial adaptations of the general scheme. Let us highlight our main contributions in order to achieve the desired convergence rate:

1. The binding preliminary step relies on Lyapunov type properties of the differential equation (1.3). We

will invoke here some rough paths techniques based on discretization schemes.

2. In the additive case, two paths driven by the same fBm differ from a drift term, which leads to a straightforward way of sticking the paths in Step 1. We are no longer able to use this trick here, and our coupling is based on a linearization of equation (1.3). The analysis of such a linearization turns out to be demanding, and is one of our main efforts in this article.

3. The different trials we have to make in a context where nontrivial correlations occur force us to consider conditioning procedures. For these conditionings, we have chosen to decompose X into a Liouville fBm plus a smooth process with singularity at $t = 0$. The rough path formalism has to be adapted to this new setting.

Those steps are sometimes delicate, and will be detailed in the remainder of the article.

Our paper is organized as follows: In Section 2 we detail our assumptions and state our main result, namely Theorem 2.5. The Markov setting for equation (1.3), as well as Lyapunov type inequalities, are given in Section 3. Our global strategy to get the convergence rate is explained at Section 4. The singular rough equations needed for the coupling step are detailed in Section 5. Then the coupling Step 1 is achieved in Section 6. Eventually, the proof of Theorem 2.5 is completed in Sections 7, 8 and 9.

2 Setting and main result

We recall here the minimal amount of rough paths considerations allowing to define and solve equation (1.3) driven by a fBm with Hurst parameter $H > 1/3$. These preliminaries will be presented using terminology taken from the so-called algebraic integration theory, which is a variant of the rough paths theory introduced in [15] (see also [11]). Then we shall state precisely the main result of this article.

2.1 Hölder spaces, rough paths and rough differential equations

For an arbitrary real interval I, a vector space V and an integer $k \geqslant 1$, we denote by $\mathcal{C}_k(I;V)$ the set of functions $g : I^k \to V$ such that $g_{t_1 \cdots t_k} = 0$ whenever $t_i = t_{i+1}$ for some $i \leqslant k-1$. Such a function is called a $(k-1)$-increment. Then, for every $f \in \mathcal{C}_1(I;V)$ and $g \in \mathcal{C}_2(I;V)$, we successively define

$$(\delta f)_{st} := f_t - f_s \quad \text{and} \quad (\delta g)_{sut} = g_{st} - g_{su} - g_{ut} \quad ,$$

for any $s < u < t \in I$.

Our analysis will rely on some regularity consideration related to Hölder spaces. We thus start by introducing Hölder type norms for 1-increments: for every $f \in \mathcal{C}_2(I;V)$, we set

$$\mathcal{N}[f;\mathcal{C}_2^\mu(I;V)] := \sup_{s,t\in I} \frac{\|f_{st}\|}{|t-s|^\mu}, \quad \text{and} \quad \mathcal{C}_2^\mu(I;V) = \{f \in \mathcal{C}_2(I;V);\, \mathcal{N}[f;\mathcal{C}_2^\mu(I;V)] < \infty\}\,.$$

Observe now that the usual Hölder spaces $\mathcal{C}_1^\mu(I;V)$ are determined in the following way: for a continuous function $f \in \mathcal{C}_1(I;V)$, define

$$\mathcal{N}[f;\mathcal{C}_1^\mu(I;V)] = \mathcal{N}[\delta f;\mathcal{C}_2^\mu(I;V)], \quad \text{and} \quad \mathcal{C}_1^\mu(I;V) = \{f \in \mathcal{C}_1(I;V);\, \mathcal{N}[f;\mathcal{C}_1^\mu(I;V)] < \infty\}\,.$$

We shall also use the supremum norm on spaces $\mathcal{C}_k(I;V)$, which will be denoted by $\mathcal{N}[\cdot\,;\mathcal{C}_k^0(I;V)]$. Notice that when the context is clear, we will simply write $\mathcal{C}_k^\mu(I)$ for $\mathcal{C}_k^\mu(I;V)$.

The rough path theory can be seen as a differential calculus with respect to a Hölder continuous noise x, under a set of abstract assumptions. These assumptions are summarized in the following definition.

Definition 2.1. *Let γ be a constant greater than $1/3$ and consider a $\mathbb{R}^d$-valued γ-Hölder path x on some fixed interval $[0,T]$. We call a* Lévy area above x *any two-index map $\mathbf{x}^{\mathbf{2}} \in \mathcal{C}_2^{2\gamma}([0,T];\mathbb{R}^{d,d})$, which satisfies, for all $s < u < t \in [0,1]$ and all $i,j \in \{1,\dots,d\}$,*

$$\delta\mathbf{x}^{\mathbf{2};ij}_{sut} = \delta x^i_{su}\, \delta x^j_{ut} \quad \textit{and} \quad \mathbf{x}^{\mathbf{2};ij}_{st} + \mathbf{x}^{\mathbf{2};ji}_{st} = \delta x^i_{st}\, \delta x^j_{st}\ . \tag{2.6}$$

The couple $\mathbf{x} := (x, \mathbf{x^2})$ *is then called a* γ-rough path *above* x*, and we will use the short notation*

$$\|\mathbf{x}\|_{\gamma;I} := \mathcal{N}[x; \mathcal{C}_1^\gamma(I;\mathbb{R}^d)] + \mathcal{N}[\mathbf{x^2}; \mathcal{C}_2^{2\gamma}(I;\mathbb{R}^{d,d})] \quad ,$$

for any interval $I \subset [0,T]$.

When the rough path $\mathbf{x}$ can be approximated by smooth functions, one talks about a canonical lift, whose precise definition is given below.

Definition 2.2. *Given a path* $x \in \mathcal{C}_1^\gamma([0,1];\mathbb{R}^d)$*, we denote by* $x^n = x^{\mathcal{P}_n}$ *the sequence of (piecewise) smooth paths obtained through the linear interpolation of* x *along the dyadic partition* $\mathcal{P}_n$ *of* $[0,1]$*. Then we will say that* x *can be* canonically lifted into a rough path *if there exists a* γ*-rough path* $\mathbf{x} := (x, \mathbf{x^2})$ *above* x *such that the sequence* $\mathbf{x}^n := (x^n, \mathbf{x^{2,}}^n)$ *defined by*

$$\mathbf{x}_{st}^{\mathbf{2},n} := \int_s^t (\delta x^n)_{su} \otimes dx_u^n \tag{2.7}$$

converges to $\mathbf{x}$ *with respect to the norm*

$$\|\mathbf{x}\|_{(0,\gamma');[0,1]} := \mathcal{N}[x; \mathcal{C}_1^0([0,1];\mathbb{R}^d)] + \mathcal{N}[x; \mathcal{C}_1^{\gamma'}([0,1];\mathbb{R}^d)] + \mathcal{N}[\mathbf{x^2}; \mathcal{C}_2^{2\gamma'}([0,1];\mathbb{R}^{d,d})] \quad ,$$

for every $0 < \gamma' < \gamma$*. In this case, we will also denote this (necessarily unique) limit* $\mathbf{x}$ *as* $\mathfrak{L}(x)$.

We finally give the definition of solution to a noisy differential equation, such as our main object of interest (1.3). We are adopting here Davie's point of view (see [6]). Namely, we characterize the solution y by a Taylor expansion up to a remainder term whose Hölder regularity is strictly greater than 1.

Definition 2.3. (**Davie**) *Let* $\mathbf{x} := (x, \mathbf{x^2})$ *be a* γ*-rough path. Then, for all smooth vector fields*

$$b : \mathbb{R}^d \to \mathbb{R}^d \quad \textit{and} \quad \sigma : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^n;\mathbb{R}^d) \ ,$$

we call $y \in \mathcal{C}_1^\gamma(I;\mathbb{R}^d)$ *a solution (on* I*) of the equation*

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d\mathbf{x}_t \quad , \quad y_{t_0} = a \ , \tag{2.8}$$

if the two-parameter path R^y *defined as*

$$R_{st}^y := (\delta y)_{st} - b(y_s)\,(t-s) - \sigma_j(y_s)\,(\delta x^j)_{st} - (D\sigma_j \cdot \sigma_k)(y_s)\,\mathbf{x}_{st}^{\mathbf{2},jk}$$

belongs to $\mathcal{C}_2^\mu(I;\mathbb{R}^d)$*, for some parameter* $\mu > 1$*. Here, the notation* $D\sigma_j \cdot \sigma_k$ *stands for*

$$(D\sigma_i \cdot \sigma_k)(v) := (D\sigma_j)(v)(\sigma_k(v)) \ , \ \textit{for every } v \in \mathbb{R}^d \ . \tag{2.9}$$

Applications of the abstract rough paths setting to a fractional Brownian motion X depends on a proper construction of the Lévy area $\mathbf{X^2}$. The reader is referred to [12, Chapter 15] for a complete review of the methods enabling this construction. It can be summarized in the following way:

Proposition 2.4. *Let* $1/3 < H < 1/2$ *be a fixed Hurst parameter. Then the fBm* X *belongs almost surely to any space* $\mathcal{C}_1^\gamma$ *for* $\gamma < H$*, and can be lifted as a canonical rough path according to Definitions 2.1 and 2.2. Furthermore, for any* $0 \leqslant s < t \leqslant T$*, the random variable* $\mathbf{X}_{st}^{\mathbf{2}}$ *satisfies the following inequality:*

$$\mathbf{E}\left[\left|\mathbf{X}_{st}^{\mathbf{2}}\right|^p\right] \leqslant c_p\,(t-s)^{2Hp}, \quad p \geqslant 1.$$

As we shall see in the next section, Proposition 2.4 will allow us to solve equation (1.3) under reasonable assumptions on the coefficients b and σ.

2.2 Assumptions and Main Result

Having defined the notion of solution to equation (1.3), we can now proceed to a description of our main result. We first have to introduce a set of hypothesis on b and σ, beginning with a boundedness assumption.

Hypothesis (H1): $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ are smooth vector fields such that

$$\sup_{v\in\mathbb{R}^d} \|(D^{(1)}b)(v)\| \ < \ \infty \ , \quad \text{and for every } \ell \geqslant 0 \ , \ \sup_{v\in\mathbb{R}^d} \|(D^{(\ell)}\sigma)(v)\| \ < \ \infty \quad . \tag{2.10}$$

The second hypothesis is the Lyapunov-type assumption alluded to in the introduction, which is classically needed for the existence of an invariant distribution.

Hypothesis (H2): There exist $C_1, C_2 > 0$ such that for every $v \in \mathbb{R}^d$, one has

$$\langle v, b(v)\rangle \leqslant C_1 - C_2\|v\|^2 \quad . \tag{2.11}$$

Finally, one needs a non-degeneracy assumption on σ.

Hypothesis (H3): For every $x \in \mathbb{R}^d$, $\sigma(x)$ is invertible and

$$\sup_{x\in\mathbb{R}^d} \|\sigma(x)^{-1}\| < +\infty.$$

We are now in a position to state our main result. One denotes by $\mathcal{L}((Y_t^{\mu_0})_{t\geqslant 0})$ the distribution of the process Y on $\mathcal{C}([0,+\infty), \mathbb{R}^d)$ starting from a (generalized) initial condition μ_0 (see Subsection 3.1.2 below for detailed definitions of initial condition and invariant distribution). We also denote by $\bar{\mathcal{Q}}\mu$ the distribution of the stationary solution (starting from an invariant distribution μ). The distribution $\bar{\mu}_0(dx)$ denotes the first marginal of $\mu_0(dx, dw)$.

Theorem 2.5. *Let $H \in (1/3, 1/2)$, and assume* (**H1**), (**H2**), (**H3**) *hold true. Then:*
(i) *There exists a unique solution of equation* (1.3) *in the sense of Definition 2.3.*
(ii) *Existence and uniqueness hold for the invariant distribution μ.*
(iii) *Let μ_0 be an initial distribution such that there exists $r > 0$ satisfying $\int |x|^r \bar{\mu}_0(dx) < \infty$. Then for each $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that*

$$\|\mathcal{L}((Y_{t+s}^{\mu_0})_{s\geqslant 0}) - \bar{\mathcal{Q}}\mu\|_{TV} \leqslant C_\varepsilon t^{-(\frac{1}{8}-\varepsilon)}.$$

Remark 2.1. Item (i) in Theorem 2.5 is classical in rough path theory, since Proposition 2.4 holds true for our fBm X. We refer to [12] for the general theory of differential equations driven by a rough path. We prove existence of the invariant distribution below in Theorem 3.4 and Corollary 3.5. It is worth noting that even though this type of result is classical, its proof is highly technical in our rough and non-Markovian context. The main part of our work is then obviously to prove item (iii) which in turns implies uniqueness and achieves the proof of (ii). Also notice that the reinforcing assumption (**H2**) is fundamental for both the Lyapunov and the coupling steps in our proofs.

Remark 2.2. As mentioned before, when $H > 1/2$, Theorem 2.5 has already be shown in [9]. However, an additional gradient type assumption on σ was needed therein, that is: σ^{-1} is the Jacobian matrix of a function $h : \mathbb{R}^d \to \mathbb{R}^d$. Up to slight adaptations (involving in particular the non-integrability of $u \to u^{-H-\frac{1}{2}}$ when $H > 1/2$), the proof developed in this paper when $H \in (1/3, 1/2)$ (especially in Step 1) extends to the case $H > 1/2$ (and does not require the gradient assumption). In other words, the above result is still true when $H > 1/2$. For the sake of simplicity, we however choose to only consider the real new case $H < 1/2$ in the sequel.

3 Existence of invariant distribution

The main result of this section is Theorem 3.4 where we establish a new Lyapunov property for rough equations and deduce that existence holds for the invariant distribution under (**H1**) and (**H2**). Before, we need to recall some background about ergodic theory for rough equations. We assume that $H < 1/2$.

3.1 Markovian structure and invariant distribution

3.1.1 Background on the Markov structure above the solutions

As shown in [20] (going back to [17] and [19]), (1.3) can be endowed with a Markovian structure. Let us briefly recall the construction. The starting point is to build an appropriate Hölder space on which $(B_t^H)_{t\in\mathbb{R}}$ can be realized through a Markov transformation. Let $\mathcal{C}_0^\infty(\mathbb{R}_-)$ the space of $\mathcal{C}^\infty$-functions with compact support w on $\mathbb{R}_-$ with values in $\mathbb{R}^d$ satisfying $w(0)=0$. Let $\mathcal{W}_\gamma$ denote the Hölder-type space being the (Polish) closure of $\mathcal{C}_0^\infty(\mathbb{R}_-)$ for the norm $\|\,.\,\|_{\mathcal{W}_\gamma}$ defined by

$$\|x\|_{\mathcal{W}_\gamma} = \sup_{s,t\in\mathbb{R}_-} \frac{|x(t)-x(s)|}{|t-s|^\gamma(1+|t|+|s|)^{\frac{1}{2}}}. \tag{3.12}$$

For any $\gamma\in(1/3,H)$, there exists a probability $\mathbb{P}_-$ on $\mathcal{W}_\gamma$ such that the canonical process is a standard d-dimensional H-fBm indexed by $\mathbb{R}_-$. In the following, we set $\mathcal{W}_-=\mathcal{W}_\gamma$ and consider $x_-\in\mathcal{W}_-$. Set $\mathcal{W}_+=\mathcal{D}_g^{0,\gamma}([0,1])$ the closure of $\mathcal{C}_0^\infty([0,1])$ for the norm $\|\,.\,\|_{\gamma;[0,1]}$. Then, with the help of operators related to the Mandelbrot representation (see [20] for more precise statements), one can define a Feller transition kernel $\hat{\mathcal{P}}$ on $\mathcal{W}_-\times\mathcal{W}_+$ such that (with a slight abuse of notation), $\hat{\mathcal{P}}(x_-,d\mathbf{x}_+)=\mathbb{P}((\mathbf{B}_t^H)_{t\in[0,1]}\in d\mathbf{x}_+|(B_t^H)_{t\in\mathbb{R}_-}=x_-)$. Then, denoting by $\mathcal{W}=\mathcal{W}_-\times\mathcal{W}_+$, $\Pi:\mathcal{W}\mapsto\mathcal{C}((-\infty,1],\mathbb{R}^d)$ the map that concatenates x_- with the path component x_+ of $\mathbf{x}_+$, and $\mathbb{P}$ the probability measure on $\mathcal{W}$ defined by $\mathbb{P}(dx_-\times dx_+)=\mathbb{P}_-(dx_-)\hat{\mathcal{P}}(x_-,dx_+)$, $\Pi^*\mathbb{P}$ corresponds to the law of $(B_t^H)_{t\in(-\infty,1]}$ on $\mathcal{C}((-\infty,1],\mathbb{R}^d)$. Denoting by Θ the -1-time shift from $\mathcal{C}((-\infty,1],\mathbb{R}^d)$ to $\mathcal{C}((-\infty,0],\mathbb{R}^d)$, the previous construction implies that a two-sided fBm (on $\mathbb{R}^d$) can be realized through the (discrete-time) Feller Markov transition kernel $\mathcal{P}$ on $\mathcal{W}$ defined by

$$\mathcal{P}(x,.)=\delta_{\Theta(x)}\otimes\hat{\mathcal{P}}(\Theta(x),.).$$

The triplet $(\mathcal{W},\mathbb{P},\mathcal{P})$ is called the noise space. Then, for a given initial condition z and a given realization $x=(x_-,\mathbf{x}_+)$ of the driving noise, we denote by $(\Phi_t(z,\mathbf{x}_+))_{t\in[0,1]}$, the unique solution to (1.3) with initial condition z. Owing to [12] and Assumption (**H1**), $(z,\mathbf{x}_+)\mapsto\Phi_t(z,\mathbf{x}_+)$ is continuous on $\mathbb{R}^d\times\mathcal{W}_+$. It follows that the solution to (1.3) can be viewed as a Feller Markov process on $\mathbb{R}^d\times\mathcal{W}$ with transition kernel $\mathcal{Q}$ defined by : $\mathcal{Q}(z,x,.)=\Psi_z^*\mathcal{P}(x,.)$ where $\Psi_z(x)=(\Phi_1(z,\mathbf{x}_+),x)$.

Remark 3.3. In [17], the construction of the Markov structure is directly realized with the underlying Wiener process. Note that such a construction would be closer to the coupling viewpoint which is introduced below.

The reader can observe that the above construction only ensures the Markovian structure above the discrete-time process $(X_n)_{n\in\mathbb{N}}$ and not for the whole process $(X_t)_{t\geqslant0}$. However, an adaptation of the previous strategy leads to the construction of a Feller Markov semi-group $(\mathcal{Q}_t)_{t\geqslant0}$ above $(Y_t)_{t\geqslant0}$ (on $\mathbb{R}^d\times\mathcal{W}_\gamma$).

3.1.2 Invariant distribution

Following [19], a probability μ on $\mathbb{R}^d\times\mathcal{W}_-$ is called a *generalized initial condition* if $\Pi_{\mathcal{W}_-}\mu=\mathbb{P}_-$ (defined in the previous section).

Definition 3.1. *Let ν be a generalized initial condition. We say that ν is an invariant distribution for $(Y_t)_{t\geqslant0}$ if for every $t\geqslant0$, $\nu\mathcal{Q}_t=\nu$.*

Definition 3.2. *We say that $V:\mathbb{R}^d\mapsto\mathbb{R}$ is a Lyapunov function for $\mathcal{Q}$ if V is continuous and positive, if $\lim_{|x|\to+\infty}V(x)=+\infty$ and if there exist $C>0$ and $\rho\in(0,1)$ such that for every $t\in[0,1]$,*

$$\int V(x)(\mu\mathcal{Q}_t)(dx,dw)\leqslant C+\rho\int V(x)\mu(dx,dw) \tag{3.13}$$

for any generalized initial condition μ on $\mathbb{R}^d\times\mathcal{W}_-$.

We have the following (classical) result:

Proposition 3.3. *The existence of a Lyapunov function V for the Feller semi-group $(\mathcal{Q}_t)_{t\geqslant 0}$ implies the existence of an invariant distribution for $(Y_t)_{t\geqslant 0}$. Furthermore, for any generalized initial condition μ such that $\int V(x)\mu(dx,dw) < +\infty$, $\sup_{t\geqslant 0}\mathbb{E}_\mu[V(Y_t)] < +\infty$.*

Proof. Let μ denote an initial condition on $\mathbb{R}^d \times \mathcal{W}_-$ such that $\int V(x)\mu(dx,dw) < +\infty$ and denote by $(\mu_t)_{t\geqslant 1}$ the sequence defined by

$$\mu_t = \frac{1}{t}\sum_0^t \mu\mathcal{Q}_s ds, t \geqslant 1. \tag{3.14}$$

By construction and by the Feller property, every weak limit of $(\mu_t)_{t\geqslant 1}$ is an invariant distribution for $(\mathcal{Q}_t)_{t\geqslant 0}$. It is thus enough to prove the tightness of $(\mu_t)_{t\geqslant 1}$: owing to the stationarity of the increments of the fBm, the second marginal of μ_t does not depend on t. $\mathcal{W}_-$ being Polish, we are thus reduced to prove the tightness of $(\nu_t)_{t\geqslant 0}$, ν_t being the first marginal of μ_t. But the definition of the Lyapunov function implies (by an iteration) $\sup_{t\geqslant 1}\nu_t(V) < +\infty$ which in turn implies the tightness (using that $V^{-1}([0,K])$ is compact for any $K > 0$). □

The aim of the next subsection is the exhibition of such a Lyapunov function V for Q. The result will actually be derived from a general (deterministic) Lyapunov property for rough differential equations.

3.2 A Lyapunov property for rough differential equations

We go back here to the general case of a rough equation

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\mathbf{x}_t \quad , \quad t \in [0,1] \quad , \quad y_0 = a \in \mathbb{R}^d \,, \tag{3.15}$$

where $\mathbf{x}$ is a given (deterministic) γ-rough path on $[0,1]$, for some fixed parameter $\gamma \in (\frac{1}{3},\frac{1}{2})$. In what follows, we will write $\|\mathbf{x}\|_\gamma$ for $\|\mathbf{x}\|_{\gamma;[0,1]}$.

Theorem 3.4. *Under Hypothesis* (**H1**) *and for every initial condition $y_0 \in \mathbb{R}^d$, Equation (3.15) admits a unique solution y on $[0,1]$, in the sense of Definition 2.3. Besides, if we assume in addition that Hypothesis* (**H2**) *holds true, then there exists a constant C (which depends on b,σ,γ,C_1,C_2, but not on $\mathbf{x}$) such that*

$$\|y_1\|^2 \leqslant e^{-C_2/2}\|y_0\|^2 + C\big\{1 + \|\mathbf{x}\|_\gamma^\mu\big\} \quad , \quad \text{with } \mu := \frac{8}{3\gamma - 1} \quad . \tag{3.16}$$

Injecting this result into the stochastic setting of Section 3.1.1 (where $\mathbf{x} := \mathbf{X}$ is the canonical rough path above the fBm), the derivation of (3.13) is immediate. It is indeed a well-known fact (see for instance [12, Theorem 15.33]) that the random variable $\|\mathbf{X}\|_\gamma$ admits finite moments of any order, and we are therefore in a position to state the desired property:

Corollary 3.5. *In the setting of Section 3.1.1 and assuming that both Hypotheses* (**H1**) *and* (**H2**) *hold true, the map $V : x \mapsto \|x\|^p$ defines a Lyapunov function for $\mathcal{Q}$, for any $p \geqslant 1$. As a consequence, there exists at least one invariant distribution ν for $(Y_t)_{t\geqslant 0}$, which additionally admits finite moments of any order.*

The rest of this section is devoted to the proof of Theorem 3.4. Under Hypothesis (**H1**), the fact that there exists at most one solution to (3.15) (in other words, the uniqueness part of our statement) is a standard result, which can for instance be found to [6, Theorem 3.3]. On the opposite, due to the unboundedness of b, it seems that the proof of existence of a global solution on $[0,1]$ cannot be found as such in the literature, and we shall therefore provide a few details below.

In brief, our strategy towards Theorem 3.4 is based on a careful analysis of the natural discrete associated with (3.15), in the same spirit as in [6]. Let us thus introduce the sequence of dyadic partitions $\mathcal{P}_n := \{t_i = t_i^n := \frac{i}{2^n}\,;\, i = 0,\dots,2^n\}$ of $[0,1]$, and consider the discrete path y^n defined on $\mathcal{P}_n$ along the iterative formula

$$y_0^n := a \quad , \quad \delta y^n_{t_i t_{i+1}} = b(y^n_{t_i})\,\delta\mathcal{T}_{t_i t_{i+1}} + \sigma(y^n_{t_i})\,\delta x_{t_i t_{i+1}} + (D\sigma\cdot\sigma)(y^n_{t_i})\,\mathbf{x}^{\mathbf{2}}_{t_i t_{i+1}} \,. \tag{3.17}$$

We shall also be led to handle the following quantities associated with y^n: for $s,t\in\mathcal{P}_n$,

$$\begin{aligned} L^{y,n}_{st} &:= \delta y^n_{st} - \sigma(y^n_s)\,\delta x_{st} - (D\sigma\cdot\sigma)(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st} \\ R^{y,n}_{st} &:= \delta y^n_{st} - b(y^n_s)\,\delta\mathcal{T}_{st} - \sigma(y^n_s)\,\delta x_{st} - (D\sigma\cdot\sigma)(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st} \\ Q^{y,n}_{st} &:= \delta y^n_{st} - \sigma(y^n_s)\,\delta x_{st}\ . \end{aligned}$$

For every $s<t\in[0,1]$, we will write $[\![s,t]\!]=[\![s,t]\!]_n:=[s,t]\cap\mathcal{P}_n$, and we extend the norms introduced in Section 2.1 to discrete paths in a natural way, namely

$$\mathcal{N}[f;\mathcal{C}^\mu_2([\![\ell_1,\ell_2\tau\wedge 1]\!])]:=\sup_{s<t\in[\![\ell_1,\ell_2]\!]}\frac{\|f_{st}\|}{|t-s|^\mu}\quad,\quad \mathcal{N}[f;\mathcal{C}^\mu_1([\![\ell_1,\ell_2\tau\wedge 1]\!])]:=\mathcal{N}[\delta f;\mathcal{C}^\mu_2([\![\ell_1,\ell_2\tau\wedge 1]\!])]\ .$$

The starting point of our analysis is the following local estimate for $R^{y,n}$, which can be obtained as a straightforward application of our forthcoming general Proposition 5.9:

Proposition 3.6. *Fix $\kappa:=\frac{1}{2}\big(\frac{1}{3}+\gamma\big)$. Then, under Hypothesis* (**H1**)*, there exists a constant c_0 (which depends only on b,σ,γ) such that if we set*

$$T_0=T_0(\|\mathbf{x}\|):=\min\Big(1,\big(c_0\{1+\|\mathbf{x}\|_\gamma\}\big)^{-1/(\gamma-\kappa)}\Big)\ ,$$

one has, for every $\tau\in\mathcal{P}_n$ satisfying $0<\tau\leqslant T_0$ and every $k\leqslant 1/\tau$,

$$\mathcal{N}[R^{y,n};\mathcal{C}^{3\kappa}_2([\![k\tau,(k+1)\tau\wedge 1]\!])]\leqslant c_0\big\{1+\|y^n_{k\tau}\|\big\}\ . \tag{3.18}$$

Corollary 3.7. *In the setting of Proposition 3.6, there exists a constant c_1 (which depends only on b,σ,γ) such that for every $\tau\in\mathcal{P}_n$ satisfying $0<\tau\leqslant T_0$ and every $k\leqslant 1/\tau$, one has*

$$\mathcal{N}[y^n;\mathcal{C}^0_1([\![k\tau,(k+1)\tau\wedge 1]\!])]\leqslant c_1\{1+\|y^n_{k\tau}\|\}\ , \tag{3.19}$$

$$\mathcal{N}[y^n;\mathcal{C}^\gamma_1([\![k\tau,(k+1)\tau\wedge 1]\!])]\leqslant c_1\{1+\|y^n_{k\tau}\|\}\{1+\|\mathbf{x}\|_\gamma\} \tag{3.20}$$

and

$$\mathcal{N}[Q^{y,n};\mathcal{C}^{2\gamma}_2([\![k\tau,(k+1)\tau\wedge 1]\!])]\leqslant c_1\{1+\|y^n_{k\tau}\|\}\{1+\|\mathbf{x}\|_\gamma\}\ . \tag{3.21}$$

Proof. For every $t\in[\![k\tau,(k+1)\tau\wedge 1]\!]$, write

$$y^n_t=y^n_{k\tau}+b(y^n_{k\tau})\,\delta\mathcal{T}_{k\tau,t}+\sigma(y^n_{k\tau})\,\delta x_{k\tau,t}+(D\sigma\cdot\sigma)(y^n_{k\tau})\,\mathbf{x}^{\mathbf{2}}_{k\tau,t}+R^{y,n}_{k\tau,t}\ ,$$

so that using (3.18), we get $\|y^n_t\|\lesssim 1+\|y^n_{k\tau}\|+\|\mathbf{x}\|_\gamma T_0^\gamma$, and (3.19) now follows from the fact that $\|\mathbf{x}\|_\gamma T_0^\gamma\leqslant\|\mathbf{x}\|_\gamma T_0^{\gamma-\kappa}\lesssim 1$.

Then, in a more general way, we have for every $s<t\in[\![k\tau,(k+1)\tau\wedge 1]\!]$

$$\delta y^n_{st}=b(y^n_s)\,\delta\mathcal{T}_{st}+\sigma(y^n_s)\,\delta x_{st}+(D\sigma\cdot\sigma)(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st}+R^{y,n}_{st}$$

and

$$Q^{y,n}_{st}=b(y^n_s)\,\delta\mathcal{T}_{st}+(D\sigma\cdot\sigma)(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st}+R^{y,n}_{st}\ .$$

Injecting (3.18) and (3.19) into these expressions easily yields (3.20) and (3.21). □

Corollary 3.8. *Under Hypothesis* (**H1**)*, Equation (3.15) admits a unique global solution y on $[0,1]$. Besides, with the previous notations, there exists a subsequence of (y^n), that we still denote by (y^n), such that*

$$\max_{i=0,\dots,2^n}\|y_{t_i}-y^n_{t_i}\|\overset{n\to\infty}{\longrightarrow}0\ . \tag{3.22}$$

Proof. Although the two local estimates (3.19)-(3.20) are not uniform as such (that is, the right-hand side still depends on y^n), they easily give rise, via an obvious iterative procedure, to a uniform estimate for $\mathcal{N}[y^n;\mathcal{C}_1^{0,\gamma}([\![0,1]\!])] := \mathcal{N}[y^n;\mathcal{C}_1^0([\![0,1]\!])]+\mathcal{N}[y^n;\mathcal{C}_1^\gamma([\![0,1]\!])]$. Still denoting by y^n the continuous path obtained through the linear interpolation of $(y^n_{t_i})_{i=0,\dots,2^n}$, we thus get a uniform estimate for $\mathcal{N}[y^n;\mathcal{C}_1^{0,\gamma}([0,1])]$, which, by a standard compactness argument, allows us to conclude about the existence of a path $y\in\mathcal{C}_1^\gamma([0,1])$, as well as a subsequence of y^n (that we still denote by y^n), such that $y^n\to y$ in $\mathcal{C}_1^{0,\gamma'}([0,1])$ for every $0<\gamma'<\gamma$.

The fact that y actually defines a solution of (3.15) is then an easy consequence of the bound (3.18). The details of this procedure can for instance be found at the end of [7, Section 3.3]. □

Let us now turn to the proof of the second part of Theorem 3.4, that is to the proof of (3.16) under Hypotheses (**H1**) and (**H2**). To this end, we introduce, for every $n\geqslant 0$, the additional discrete path $z^n:\mathcal{P}_n\to\mathbb{R}$ defined for every $t\in\mathcal{P}_n$ as

$$z^n_t := \frac{1}{2}\|y^n_t\|^2 \ .$$

In the same vein as above, we will lean on the following quantities related to z^n: for every $s,t\in\mathcal{P}_n$,

$$\begin{aligned} R^{z,n}_{st} &:= \delta z^n_{st} - \langle y^n_s, b(y^n_s)\rangle\,\delta\mathcal{T}_{st} - \langle y^n_s,\sigma(y^n_s)\rangle\,\delta x_{st} - \Sigma(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st}\\ L^{z,n}_{st} &:= \delta z^n_{st} - \langle y^n_s,\sigma(y^n_s)\rangle\,\delta x_{st} - \Sigma(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st}\\ Q^{z,n}_{st} &:= \delta z^n_{st} - \langle y^n_s,\sigma(y^n_s)\rangle\,\delta x_{st}\ , \end{aligned}$$

where we have set

$$\Sigma(y^n_s) := \langle\sigma(y^n_s),\sigma(y^n_s)\rangle + \langle y^n_s,(D\sigma\cdot\sigma)(y^n_s)\rangle\ .$$

Just to be clear, the notation for the second-order term in $R^{z,n},L^{z,n}$ specifically refers to the sum

$$\Sigma(y^n_s)\,\mathbf{x}^{\mathbf{2}}_{st} = \big\{\langle\sigma_j(y^n_s),\sigma_k(y^n_s)\rangle + \langle y^n_s,(D\sigma_j\cdot\sigma_k)(y^n_{st})\rangle\big\}\,\mathbf{x}^{\mathbf{2},jk}_{st}\ .$$

Finally, along the same lines as in the subsequent Section 5, we set, if $s=\frac{p}{2^n}$ and $t=\frac{q}{2^n}$ and $G:[\![0,1]\!]\to\mathbb{R}^d$,

$$\mathcal{M}^\mu\big[G;[\![s,t]\!]\big] := \sup_{p\leqslant i\leqslant q}\frac{\|G_{t_it_{i+1}}\|}{|t_{i+1}-t_i|^\mu}\ .$$

Let us start with a few estimates on $R^{z,n}$, for which Hypothesis (**H2**) is still not required:

Lemma 3.9. *Under Hypothesis* (**H1**) *and with the above notations, there exists a constant c_2 (which depends only on b,σ,γ) such that for every $s<t\in\mathcal{P}_n$, one has*

$$\mathcal{M}^{3\gamma}\big[R^{z,n};[\![s,t]\!]\big]\leqslant c_2\{1+\|\mathbf{x}\|_\gamma^2\}\{1+\mathcal{N}[y^n;\mathcal{C}_1^0([\![s,t]\!])]^2\}\ . \tag{3.23}$$

Proof. We have

$$\delta z^n_{t_it_{i+1}} = \langle y^n_{t_i},\delta y^n_{t_it_{i+1}}\rangle + \frac{1}{2}\langle\delta y^n_{t_iti+1},\delta y^n_{t_it_{i+1}}\rangle\ ,$$

and so, injecting (3.17) into the first term immediately gives, thanks to the second identity in (2.6),

$$R^{z,n}_{t_it_{i+1}} = \frac{1}{2}\langle\delta y^n_{t_iti+1},\delta y^n_{t_it_{i+1}}\rangle - \langle\sigma(y^n_{t_i}),\sigma(y^n_{t_i})\rangle\mathbf{x}^{\mathbf{2}}_{t_it_{i+1}} = \langle\sigma(y^n_{t_i})\,\delta x_{t_it_{i+1}} + \frac{1}{2}Q^{y,n}_{tit_{i+1}},Q^{y,n}_{tit_{i+1}}\rangle\ . \tag{3.24}$$

Finally, since $Q^{y,n}_{t_it_{i+1}} = b(y^n_{t_i})\,\delta\mathcal{T}_{t_it_{i+1}} + (D\sigma\cdot\sigma)(y^n_{t_i})\,\mathbf{x}^{\mathbf{2}}_{t_it_{i+1}}$, it is immediate that

$$\|Q^{y,n}_{t_it_{i+1}}\|\lesssim|t_{i+1}-t_i|\{1+\|y^n_{t_i}\|\}+|t_{i+1}-t_i|^{2\gamma}\|\mathbf{x}\|_\gamma\ .$$

Going back to (3.24), we get the conclusion. □

Proposition 3.10. *Assume Hypothesis* (**H1**) *holds true and let $T_0=T_0(\|\mathbf{x}\|_\gamma)$ be the time defined in Proposition 3.6. Then there exists a constant c_3 (which depends only on b,σ,γ) such that for every $\tau\in\mathcal{P}_n$ satisfying $0<\tau\leqslant T_0$ and every $k\leqslant 1/\tau$, one has*

$$\mathcal{N}[R^{z,n};\mathcal{C}_2^{3\gamma}([\![k\tau,(k+1)\tau\wedge 1]\!])]\leqslant c_3\{1+\|\mathbf{x}\|_\gamma^3\}\{1+z^n_{k\tau}\}\ .$$

Proof. Thanks to Lemma 5.6, we can rely on the estimate

$$\mathcal{N}[R^{z,n};\mathcal{C}_2^{3\gamma}([\![k\tau,(k+1)\tau\wedge 1]\!])]\lesssim\mathcal{M}^{3\gamma}\big[R^{z,n};[\![k\tau,(k+1)\tau\wedge 1]\!]\big]+\mathcal{N}[\delta R^{z,n};\mathcal{C}_3^{3\gamma}([\![k\tau,(k+1)\tau\wedge 1]\!])]\;.$$

As far as the first term is concerned, combining (3.23) and (3.19) allows us to assert that

$$\mathcal{M}^{3\gamma}\big[R^{z,n};[\![k\tau,(k+1)\tau\wedge 1]\!]\big]\lesssim\{1+\|\mathbf{x}\|_\gamma^2\}\{1+z_{k\tau}^n\}\;.$$

Then, for every $s<u<t\in[\![k\tau,(k+1)\tau\wedge 1]\!]$, decompose $\delta R_{sut}^{z,n}$ as

$$\delta R_{sut}^{z,n}=-\delta(\langle y^n,b(y^n)\rangle)_{su}\,\delta\mathcal{T}_{ut}+\delta L_{sut}^{z,n}\;.$$

On the one hand, one has, by (3.19) and (3.20),

$$\big|\delta(\langle y^n,b(y^n)\rangle)_{su}\big|\leqslant\big|\langle\delta y_{su}^n,b(y_u^n)\rangle\big|+\big|\langle y_s^n,\delta b(y^n)_{su}\rangle\big|\lesssim|u-s|^\gamma\{1+\|\mathbf{x}\|_\gamma\}\{1+z_{k\tau}^n\}\;.$$

On the other hand, combining Chen's identity with elementary Taylor expansions easily leads us to the decomposition

$$\delta L_{sut}^{z,n}=\big\{I_{su}^i+II_{su}^i+III_{su}^i+IV_{su}^i\big\}\,\delta x_{ut}^i+\delta\Sigma_{ij}(y^n)_{su}\,\mathbf{x}_{ut}^{\mathbf{2},ij}\;,$$

with

$$\Sigma_{ij}(y^n):=\langle\sigma_i(y^n),\sigma_j(y^n)\rangle+\langle y^n,(D\sigma_i\cdot\sigma_j)(y^n)\rangle\;,$$

$$I_{su}^i:=\langle\delta y_{su}^n,\delta\sigma_i(y^n)_{su}\rangle\;,\quad II_{su}^i:=\langle\sigma_i(y_s^n),Q_{su}^{y,n}\rangle\;,\quad III_{su}^i:=\int_0^1 d\xi\,\langle y_s^n,D\sigma_i(y_s^n+\xi\,\delta y_{su}^n)Q_{su}^{y,n}\rangle$$

and finally

$$IV_{su}^i:=\int_0^1 d\xi\,\langle y_s^n,\big[D\sigma_i(y_s^n+\xi\,\delta y_{su}^n)-D\sigma_i(y_s^n)\big](\sigma_j(y_s^n))\rangle\,\delta x_{su}^j\;.$$

With the above expressions in mind and using the three estimates (3.19), (3.20) and (3.21), it is not hard to check that

$$\big|\delta L_{sut}^{z,n}\big|\lesssim|t-s|^{3\gamma}\{1+\|\mathbf{x}\|_\gamma^3\}\{1+z_{k\tau}^n\}\;,$$

which achieves the proof of our assertion.

□

Let us finally involve Hypothesis (**H2**) into the picture:

Corollary 3.11. *Assume Hypotheses* (**H1**) *and* (**H2**) *hold true and let* $T_0=T_0(\|\mathbf{x}\|_\gamma)$ *be the time defined in Proposition 3.6. Then there exist constants* c_4,c_5 *(both depending only on* b,σ,γ,C_1,C_2*) such that if we set*

$$T_1=T_1(\|\mathbf{x}\|_\gamma):=\min\left(T_0,\frac{2}{C_2},\left(\frac{1}{c_4\{1+\|\mathbf{x}\|_\gamma^3\}}\right)^{1/(3\gamma-1)}\right)\;,$$

one has, for every $\tau\in\mathcal{P}_n$ *satisfying* $0<\tau\leqslant T_1$ *and every* $k\leqslant 1/\tau$,

$$z_{(k+1)\tau\wedge 1}^n\leqslant\Big(1-\frac{C_2}{2}\tau\Big)z_{k\tau}^n+c_5\{1+\|\mathbf{x}\|_\gamma^2\}\tau^{2\gamma-1}\;,\tag{3.25}$$

where we recall that the two parameters C_1,C_2 *have been introduced in Hypothesis (H2).*

Proof. Using Hypothesis (**H2**), we get that for every $\tau\in\mathcal{P}_n$ and every such that $k\leqslant 1/\tau$,

$$\begin{aligned}z_{(k+1)\tau\wedge 1}^n&=z_{k\tau}^n+\langle y_{k\tau}^n,b(y_{k\tau}^n)\rangle\,\delta\mathcal{T}_{k\tau,(k+1)\tau\wedge 1}\\&\quad+\langle y_{k\tau}^n,\sigma(y_{k\tau}^n)\rangle\,\delta x_{k\tau,(k+1)\tau\wedge 1}+\Sigma(y_{k\tau}^n)\,\mathbf{x}_{k\tau,(k+1)\tau\wedge 1}^{\mathbf{2}}+R_{k\tau,(k+1)\tau\wedge 1}^{z,n}\\&\leqslant(1-C_2\tau)z_{k\tau}^n+C_1\\&\quad+\langle y_{k\tau}^n,\sigma(y_{k\tau}^n)\rangle\,\delta x_{k\tau,(k+1)\tau\wedge 1}+\Sigma(y_{k\tau}^n)\,\mathbf{x}_{k\tau,(k+1)\tau\wedge 1}^{\mathbf{2}}+R_{k\tau,(k+1)\tau\wedge 1}^{z,n}\;,\end{aligned}$$

and so, thanks to Proposition 3.10, we can conclude that for every $0 < \tau \leqslant \min\big(T_0, \frac{2}{C_2}\big)$ and every $k \leqslant 1/\tau$, one has

$$z^n_{(k+1)\tau\wedge 1} \leqslant (1 - C_2\tau) z^n_{k\tau} + C_1 + c_4\Big[\|\mathbf{x}\|_\gamma \tau^\gamma \big\{1 + (z^n_{k\tau})^{\frac{1}{2}}\big\} + \frac{C_2}{4}\tau^{3\gamma}\{1 + \|\mathbf{x}\|^3_\gamma\}\{1 + z^n_{k\tau}\}\Big] ,$$

for some constant $c_4 = c_4(b, \sigma, \gamma, C_2)$. Now, by the very definition of T_1, we know that if $0 < \tau \leqslant T_1$, then

$$c_4\tau^{3\gamma}\{1 + \|\mathbf{x}\|^3_\gamma\} \leqslant \tau\big(c_4\tau^{3\gamma-1}\{1 + \|\mathbf{x}\|^3_\gamma\}\big) \leqslant \tau .$$

To achieve the proof, it now suffices to use the basic inequality

$$c_4\|\mathbf{x}\|_\gamma \tau^\gamma (z^n_{k\tau})^{\frac{1}{2}} \leqslant \frac{C_2}{4}\tau z^n_{k\tau} + \frac{c_4^2}{C_2}\|\mathbf{x}\|^2_\gamma \tau^{2\gamma-1} .$$

□

At this point, we are very close to (3.16). With the notations of Corollary 3.11, consider n large enough such that we can exhibit $\tau_0 \in \mathcal{P}_n$ satisfying $\frac{1}{2}T_1 \leqslant \tau_0 \leqslant T_1$, and then let K be the integer such that $(K-1)\tau_0 \leqslant 1 < K\tau_0$. Iterating the bound (3.25) with $\tau = \tau_0$ yields that

$$\begin{aligned} z^n_1 &\leqslant \Big(1 - \frac{C_2}{2}\tau_0\Big)^K z^n_0 + c_5\, K\{1 + \|\mathbf{x}\|^2_\gamma\}\tau_0^{2\gamma-1} \\ &\leqslant \Big(1 - \frac{C_2}{2}\tau_0\Big)^{\frac{1}{\tau_0}} z^n_0 + c_5\, K\{1 + \|\mathbf{x}\|^2_\gamma\}\tau_0^{2\gamma-1} \\ &\leqslant e^{-C_2/2} z^n_0 + c_5\, K\{1 + \|\mathbf{x}\|^2_\gamma\}\tau_0^{2\gamma-1} . \end{aligned}$$

Thanks to (3.22), the conclusion is now immediate, by noting that $K\tau_0^{2\gamma-1} \lesssim T_1^{2\gamma-2}$ and then using the explicit description of T_1, T_0 in terms of $\|\mathbf{x}\|_\gamma$.

4 Sketch of the strategy

We now turn to the second part of Theorem 2.5 about the convergence in total variation of the process towards the stationary solution. This result is based on a coupling method first introduced in [17]. We thus begin by recalling the details of the strategy. To this end, we first introduce some notations about the Mandelbrot-Van Ness representation of the fBm.

4.1 Decomposition of the fBm

As recalled in (1.4), the Mandelbrot-Van Ness formula allows us to realize any fBm $(X_t)_{t\geqslant 0}$ (with Hurst parameter $H \in (0,1)$) through a standard two-sided Brownian motion $(W_t)_{t\in\mathbb{R}}$. The representations immediately gives rise to the decomposition

$$X_t = D_t + Z_t , \tag{4.26}$$

where

$$D_t := \alpha_H \int_{-\infty}^0 \{(t-r)^{H-\frac{1}{2}} - (-r)^{H-\frac{1}{2}}\}\, dW_r \quad \text{and} \quad Z_t := \alpha_H \int_0^t (t-r)^{H-\frac{1}{2}}\, dW_r . \tag{4.27}$$

is seen a the 'past' component encoding the 'memory' of W, while

$$Z_t := \alpha_H \int_0^t (t-r)^{H-\frac{1}{2}}\, dW_r$$

stands for the 'innovation' process (when looking at X after time 0).

It turns out that, away from 0, the process D so defined is smooth (see Lemma 6.5 for details), so that the roughness of X is essentially inherited from that of Z. This basic observation will be one the keys of

our analysis, at every step of the strategy (it is worth noting that similar ideas already appear in [20]). All along the procedure, we will thus be led to control the past of the process through the quantity

$$|||D|||_{1;\gamma} := \sup_{t\in(0,1]} t^{1-\gamma}|D'(t)| \ , \tag{4.28}$$

for some fixed parameter $\gamma \in (0, H)$. Let us more generally introduce the following class of functions:

Notation 4.1. *For every $k \geqslant 1$ and every $\gamma \in (0,1)$, we denote by $\mathcal{E}^k_\gamma$ the space of paths $f : [0,1] \to \mathbb{R}^d$ which are continuous on $[0,1]$, k-times differentiable on $(0,1]$, and such that*

$$|||f|||_{k;\gamma} := \max_{1\leqslant \ell\leqslant k} \sup_{t\in(0,1]} t^{\ell-\gamma}|f^{(\ell)}(t)| \ < \ \infty \ . \tag{4.29}$$

4.2 The general 3-step scheme

Let $(X_t)_{t\in\mathbb{R}}$ and $(\widetilde{X}_t)_{t\in\mathbb{R}}$ denote two fractional Brownian motions with common Hurst parameter $H \in (1/3, 1)$. Then, denote by $(Y_t, \widetilde{Y}_t)$, a couple of solutions to (1.3):

$$\begin{cases} dY_t = b(Y_t)dt + \sigma(Y_t)\, d\mathbf{X}_t \\ d\widetilde{Y}_t = b(\widetilde{Y}_t)dt + \sigma(\widetilde{Y}_t)\, d\widetilde{\mathbf{X}}_t \end{cases} \tag{4.30}$$

with initial conditions $(Y_0, (X_t)_{t\leqslant 0})$, $(\widetilde{Y}_0, (X_t)_{t\leqslant 0})$. We denote by $(\mathcal{F}_t)_{t\geqslant 0}$ the usual augmentation of the filtration $(\sigma(X_s, \widetilde{X}_s, (Y_0, \widetilde{Y}_0))_{s\leqslant t})_{t\geqslant 0}$. To initiate the coupling procedure without "weight of the past", we will certainly assume that $a.s$,

$$(X_t)_{t\leqslant 0} = (\widetilde{X}_t)_{t\leqslant 0}$$

and that the initial distribution $\widetilde{\mu}$ of $(Y_0, \widetilde{Y}_0)$ is of the form

$$\widetilde{\mu}(dx, d\widetilde{x}, dw) = \nu_0(w, dx)\nu(w, d\widetilde{x})\mathbb{P}_H(dw) \tag{4.31}$$

where $\mathbb{P}_H$ denotes the distribution of a fBm $(X_t)_{t\leqslant 0}$ on $\mathcal{C}(\mathbb{R}^-, \mathbb{R}^d)$ and the transitions probabilities $\nu_0(., dx)$ and $\nu(., d\widetilde{x})$ correspond respectively to the conditional distributions of Y_0 and $\widetilde{Y}_0$ given $(X_t)_{t\leqslant 0}$. Furthermore, we assume that

$$\nu_0(w, d\widetilde{x})\mathbb{P}_H(dw) = \mu_0 \quad \text{and that } \nu(w, d\widetilde{x})\mathbb{P}_H(dw) = \mu. \tag{4.32}$$

In other words, $\tilde{Y}$ is a stationary solution whereas Y starts with a given initial condition μ_0.

The processes $(X_t)_{t\in\mathbb{R}}$ and $(\widetilde{X}_t)_{t\in\mathbb{R}}$ can be realized through the decomposition introduced in the previous subsection with respect to some two-sided Brownian motions respectively denoted by W and $\widetilde{W}$. In particular, the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is also generated by $(\sigma(W_s, \widetilde{W}_s, (Y_0, \widetilde{Y}_0))_{s\leqslant t})_{t\geqslant 0}$.

Furthermore, we will assume in all the proof that on $[0, \infty)$, W and $\widetilde{W}$ (resp. X and $\widetilde{X}$) differ by a (random) drift term denoted by $g_{_W}$ (resp. g_X):

$$d\widetilde{W}_t = dW_t + g_{_W}(t)dt \quad \text{and} \quad d\widetilde{X}_t = dX_t + g_X(t)dt. \tag{4.33}$$

Note that the functions $g_{_W}$ and g_X are linked by the following formulas when make sense (see [17], Lemma 4.2 for details):

$$g_{_W}(t) = \alpha_H \frac{d}{dt}\int_{-\infty}^t (t-s)^{\frac{1}{2}-H} g_X(s)ds \tag{4.34}$$

$$g_X(t) = \gamma_H \alpha_{1-H} \frac{d}{dt}\int_{-\infty}^t (t-s)^{H-\frac{1}{2}} g_{_W}(s)ds. \tag{4.35}$$

The idea is to build $g_{_W}$ (resp. g_X) in order to stick Y and $\widetilde{Y}$. We set

$$\tau_\infty := \inf\{t \geqslant 0, \ Y_s = \widetilde{Y}_s \ \text{ for all } s \geqslant t\} \ .$$

As usual, this coupling will be achieved after a series of trials. As mentioned in the introduction, each trial is decomposed in three steps:

- Step 1: Try to couple the positions with a *controlled cost* (in a sense made precise below).
- Step 2 (specific to non-Markov processes): Try to keep the paths fastened together (*i.e.* to ensure that $g_X(t) = 0$)
- Step 3: If Step 2 fails, wait a sufficiently long time in order that in the next trial, Step 1 be achieved with a *controlled cost* and with (uniformly lower-bounded away from 0) probability. During this step, we suppose that $g_W(t) = 0$.

Let us make a few precisions:

$\triangleright$ We denote by $\tau_0 \geqslant 0$ the beginning of the first trial and by τ_k, $k \geqslant 1$, the end of each trial. We will assume in the sequel that $\tau_0 = 0$.

If $\tau_k = +\infty$, the coupling tentative has been successful. Otherwise, τ_k is the end of Step 3 of trial k. We will assume that

$$\text{for all } t \in \mathbb{R}_-, \quad W_t = \widetilde{W}_t \quad a.s. \quad \text{or equivalently that} \quad g_W(t) = g_X(t) = 0 \quad \text{on } \mathbb{R}_-.$$

$\triangleright$ Step 1 is carried out on each interval $[\tau_{k-1}, \tau_{k-1}+1]$. The "cost" of coupling is represented by the function g_W that one needs to build on $[\tau_{k-1}, \tau_{k-1}+1]$ in order to get Y and $\widetilde{Y}$ stuck together at time $\tau_{k-1}+1$. Oppositely to the Markovian case, this cost does not only depend on the positions of $Y_{\tau_{k-1}}$ and $\widetilde{Y}_{\tau_{k-1}}$ but also on the past of the Brownian motions, which have a (strong) influence on the dynamics of X and $\widetilde{X}$.

If Step 1 fails (which includes the case where one does not attempt the coupling), one begins Step 3 (see below). Otherwise, one begins Step 2.

$\triangleright$ Step 2 is in fact a series of trials on some intervals I_ℓ with length

$$|I_\ell| = c_2 2^\ell \tag{4.36}$$

where c_2 is a constant greater than one which will be calibrated in the sequel. More precisely, one successively tries to keep Y and $\widetilde{Y}$ as being equal on intervals $[\tau_{k-1} + 1 + c_2 \sum_{u=1}^{\ell-1} 2^k, \tau_{k-1} + 1 + c_2 \sum_{u=1}^{\ell} 2^k]$ (with the convention $\sum_{\varnothing} = 0$). Equivalently, this means that on these successive intervals,

$$g_X(t) = 0 \quad a.s.$$

Thus, the first natural question is the following: how to build the coupling $(W, \widetilde{W})$ of the innovations in order to ensure this property, *i.e.* what is the corresponding function g_W on these successive intervals which ensures that $g_X = 0$? The answer is given by Lemma 4.3 of [17] that we choose to recall here in a slightly different way:

Lemma 4.1. Let τ, t_0 and ε be some positive numbers. Assume that $g_W(t) = 0$ on $(-\infty, 0] \cup [\tau - \varepsilon, \tau]$. Then, $g_X(t) = 0$ for all $t \in [\tau, \tau + t_0]$ if and only if

$$\forall t \in (\tau, \tau + t_0], \quad g_W^\tau(t) = (\mathcal{R}_0 g_W^\tau)(t) \tag{4.37}$$

where $g_W^\tau(.) = g_W(\tau + .)$ and for $T \geqslant 0$, for a given function g (for which the definition below makes sense)

$$\forall t > 0, \quad (\mathcal{R}_T g)(t) = C_H \int_{-\infty}^{0} \frac{t^{\frac{1}{2}-H}(T-s)^{H-\frac{1}{2}}}{t+T-s} g(s) ds, \quad t \in (0, +\infty)$$

where C_H is a constant (depending only on H).

Remark 4.4. The fact that $g_W(t) = 0$ on $(-\infty, 0] \cup [\tau - \varepsilon, \tau]$ implies that $\mathcal{R}_T g$ is well-defined for every $T > 0$. Also, note that $\mathcal{R}_T$ is a linear operator; this property will be convenient to separate each contribution of the past, *i.e.* coming from each attempt.

The attempt is successful if Step 1 and Step 2 (*i.e.* all the sub-attempts of this step) are. To ensure a positive probability to the success of the k-th attempt, one needs certainly to impose some conditions on the system at time τ_{k-1}.

In the continuity of [17] and [9], we thus introduce an admissibility condition:

Definition 4.2. *Let K and α be some positive constants, $\gamma \in (0,H)$, and fix a time $\tau \geqslant 0$. Then we will say that a state $\pi := (a, \tilde{a}, w, \widetilde{w}) \in (\mathbb{R}^d)^2 \times \mathcal{C}(]-\infty,\tau];\mathbb{R}^d)^2$ is (K,α,γ)-admissible (at time τ) if $\tau < +\infty$ and if the following conditions are satisfied:*

(i) It holds that $\widetilde{w} = w + \int_{-\infty}^{\cdot} g_{\scriptscriptstyle W}(s)\,ds$, with $g_{\scriptscriptstyle W}$ satisfying

$$\sup_{T\geqslant 0}\int_0^{+\infty}(1+t)^{2\alpha}|(\mathcal{R}_T|g_{\scriptscriptstyle W}^{\tau}|)(t)|^2dt \leqslant 1\ ; \tag{4.38}$$

(ii) It holds that

$$|a| + |\tilde{a}| + |||D^{(\tau)}(w)\,|||_{1;\gamma} + |||D^{(\tau)}(\widetilde{w})|||_{1;\gamma} \leqslant K \tag{4.39}$$

where we have set, for $t \in [0,1]$,

$$D^{(\tau)}(w)_t := \int_{-\infty}^{\tau}\Big((t+\tau-r)_+^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}}\Big)\,dw_r\ .$$

Remark 4.3. We are aware that, following the subsequent Lemma 6.5, the above transformation $D^{(\tau)}$ is only defined on a subspace $\Omega_-^{\tau} \subset \mathcal{C}(]-\infty,\tau];\mathbb{R}^d)$ of full Wiener measure (obtained through a shifting of Ω_-). Nevertheless, using the stability properties reported in the very same Lemma 6.5, it is readily checked that, on top of being of full Wiener measure, such a space is left invariant by the successive constructions of our strategy. We can therefore assume that at any time τ, the noise component of the system under consideration takes value in Ω_-^{τ}, which allows us to justify this slight abuse of notation.

If the system is admissible, that is if

$$\big(Y(\omega)_{\tau_{k-1}(\omega)}, \widetilde{Y}(\omega)_{\tau_{k-1}(\omega)}, (W(\omega)_t)_{t\leqslant\tau_{k-1}(\omega)}, (\widetilde{W}(\omega)_t)_{t\leqslant\tau_{k-1}(\omega)}\big)$$

defines a (K,α,γ)-admissible state at time $\tau_{k-1}(\omega)$, one attempts the coupling. Otherwise, one waits for the next one, *i.e.* one sets $g_{\scriptscriptstyle W}(t) = 0$ on $[\tau_{k-1},\tau_k]$ (One will come back below on the duration $\Delta\tau_k := \tau_k - \tau_{k-1}$).

Regarding Lemma 4.1, one can remark that Condition 4.38 plays a fundamental role in Step 2. More precisely, it can be understood as a sufficient condition to ensure the success of the series of attempts involved by Step 2.

The second assumption is mainly linked to Step 1. Roughly, it ensures that each marginal is in a sufficiently controllable state to couple the positions with a bounded cost. The fact that the positions live in a compact set at the beginning of the attempt is classical. The second condition (on $D^{(\tau)}$) is of course specific to this non-Markovian setting.

Finally, note that the first condition will be ensured with probability 1 at the beginning of the attempt, whereas, of course, the second one will occur only with a (lower-bounded) positive probability. We denote by

$$A_k(K,\alpha,\gamma) = \{\omega :\ \text{system } (K,\alpha,\gamma)\text{-admissible at time } \tau_{k-1}(\omega)\}. \tag{4.40}$$

$\triangleright$ If the coupling attempt fails, one begins Step 3, *i.e.* one waits sufficiently before another attempt. This waiting time will be chosen exponentially proportional to the length of the failing attempt. More precisely, let $\ell \geqslant 1$ denote the numbers of trials in Step 2 and adopt the convention, that $\ell = 0$ if Step 1 fails (including the case where the coupling is not attempted since the system is not admissible at time τ_{k-1}). Set

$$F_{k,\ell} = \{\text{At trial } k, \text{ Step 2 fails after } \ell \text{ attempts}\}. \tag{4.41}$$

One assumes in the sequel that for every $\omega \in F_{k,\ell}$,

$$\tau_k - \tau_{k-1}^3 = \Delta_3(\ell,k) \quad \text{with} \quad \Delta_3(\ell,k) := c_3 a_k 2^{\beta\ell} \tag{4.42}$$

where τ_{k-1}^3 denotes the beginning of Step 3 (in the k-*th* attempt). where $c_3 \geqslant 2c_2$, $\beta \in [1,+\infty)$ and $(a_k)_{k\geqslant 1}$ is an increasing deterministic sequence. These quantities will be calibrate these quantities later. At this stage, we can however remark that conditionally to $F_{k,\ell}$, the length of each step is deterministic. During this waiting time, one certainly assumes that

$$g_{\scriptscriptstyle W}(t) = 0 \quad \text{on } [\tau_{k-1}^3,\tau_k], \quad \text{i.e.} \quad W_t - \widetilde{W}_{\tau_{k-1}^3} = W_t - \widetilde{W}_{\tau_{k-1}^3}. \tag{4.43}$$

4.3 Heuristic description of the coupling system in Step 1

At time τ_{k-1}, set $a_0 := Y_{\tau_{k-1}}$ and $a_1 = \widetilde{Y}_{\tau_{k-1}}$. As reported in the previous section, our aim in Step 1 (the *hitting* step) will be to build g_X on the interval $[\tau_{k-1}, \tau_{k-1}+1]$ in such a way that $Y_{\tau_{k-1}+1} = \widetilde{Y}_{\tau_{k-1}+1}$ with strictly positive probability. This construction will actually be the topic of both Sections 5 and 6. However, let us try here to give an idea, at some heuristic level, of the motivations behind this forthcoming strategy.

To this end, let us simplify the framework by assuming that $\tau_{k-1} = 0$, $d = 1$, and consider for the moment the case of a smooth deterministic driver x. In brief, our purpose is to exhibit a triplet of paths $(y^0_t, y^1_t, g_t)_{t\in[0,1]}$ satisfying the system

$$\begin{cases} dy^0_t = b(y^0_t)\,dt + \sigma(y^0_t)\,dx_t \\ dy^1_t = b(y^1_t)\,dt + \sigma(y^1_t)\,\big(dx_t + g_t\,dt\big) \end{cases},$$

as well as the constraints $y^0_0 = a_0$, $y^1_0 = a_1$ and $y^0_1 = y^1_1$. Using our invertibility assumptions on σ (that is, Hypothesis (**H3**)), this amounts to finding $(y^0_t, y^1_t, h_t)_{t\in[0,1]}$ such that

$$\begin{cases} dy^0_t = b(y^0_t)\,dt + \sigma(y^0_t)\,dx_t \\ dy^1_t = b(y^1_t)\,dt + \sigma(y^1_t)\,dx_t + h_t\,dt \end{cases},$$

and $y^0_0 = a_0$, $y^1_0 = a_1$, $y^0_1 = y^1_1$. In fact, let us consider the slightly more general issue of exhibiting a family of paths $(y^\xi_t, h^\xi_t)_{t\in[0,1],\xi\in[0,1]}$ that satisfy the equation

$$dy^\xi_t = b(y^\xi_t)\,dt + \sigma(y^\xi_t)\,dx_t + h^\xi_t\,dt \tag{4.44}$$

as well as the constraints $y^\xi_0 = a_0 + \xi(a_1 - a_0)$, $y^0_1 = y^1_1$ and $h^0_\cdot \equiv 0$. Then, using the basic identity $y^0_1 - y^1_1 = \int_0^1 d\xi\,\partial_\xi y^\xi_1$, we are led to the following sufficient formulation of the problem: finding a family $(y^\xi_t, h^\xi_t)_{t\in[0,1],\xi\in[0,1]}$ that satisfies both (4.44) and the constraints $y^\xi_0 = a_0 + \xi(a_1 - a_0)$, $\partial_\xi y_1 \equiv 0$, $h^0_\cdot \equiv 0$.

A natural way to answer the latter question is to let the so-called tangent path (associated with y) come into the picture. Namely, set $h^\xi_t := -\int_0^\xi d\eta\,\jmath^\eta_t$, where for each ξ, $\jmath^\xi$ stands for the solution of the equation

$$d\jmath^\xi_t = b'(y^\xi_t)\jmath^\xi_t\,dt + \sigma'(y^\xi_t)\jmath^\xi_t\,dx_t \quad , \quad \jmath^\xi_0 = a_1 - a_0\ .$$

With this specific choice of h^ξ_t, it is readily checked that the two paths $t \mapsto \partial_\xi y^\xi_t$ and $t \mapsto \jmath^\xi_t(1-t)$ satisfy the very same equation

$$dz^\xi_t = \big[b'(y^\xi_t)z^\xi_t - \jmath^\xi_t\big]\,dt + \sigma'(y^\xi_t)z^\xi_t\,dx_t\ .$$

Accordingly, $\partial_\xi y^\xi_t = \jmath^\xi_t(1-t)$ and the above constraints $\partial_\xi y_1 \equiv 0$, $h^0_\cdot \equiv 0$ are indeed satisfied.

As a conclusion of these considerations, the problem now reduces to solving the parametric (or functional-valued) system

$$\begin{cases} dy^\xi_t = \big[b(y^\xi_t) - \int_0^\xi d\eta\,\jmath^\eta_t\big]\,dt + \sigma(y^\xi_t)\,dx_t \\ d\jmath^\xi_t = b'(y^\xi_t)\jmath^\xi_t\,dt + \sigma'(y^\xi_t)\jmath^\xi_t\,dx_t \end{cases}, \tag{4.45}$$

with initial conditions $y^\xi_0 = a_0 + \xi(a_1 - a_0)$, $\jmath^\xi_0 = a_1 - a_0$, and ξ varying in $[0,1]$. This new expression of the problem can of course be extended to the case of rough x (and to any dimension d), with $(y^\xi_t, \jmath^\xi_t)$ understood as a rough solution of (4.45), in the sense of Definition 2.3. Just as above, once endowed with such a solution $(y, \jmath)$, defining g_X as $g_X(t) := -\sigma(y^1_t)^{-1}\int_0^1 d\eta\,\jmath^\eta_t$ would then close the procedure.

Unfortunately, as the reader can easily check it, the vector fields involved in (4.45) do not meet the usual boundedness assumptions that guarantee the existence of a *global* rough solution defined on $[0,1]$ (compare for instance with the conditions in [6, Theorem 6.1] or in [12, Theorem 10.26]). In fact, we have not been able to establish such a global existence in the general situation, and we even suspect that an explosion

phenomenon might occur in some cases. What we will prove in the next section is a weaker result according to which global existence on $[0,1]$ holds provided the norm $\|\mathbf{x}\|_{\gamma;[0,1]}$ is small enough (depending on a_0 and a_1).

Going back to the stochastic setting of our study (where $x = X$ is a fBm), the latter existence result is still not satisfying though, because the required smallness condition on $\|\mathbf{X}\|_{\gamma;[0,1]}$ implicitly involves the past trajectory $(W_t)_{t\leqslant 0}$, which is somehow fixed (and not necessarily small) at this stage of our three-step procedure. In order to overcome this difficulty, we shall appeal (once again) to the 'past-innovation' decomposition (4.26) of X, and rewrite the *hitting* system as

$$\begin{cases} dY_t^\xi = \big[b(Y_t^\xi)\,dt - \int_0^\xi d\eta\, J_t^\eta\, dt + \sigma(Y_t^\xi)\,dD_t\big]\,dt + \sigma(Y_t^\xi)\,dZ_t \\ dJ_t^\xi = \big[b'(Y_t^\xi)J_t^\xi\,dt + \sigma'(Y_t^\xi)J_t^\xi\,dD_t\big] + \sigma'(Y_t^\xi)J_t^\xi\,dZ_t \end{cases}, \tag{4.46}$$

with initial conditions $Y_0^\xi = a_0 + \xi(a_1 - a_0)$, $J_0^\xi = a_1 - a_0$, $\xi \in [0,1]$. Now recall that, at the sole price of a singularity at time 0, the process D can be considered as smooth and therefore the whole terms into brackets in (4.46) can somehow be seen as drift terms, to be distinguished from the real rough perturbation driven by Z. Based on these properties and still using a pathwise approach (as developed in Section 5 below), we will derive the following refined version of the previous existence statement: there exists a unique global solution to (4.46) on $[0,1]$ provided the norm of $\|\mathbf{Z}\|_{\gamma;[0,1]}$ (which no longer depends on $(W_t)_{t\leqslant 0}$) is small enough. This result, essentially summed up by Proposition 6.2, will turn out to be sufficient for our purpose.

5 Singular rough equations

This section is devoted to the presentation of a natural setting to study the hitting system (4.46) (properly extended to $\mathbb{R}^d$) and exhibit sharp conditions on (D,Z) for this system to admit a unique global solution on $[0,1]$. To this end, it will turn out to be fundamental that the trajectories of the process D should somehow be considered as differentiable paths that induce some drift term in the equation. However, as we have evoked it in Section 4.1 (see also Lemma 6.5 below), this differentiability assumption is not exactly satisfied, due to a possible singularity at time 0 for the derivative of D.

With these observations in mind, the purpose of the section is essentially twofold:

• Introduce appropriate *singular* extensions of the Hölder spaces defined in Section 2.1 and then extend the classical study of rough systems to this setting, for general Banach-valued equations;

• Exhibit sufficient conditions on the driver for these rough singular equations to have a unique global solution, even in situations where the usual boundedness requirements on the vector fields are not met (see Hypothesis (VF2) below), which is the case for (4.46).

The effective application of these general (pathwise) considerations to the particular fractional system (4.46) will then be analyzed in Section 6.1.

From now on and for the rest of the section, we fix two parameters: $\gamma \in (\frac{1}{3}, \frac{1}{2})$ (for the general Hölder roughness) and $\beta \in [\gamma, 1]$ (encoding the singularity at time 0).

5.1 Singular rough solutions and well-posedness results

Our singular extensions of the usual Hölder spaces are specifically defined through the following seminorms: given a Banach space V, an interval $I \subset [0,1]$ and two parameters $\alpha \in (0,1], \mu \geqslant \alpha$, set, for any map $f : I^2 \to V$, resp. $f : I^3 \to V$,

$$\mathcal{N}[f; \mathcal{C}_{2;\beta}^{\alpha,\mu}(I;V)] := \max\left(\sup_{s<t\in I} \frac{\|f_{st}\|_V}{|t-s|^\alpha}, \sup_{0<s<t\in I} \frac{\|f_{st}\|_V}{|t-s|^\mu\, s^{\beta-1}} \right), \tag{5.47}$$

resp.

$$\mathcal{N}[f; \mathcal{C}_{3;\beta}^{\alpha,\mu}(I;V)] := \max\left(\sup_{s<u<t\in I} \frac{\|f_{sut}\|_V}{|t-s|^\alpha}, \sup_{0<s<u<t\in I} \frac{\|f_{sut}\|_V}{|t-s|^\mu\, s^{\beta-1}} \right), \tag{5.48}$$

and then

$$\mathcal{C}^{\alpha,\mu}_{1;\beta}(I;V) := \big\{ f \in \mathcal{C}_1(I;V) : \ \delta f \in \mathcal{C}^{\alpha,\mu}_{2;\beta}(I;V) \big\} \ . \tag{5.49}$$

Of course, it holds that $\mathcal{C}^{\alpha,\mu}_{i;\beta}(I;V) \subset \mathcal{C}^{\alpha}_{i}(I;V)$ and $\mathcal{C}^{\alpha,\mu}_{i;1}(I;V) = \mathcal{C}^{\mu}_{i}(I;V)$. What actually led us to the above definitions is the following readily-checked (and relatively sharp) inclusion:

Lemma 5.1. *It holds that $\mathcal{E}^1_\gamma([0,1];\mathbb{R}^d) \subset \mathcal{C}^{\gamma,1}_{1;\gamma}([0,1];\mathbb{R}^d)$ and for every $g \in \mathcal{E}^1_\gamma([0,1];\mathbb{R}^d)$,*

$$\mathcal{N}[\delta g; \mathcal{C}^{\gamma,1}_{2;\gamma}([0,1];\mathbb{R}^d)] \leqslant c_\gamma |\!|\!| g |\!|\!|_{1;\gamma} \ . \tag{5.50}$$

Let us now introduce the related notion of a singular rough solution. In the sequel, given two Banach spaces V, W and a smooth map $F : V \to W$, we will denote by $D^{(\ell)}F : V \to \mathcal{L}(V^{\otimes \ell}; W)$ the ℓ-th derivative of F, understood in the usual Fréchet sense.

Definition 5.2. *Consider a path $h \in \mathcal{C}^{\gamma,1}_{1;\beta}([0,1];\mathbb{R}^m)$ and a γ-rough path $\mathbf{z} = (z, \mathbf{z}^{\mathbf{2}})$, in the sense of Definition 2.1. Then, for any fixed Banach space V, any interval $I = [t_0, t_1] \subset [0,1]$, any $v_0 \in V$ and all smooth vector fields*

$$B : V \to \mathcal{L}(\mathbb{R}^m; V) \quad , \quad \Sigma : V \to \mathcal{L}(\mathbb{R}^n; V) \ ,$$

we call $y \in \mathcal{C}^\gamma_1(I;V)$ a solution (on I) of the equation

$$dy_t = B(y_t)\, dh_t + \Sigma(y_t)\, d\mathbf{z}_t \quad , \quad y_{t_0} = v_0 \ , \tag{5.51}$$

if the two-parameter path R^y defined as

$$R^y_{st} := (\delta y)_{st} - B_i(y_s)\, (\delta h^i)_{st} - \Sigma_j(y_s)\, (\delta z^j)_{st} - (D\Sigma_j \cdot \Sigma_k)(y_s)\, \mathbf{z}^{\mathbf{2},jk}_{st}$$

belongs to $\mathcal{C}^{\gamma,\mu}_{2;\beta}(I;V)$, for some parameter $\mu > 1$. Here, the notation $D\Sigma_j \cdot \Sigma_k$ stands for

$$(D\Sigma_j \cdot \Sigma_k)(v) := (D\Sigma_j)(v)(\Sigma_k(v)) \ , \ \textit{for every } v \in V \ .$$

Remark 5.3. We are aware that the space $\mathcal{E}^1_\gamma$ could also be continuously embedded into the space of paths with finite 1-variation, so that the whole problem could certainly receive an analog treatment (with h still considered as inducing a drift term) in the p-variation setting used in [6] or in [12], instead of our singular Hölder setting. Nevertheless, switching the equation to a p-variation framework could expose us to the risk of a loss of topological sharpness in the results, with solutions possibly leaving the space of Hölder paths (see for instance the general definition of a solution in [6, Definition 3.1]). This is not the case in the above formulation, where the solution is still expected to belong to $\mathcal{C}^\gamma_1$.

Let us now turn to the presentation of the main results of this section about existence/uniqueness of a solution for the rough singular equation (5.51). We will either be concerned with the classical situation of bounded vector fields (Hypothesis (VF1)) or the more general possibility of linear growth (Hypothesis (VF2)).

Hypothesis (VF1). The vector field Σ and all its derivatives are uniformly bounded on V. Besides, the derivative $D^{(1)}B : V \to \mathcal{L}(V; \mathcal{L}(\mathbb{R}^m; V))$ is uniformly bounded on V.

Hypothesis (VF2). The following bounds on B and Σ hold true: for all $\ell \geqslant 0$,

$$\|(D^{(\ell)}B)(v)\| \lesssim 1 + \|v\| \quad , \quad \|(D^{(\ell)}\Sigma)(v)\| \lesssim 1 + \|v\| \ , \tag{5.52}$$

and also, for every $v, w \in V$,

$$\|(D\Sigma \cdot \Sigma)(v)\| \lesssim 1 + \|v\| \quad , \quad \|(D\Sigma \cdot \Sigma)(v) - (D\Sigma \cdot \Sigma)(w)\| \lesssim \|v - w\| \{1 + \|v\|\} \ . \tag{5.53}$$

Theorem 5.4 ((VF1)-situation)**.** *Under Hypothesis (VF1), and for any $v_0 \in V$, Equation (5.51) admits a unique solution on $[0,1]$ with initial condition v_0, in the sense of Definition 5.2.*

Theorem 5.5 ((VF2)-situation)**.** *Under Hypothesis (VF2), the following assertions hold true:*

(i) For any $v_0 \in V$, Equation (5.51) admits at most one solution on $[0,1]$ with initial condition v_0, in the sense of Definition 5.2.

(ii) For every $K \geqslant 1$, there exists $M_K > 0$ such that if $\|v_0\| \leqslant K$, $\mathcal{N}[\delta h; \mathcal{C}^{\gamma;1}_{2;\beta}([0,1])] \leqslant K$ and $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$, then Equation (5.51) admits a unique solution y on $[0,1]$ with initial condition v_0, in the sense of Definition 5.2. Besides,

$$\mathcal{N}[y; \mathcal{C}^0_1([0,1];V)] + \mathcal{N}[y; \mathcal{C}^\gamma_1([0,1];V)] \leqslant C(K) \ , \tag{5.54}$$

for some growing function $C : \mathbb{R}^+ \to \mathbb{R}^+$.

Just as in Section 3.2, and in the same spirit as in [6], our proof for both Theorem 5.4 and Theorem 5.5 relies on the examination of the discrete scheme associated with the equation. Set $t_i = t^n_i := \frac{i}{2^n}$, $\mathcal{P}_n := \{t_i : \ i = 0, \dots, 2^n\}$ and define y^n along the iterative formula: $y^n_0 = v_0$ and

$$(\delta y^n)_{t_i t_{i+1}} = B(y^n_{t_i})\,(\delta h)_{t_i t_{i+1}} + \Sigma(y^n_{t_i})\,(\delta z)_{t_i t_{i+1}} + (D\Sigma \cdot \Sigma)(y^n_{t_i})\,\mathbf{z}^{\mathbf{2}}_{t_i t_{i+1}} \ .$$

Then, for every $s, t \in \mathcal{P}_n$, set

$$R^n_{st} := (\delta y^n)_{st} - B(y^n_s)\,(\delta h)_{st} - \Sigma(y^n_s)\,(\delta z)_{st} - (D\Sigma \cdot \Sigma)(y^n_s)\,\mathbf{z}^{\mathbf{2}}_{st} \ ,$$

noting in particular that $R^n_{t_i t_{i+1}} = 0$. We will also consider the paths

$$L^n_{st} := (\delta y^n)_{st} - \Sigma(y^n_s)\,(\delta z)_{st} - \ (D\Sigma \cdot \Sigma)(y^n_s)\,\mathbf{z}^{\mathbf{2}}_{st} \quad \big(= R^n_{st} + B(y^n_s)(\delta h)_{st} \ \big)$$

and

$$Q^n_{st} := (\delta y^n)_{st} - \Sigma(y^n_s)\,(\delta z)_{st} \ .$$

Finally, for every $s < t \in [0,1]$, we will write $[\![s,t]\!] = [\![s,t]\!]_n := [s,t] \cap \mathcal{P}_n$, and we extend (or rather restrict) the norms (5.47)-(5.49) to discrete paths as

$$\mathcal{N}[f; \mathcal{C}^{\alpha,\mu}_{2;\beta}([\![s,t]\!];V)] := \max\left(\sup_{u<v\in[\![s,t]\!]} \frac{\|f_{uv}\|_V}{|v-u|^\alpha}, \sup_{0<u<v\in[\![s,t]\!]} \frac{\|f_{uv}\|_V}{|v-u|^\mu\, u^{\beta-1}} \right) ,$$

with a similar definition for $\mathcal{N}[f; \mathcal{C}^{\alpha,\mu}_{i;\beta}([\![s,t]\!];V)]$, $i \in \{1,3\}$.

The whole key towards the desired estimates lies in the following "singular sewing lemma":

Lemma 5.6. *Let $0 < \alpha \leqslant \lambda \leqslant 1$, $\mu_1 \geqslant 1$ and $\mu_2 > 1$. Then there exists a constant $c_{\alpha,\lambda,\mu_1,\mu_2} > 0$ such that for every path $G : \mathcal{P}^2_n \to V$ and all $s \leqslant t \in \mathcal{P}_n$, one has*

$$\mathcal{N}[G; \mathcal{C}^{\alpha,\mu_1\wedge\mu_2}_{2;\lambda}([\![s,t]\!];V)] \leqslant c_{\alpha,\lambda,\mu_1,\mu_2}\big\{\mathcal{M}^{\alpha,\mu_1}_\lambda\big[G;[\![s,t]\!]\big] + \mathcal{N}[\delta G; \mathcal{C}^{\alpha,\mu_2}_{3;\lambda}([\![s,t]\!];V)]\big\} \ ,$$

where we have set, if $s = \frac{p}{2^n}$ and $t = \frac{q}{2^n}$,

$$\mathcal{M}^{\alpha,\mu_1}_\lambda\big[G;[\![s,t]\!]\big] := \max\left(\sup_{p\leqslant i\leqslant q} \frac{\|G_{t_i t_{i+1}}\|}{|t_{i+1}-t_i|^\alpha} \ , \ \sup_{p+1\leqslant i\leqslant q} \frac{\|G_{t_i t_{i+1}}\|}{|t_{i+1}-t_i|^{\mu_1} t_i^{\lambda-1}} \right) .$$

Proof. See Appendix B. □

5.2 Existence of a solution in the (VF2)-situation

Proposition 5.7. *Let Hypothesis (VF2) prevail and assume additionally that*

$$\mathcal{N}[\delta h; \mathcal{C}^{\gamma;1}_{2;\beta}([0,1];V)] \leqslant K \ , \ \textit{for some } K \geqslant 1 \ .$$

Then there exists a constant c_0 (which depends only on B, Σ, γ and β) such that if we set $T_0 = T_0(K) := \min\big(1, (c_0 K)^{-6/(3\gamma-1)}\big)$, the following assertion holds true for every $k \leqslant 1/T_0$: if $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant \big(1 + \|y^n_{kT_0}\|\big)^{-1}$, then

$$\mathcal{N}[L^n; \mathcal{C}^{\gamma;1}_{2;\beta}([\![kT_0, (k+1)T_0 \wedge 1]\!];V)] \leqslant c_0\, K\big\{1 + \|y^n_{kT_0}\|\big\} \ . \tag{5.55}$$

Proof. The strategy consists in an iteration procedure over the points of the partition. So, assume that (5.55) holds true on an interval $[\![0,t_q]\!]$, with $t_q \leqslant T_0$ (for some time T_0 to be determined along the proof). In other words, assume that

$$\mathcal{N}[L^n;\mathcal{C}^{\gamma,1}_{2;\beta}([\![0,t_q]\!];V)] \leqslant c_K\big\{1+\|y^n_0\|\big\}\ , \tag{5.56}$$

where we denote from now on $c_K := c_0K$ (for some constant c_0 to be fixed later on). Due to (5.52) and (5.53), it is then easy to check that the following bounds hold true as well:

$$\mathcal{N}[y^n;\mathcal{C}^0_1([\![0,t_q]\!])] \lesssim \{1+\|y^n_0\|\}\{1+c_KT_0^\gamma\}\ , \tag{5.57}$$

and

$$\max\big(\mathcal{N}[y^n;\mathcal{C}^\gamma_1([\![0,t_q]\!])],\mathcal{N}[Q^n;\mathcal{C}^{\gamma;2\gamma}_{2;\beta}([\![0,t_q]\!])]\big) \lesssim \{1+\|y^n_0\|\}\{1+c_K\}\ . \tag{5.58}$$

Now, in order to extend (5.56) on $[\![0,t_{q+1}]\!]$ (assuming that $t_{q+1}\leqslant T_0$), let us first apply Lemma 5.6 to L^n and assert that

$$\mathcal{N}[L^n;\mathcal{C}^{\gamma,1}_{2;\beta}([\![0,t_{q+1}]\!])] \lesssim \mathcal{M}^{\gamma,1}_\beta\big[L^n;[\![0,t_{q+1}]\!])\big] + \mathcal{N}[\delta L^n;\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![0,t_{q+1}]\!])]\ , \tag{5.59}$$

where we set from now on $\kappa := \frac{1}{2}\big(\frac{1}{3}+\gamma\big)$, so that $1<3\kappa<3\gamma$.

As far as the first term is concerned, we can use the fact that $R^n_{t_it_{i+1}}=0$, and then (5.52) and (5.57), to deduce that

$$\begin{aligned}\mathcal{M}^{\gamma,1}_\beta\big[L^n;[\![0,t_{q+1}]\!])\big] &= \mathcal{N}[B(y^n)\,\delta h;\mathcal{C}^{\gamma,1}_{2;\beta}([\![0,t_{q+1}]\!])]\\ &\leqslant K\cdot\mathcal{N}[B(y^n);\mathcal{C}^0_1([\![0,t_q]\!]] \lesssim K\{1+\|y^n_0\|\}\{1+c_KT_0^\gamma\}\ .\end{aligned}$$

In order to estimate $\mathcal{N}[\delta L^n;\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![0,t_{q+1}]\!])]$, let us first rely on Chen relation and decompose the increments of L^n as $\delta L^n = I^i\,\delta z^i + II^i\,\delta z^i + III^{ij}\,\mathbf{z}^{\mathbf{2},ij}$, where we have set

$$I^i_{st} := \int_0^1 d\lambda\,(D\Sigma_i)(y^n_s+\lambda(\delta y^n)_{st})\,Q^n_{st}\ , \tag{5.60}$$

$$II^i_{st} := \int_0^1 d\lambda\,\big[(D\Sigma_i)(y^n_s+\lambda(\delta y^n)_{st})-(D\Sigma_i)(y^n_s)\big]\,\Sigma_j(y^n_s)\,(\delta z^j)_{st}\ , \tag{5.61}$$

$$III^{ij}_{st} := \delta(D\Sigma_i\cdot\Sigma_j)(y^n)_{st}\ . \tag{5.62}$$

For $I^i\,\delta z^i$, we can combine (5.52), (5.57) and (5.58) to get that

$$\begin{aligned}\mathcal{N}[I^i\,\delta z^i;\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![0,t_{q+1}]\!])] &\lesssim T_0^{3(\gamma-\kappa)}\big(\{1+\|y^n_0\|\}\{1+c_K\}\big)^2\,\|\mathbf{z}\|_{\gamma;[0,1]}\\ &\lesssim \{1+\|y^n_0\|\}\{1+T_0^{3(\gamma-\kappa)}c_K^2\}\cdot\Big(\{1+\|y^n_0\|\}\|\mathbf{z}\|_{\gamma;[0,1]}\Big)\\ &\lesssim \{1+\|y^n_0\|\}\{1+T_0^{3(\gamma-\kappa)}c_K^2\}\ ,\end{aligned}$$

where we have used the assumption $\{1+\|y^n_0\|\}\|\mathbf{z}\|_{\gamma;[0,1]}\leqslant 1$ to derive the third inequality.

With similar arguments, we can show that

$$\begin{aligned}\mathcal{N}[II^i\,\delta z^i;\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![0,t_{q+1}]\!])] &\lesssim \{1+\|y^n_0\|\}\{1+T_0^{3(\gamma-\kappa)}c_K^3\}\cdot\Big(\{1+\|y^n_0\|\}^2\|\mathbf{z}\|^2_{\gamma;[0,1]}\Big)\\ &\lesssim \{1+\|y^n_0\|\}\{1+T_0^{3(\gamma-\kappa)}c_K^3\}\ .\end{aligned}$$

Finally, thanks to the second estimate in (5.53), we obtain that

$$\mathcal{N}[III^{ij}\,\mathbf{z}^{\mathbf{2},ij};\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![0,t_{q+1}]\!])] \lesssim \{1+\|y^n_0\|\}\{1+T_0^{3(\gamma-\kappa)}c_K^2\}\ .$$

Going back to (5.59), we have shown that, for some constant c_1 depending only on B, Σ and (γ,κ,β),

$$\mathcal{N}[L^n;\mathcal{C}^{\gamma,1}_{2;\beta}([\![0,t_{q+1}]\!])] \leqslant \{1+\|y^n_0\|\}\cdot\Big(c_1K\{1+T_0^{3(\gamma-\kappa)}c_K^3\}\Big)\ .$$

Let us now set $c_0 := 2c_1$, $c_K := c_0 K$ and $T_0 := \min(1, (2c_1 K)^{-1/(\gamma-\kappa)})$, in such a way that

$$c_1 K\{1 + T_0^{3(\gamma-\kappa)} c_K^3\} \leqslant c_K \ ,$$

and accordingly $\mathcal{N}[L^n; \mathcal{C}^{\gamma,1}_{2;\beta}([\![0, t_{q+1}]\!])] \leqslant c_K\{1 + \|y_0^n\|\}$ as desired.

This iteration procedure allows us to extend the bound (5.56) over the interval $[\![0, T_0]\!]$. Then it is easy to see that the very same arguments can be used for any interval $[\![kT_0, (k+1)T_0]\!]$, which completes the proof. □

Corollary 5.8. *Let Hypothesis (VF2) prevail and assume additionally that*

$$\mathcal{N}[\delta h; \mathcal{C}^{\gamma,1}_{2;\beta}([0,1]; V)] \leqslant K \text{ and } \|v_0\| \leqslant K \ , \text{ for some } K \geqslant 1 \ .$$

Then there exists $M_K > 0$ such that if $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$, one has

$$\sup_{n\geqslant 0} \max\left(\mathcal{N}[y^n; \mathcal{C}^0_1([\![0,1]\!])], \mathcal{N}[y^n; \mathcal{C}^\gamma_1([\![0,1]\!])], \mathcal{N}[Q^n; \mathcal{C}^{\gamma,2\gamma}_{2;\beta}([\![0,1]\!])], \mathcal{N}[L^n; \mathcal{C}^{\gamma,1}_{2;\beta}([\![0,1]\!])]\right) \leqslant C(K) \ , \quad (5.63)$$

for some growing function $C : \mathbb{R}^+ \to \mathbb{R}^+$. As a result, under the same assumptions and if $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$, it holds that

$$\sup_{n\geqslant 0} \mathcal{N}[R^n; \mathcal{C}^{\gamma,3\gamma}_{2;\beta}([\![0,1]\!])] \leqslant D(K) \ , \quad (5.64)$$

for some growing function $D : \mathbb{R}^+ \to \mathbb{R}^+$.

Proof. Using (5.55) as well as its spin-offs (5.57) and (5.58), it is not hard to exhibit a growing sequence (c_k) that depends only on $(B, \Sigma, \gamma, \beta)$ (and not on K) such that the following property holds true: for every $k \geqslant 0$, if $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant (1 + c_k\{1+K\})^{-1}$, then one has both

$$\mathcal{N}[y^n; \mathcal{C}^0_1(I_k)] \leqslant c_{k+1}\{1+K\} \quad (5.65)$$

and

$$\max\left(\mathcal{N}[y^n; \mathcal{C}^\gamma_1(I_k)], \mathcal{N}[Q^n; \mathcal{C}^{\gamma,2\gamma}_{2;\beta}(I_k)], \mathcal{N}[L^n; \mathcal{C}^{\gamma,1}_{2;\beta}(I_k)]\right) \leqslant c_{k+1}\{1+K^2\} \ , \quad (5.66)$$

where we have set $I_k := [\![kT_0, (k+1)T_0]\!]$. As a result, if we denote by N_K the smallest integer such that $T_0 N_K \geqslant 1$ and assume that $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K := (1 + c_{N_K}(1+K))^{-1}$, then both bounds (5.65) and (5.66) hold true for $k = 0, \dots, N_K - 1$. The extension of these local bounds into global ones (that is, on the interval $[\![0,1]\!]$) is then a matter of standard arguments, which achieves the proof of (5.63).

As far as (5.64) is concerned, apply first Lemma 5.6 to the path R^n, which, since $R^n_{t_i t_{i+1}} = 0$, entails that

$$\mathcal{N}[R^n; \mathcal{C}^{\gamma,3\gamma}_{2;\beta}([\![0,1]\!]; V)] \lesssim \mathcal{N}[\delta R^n; \mathcal{C}^{\gamma,3\gamma}_{3;\beta}([\![0,1]\!]; V)]\} \ .$$

Then, just as in the proof of Proposition 5.7, observe that we can decompose the increments of R^n as

$$(\delta R^n)_{sut} = \delta B(y^n)_{su}(\delta h)_{ut} + (\delta L^n)_{sut} = \delta B(y^n)_{su}(\delta h)_{ut} + I^i_{su}\delta z^i_{ut} + II^i_{su}\delta z^i_{ut} + III^{ij}_{su}\mathbf{z}^{\mathbf{2},ij}_{ut} \ , \quad (5.67)$$

where the paths (I, II, III) have been defined through (5.60)-(5.62). The conclusion is now easy to derive from the bound (5.63).

□

Proof of Theorem 5.5, point (ii). Consider the sequence (still denoted by y^n) of continuous paths on $[0,1]$ defined through the linear interpolation of the points of the previous (discrete) sequence y^n. Define M_K as in Corollary 5.8 and assume that $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$. Then it is readily checked that (5.63) gives rise to a uniform bound for $\mathcal{N}[y^n; \mathcal{C}^\gamma_1([0,1]; V)]$, and we can therefore conclude about the existence of a path $y \in \mathcal{C}^\gamma_1([0,1]; V)$, as well as a subsequence of y^n (that we still denote by y^n), such that $y^n \to y$ in $\mathcal{C}^\kappa_1([0,1]; V)$ for every $0 < \kappa < \gamma$.

The fact that y actually defines a solution of (5.51) is essentially obtained by passing to the limit in the uniform estimate (5.64). The details of this (easy) procedure can for instance be found at the end of [7, Section 3.3]. As for the bound (5.54), it is a straightforward consequence of (5.63). □

5.3 Existence of a solution in the (VF1)-situation

Under Hypothesis (VF1), the exhibition of a uniform bound for $\mathcal{N}[R^n;\mathcal{C}^{\gamma,\mu}_{2;\beta}([\![0,1]\!])]$ (with $\mu>1$) essentially follows the same general procedure as in the classical ('non-singular') situation treated in [6] or [12]. As we here consider slightly more specific topologies, let us briefly review the result at the core of this procedure.

Proposition 5.9. *Let Hypothesis (VF1) prevail and assume additionally that*

$$\mathcal{N}[\delta h;\mathcal{C}^{\gamma,1}_{2;\beta}([0,1];\mathbb{R}^m)]\leqslant K\ ,\ \textit{for some } K\geqslant 1\ .$$

Also, fix a parameter κ such that $1<3\kappa<3\gamma$. Then there exists a a constant c_0 (which depends only on B, Σ, γ, β and κ) such that if we set

$$T_0=T_0(\|\mathbf{z}\|_\gamma,K):=\min\Big(1,\Big(c_0\big\{1+\|\mathbf{z}\|_\gamma\big\}K\Big)^{-1/(\gamma-\kappa)}\Big)\ ,$$

the following property holds true: for every $0<T_1<T_0$ and every $k\leqslant 1/T_1$,

$$\mathcal{N}[R^n;\mathcal{C}^{\gamma,3\kappa}_{2;\beta}([\![kT_1,(k+1)T_1\wedge 1]\!])]\leqslant c_0\,K\big\{1+\|y^n_{kT_1}\|\big\}\ . \tag{5.68}$$

Proof. Just as in the proof of Proposition 5.7, the strategy consists in an iteration procedure over the points of $\mathcal{P}_n$. The argument actually relies on the following two readily-checked assertions: (i) If $\mathcal{N}[R^n;\mathcal{C}^{\gamma,3\kappa}_{2;\beta}([\![s,t]\!])]\leqslant c_0\,K\{1+\|y^n_s\|\}$, then one has

$$\max\big(\mathcal{N}[y^n;\mathcal{C}^\gamma_1([\![s,t]\!])],\mathcal{N}[Q^n;\mathcal{C}^{\gamma,2\gamma}_{2;\beta}([\![s,t]\!])]\big)\leqslant c_1\,K\big[\|\mathbf{z}\|_\gamma+\big\{1+\|y^n_s\|\big\}\big\{1+c_0\big\}\big\{1+K|t-s|^\gamma\big\}\big]$$

for some constant c_1 that depends only on (B,Σ); (ii) With decomposition (5.67) in mind, one has

$$\mathcal{N}[\delta R^n;\mathcal{C}^{\gamma,3\kappa}_{3;\beta}([\![s,t]\!])]\leqslant c_2\,|t-s|^{3(\gamma-\kappa)}\big[\mathcal{N}[y^n;\mathcal{C}^\gamma_1([\![s,t]\!])]\{1+K+\|\mathbf{z}\|^2_\gamma\}+\mathcal{N}[Q^n;\mathcal{C}^{\gamma,2\gamma}_{2;\beta}([\![s,t]\!])]\|\mathbf{z}\|_\gamma\big]$$

for some constant c_2 that depends only on (B,Σ).

It is now easy to inject (i) and (ii) into the iteration scheme exhibited in the previous section for L^n (note that we can additionally use the fact that $R^n_{t_it_{i+1}}=0$ here). The details of the procedure are therefore left to the reader. □

5.4 Uniqueness of the solution

It is a well-known fact that uniqueness statements are usually less demanding than existence statements as far as global boundedness of the vector fields is concerned. Accordingly, in opposition with the previous existence proof (where specific sharp estimates had to be displayed), the strategy towards uniqueness essentially follows the same lines as in the standard situation. We briefly review the transposition of the main arguments in this singular setting.

Assume here that either Hypothesis (VF1) or Hypothesis (VF2) prevails and consider two solutions $U,\widetilde{U}$ of (5.51) with identical initial conditions. Then set

$$R_{st}=R(y)_{st}:=(\delta y)_{st}-B_i(y_s)\,(\delta h^i)_{st}-\Sigma_j(y_s)\,(\delta z^j)_{st}-(D\Sigma_j\cdot\Sigma_k)(y_s)\,\mathbf{z}^{\mathbf{2},jk}_{st}\ ,$$

$$Q_{st}=Q(y)_{st}:=(\delta y)_{st}-\Sigma_j(y_s)\,(\delta z^j)_{st}\ ,$$

and similarly $\widetilde{R}:=R(\widetilde{y})$, $\widetilde{Q}:=Q(\widetilde{y})$. Also, fix μ, resp. $\tilde{\mu}>1$ such that $\mathcal{N}[R;\mathcal{C}^{\gamma,\mu}_{2;\beta}([0,1])]<\infty$, resp. $\mathcal{N}[R;\mathcal{C}^{\gamma,\mu}_{2;\beta}([0,1])]<\infty$, as well as a parameter κ satisfying both $\frac{1}{3}<\kappa<\gamma$ and $3\kappa<\mu\wedge\tilde{\mu}$.

Lemma 5.10. *There exists a finite constant $c_{R,\widetilde{R}}>0$ such that for every $s<t\in\mathcal{P}_n$, one has*

$$\mathcal{N}[R-\widetilde{R};\mathcal{C}^{\kappa,3\kappa}_{2;\beta}([\![s,t]\!])]\leqslant c_{R,\widetilde{R}}\cdot\big\{2^{-n\varepsilon}+\mathcal{N}[\delta(R-\widetilde{R});\mathcal{C}^{\kappa,3\kappa}_{3;\beta}([\![s,t]\!])]\big\}\ , \tag{5.69}$$

where $\varepsilon:=\inf(\gamma-\kappa,(\mu\wedge\tilde{\mu})-3\kappa)>0$.

Proof. It is a mere application of Lemma 5.6. Observe indeed that

$$\begin{aligned}\mathcal{M}_\beta^{\kappa,3\kappa}\big[R-\widetilde{R};[\![s,t]\!]\big] &\leqslant \mathcal{M}_\beta^{\kappa,3\kappa}\big[R;[\![s,t]\!]\big]+\mathcal{M}_\beta^{\kappa,3\kappa}\big[\widetilde{R};[\![s,t]\!]\big]\\ &\leqslant c_{R,\widetilde{R}}\,\{2^{-n(\gamma-\kappa)}+2^{-n(\mu_1-3\kappa)}+2^{-n(\mu_2-3\kappa)}\}\ .\end{aligned}$$

□

Lemma 5.11. *There exists a finite constant $C_{y,\widetilde{y}}>0$ such that for every $s<t\in\mathcal{P}_n$, one has*

$$\mathcal{N}[\delta(R-\widetilde{R});\mathcal{C}_{3;\beta}^{\kappa,3\kappa}([\![s,t]\!])]\leqslant C_{y,\widetilde{y}}\,|t-s|^{\gamma-\kappa}\,\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y,\widetilde{y});[\![s,t]\!]\big]\ , \tag{5.70}$$

where we have set

$$\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y,\widetilde{y});[\![s,t]\!]\big]:=\mathcal{N}[y-\widetilde{y};\mathcal{C}_1^0([\![s,t]\!])]+\mathcal{N}[y-\widetilde{y};\mathcal{C}_1^\kappa([\![s,t]\!])]+\mathcal{N}[Q-\widetilde{Q};\mathcal{C}_{2;\beta}^{\kappa,2\gamma}([\![s,t]\!])]\ . \tag{5.71}$$

Proof. First, note that the increments of R (or $\widetilde{R}$) can be decomposed just as the increments of R^n in the proof of Corollary 5.8 (see (5.67)), which allows us to write

$$\delta(R-\widetilde{R})_{sut}=\delta(B(y)-B(\widetilde{y}))_{su}\,\delta h_{ut}+\big[I^i_{su}-\widetilde{I}^i_{su}\big]\,\delta z^i_{ut}+\big[II^i_{su}-\widetilde{II}^i_{su}\big]\,\delta z^i_{ut}+\big[III^{ij}_{su}-\widetilde{III}^{ij}_{su}\big]\,\mathbf{z}^{\mathbf{2},ij}_{ut}\ ,$$

where the paths I,II,III, resp. $\widetilde{I},\widetilde{II},\widetilde{III}$, are defined along (5.60)-(5.62) (replace (y^n,Q^n) with (y,Q), resp. $(\widetilde{y},\widetilde{Q})$). The bound (5.70) is then obtained through standard differential-calculus arguments based on (5.52)-(5.53). □

Proof of Theorem 5.5, point (i), and uniqueness property of Theorem 5.4. Consider the above setting and notations. First, going back to the very definitions of (K,R) and $(\widetilde{K},\widetilde{R})$, it is not hard to check that for every $s<t\in\mathcal{P}_n$, one has, with the notation (5.71),

$$\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y,\widetilde{y});[\![s,t]\!]\big]\leqslant c_{y,\widetilde{y}}\Big\{\|y_s-\widetilde{y}_s\|+|t-s|^\kappa\,\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y;\widetilde{y});[\![s,t]\!]\big]+\mathcal{N}\big[R-\widetilde{R};\mathcal{C}_{2;\beta}^{\kappa,3\kappa}([\![s,t]\!]\big]\Big\}\ ,$$

where the constant $c_{y,\widetilde{y}}$ does not depend on n. We can then combine (5.69)-(5.69) and assert that for every $s<t\in\mathcal{P}_n$,

$$\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y,\widetilde{y});[\![s,t]\!]\big]\leqslant c_{y,\widetilde{y}}\Big\{\|y_s-\widetilde{y}_s\|+|t-s|^{\gamma-\kappa}\,\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y;\widetilde{y});[\![s,t]\!]\big]+2^{-n\varepsilon}\Big\}\ .$$

The uniqueness result is now immediate. Indeed, for $T_0>0$ such that $c_{y,\widetilde{y}}\,T_0^{\gamma-\kappa}\leqslant\frac{1}{2}$, and since $y_0=\widetilde{y}_0$, we first get that

$$\mathcal{N}\big[y-\widetilde{y};\mathcal{C}_1^0([\![0,T_0]\!])\big]\leqslant\mathcal{N}_\beta^{\kappa,2\gamma}\big[(y,\widetilde{y});[\![0,T_0]\!]\big]\leqslant C_{y,\widetilde{y}}\cdot 2^{-n\varepsilon}\ ,$$

and accordingly $y_t=\widetilde{y}_t$ for every $t\in[0,T_0]$. The argument can then be repeated on $[T_0,2T_0]$, $[2T_0,3T_0]$, and so on. □

6 Hitting step

Keeping in mind the strategy sketched out in Section 4.3, the route to Step 1, that is the hitting step, is now quite clear: we need to check that the vector fields involved in the hitting system (4.46) do satisfy the assumptions of the previous section, and then see how the conditions therein exhibited (for the driver) can be injected into the general coupling machinery.

From now on and for the rest of the section, we fix $H\in(\frac{1}{3},\frac{1}{2})$, $\gamma\in(\frac{1}{3},H)$, as well as vector fields $b:\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\to\mathcal{L}(\mathbb{R}^d,\mathbb{R}^d)$ satisfying Hypotheses (**H1**) and (**H3**) (note that Hypothesis (**H2**) is not required at this stage of the procedure).

6.1 Rough hitting

This first section focuses on the hitting issue at the level of the general (deterministic) rough system, and therefore it settles the bases for our forthcoming stochastic analysis. Let us fix two paths $h \in \mathcal{E}^1_\gamma([0,1];\mathbb{R}^d)$, $z \in \mathcal{C}^\gamma([0,1];\mathbb{R}^d)$, and assume that z can be canonically lifted into a γ-rough path $\mathbf{z} := \mathfrak{L}(z)$, in the sense of Definition 2.2.

Lemma 6.1. *Consider the Banach space $V_2 := \mathcal{W}^{1,\infty}([0,1];\mathbb{R}^d) \times L^\infty([0,1];\mathbb{R}^d)$, and define the vector fields (B,Σ) on V along the following formulas:*

$$B_0^i(y,\jmath)(\xi) := \begin{cases} b^i(y(\xi)) - \int_0^\xi d\eta\, \jmath^i(\eta) \\ (\partial_k b^i)(y(\xi))\jmath^k(\xi) \end{cases} \quad , \quad B_j^i(y,\jmath)(\xi) = \Sigma_j^i(y,\jmath)(\xi) := \begin{cases} \sigma_j^i(y(\xi)) \\ (\partial_k \sigma_j^i)(y(\xi))\jmath^k(\xi) \end{cases}$$

for $i,j = 1,\dots,d$. Then, under Hypothesis (**H1**)*, the pair (B,Σ) satisfies Hypothesis (VF2).*

Proof. We have the following explicit expressions:

$$(D\Sigma_j^i)(y,\jmath)(y_1,\jmath_1) = \begin{cases} (\partial_k \sigma_i^j)(y)\, y_1^k \\ (\partial_k \partial_\ell \sigma_j^i)(y)\, y_1^\ell\, \jmath^k + (\partial_k \sigma_i^j)(y)\, \jmath_1^k \end{cases} ,$$

and more generally, for every $q \geqslant 1$,

$$\begin{aligned} &(D^{(q)}\Sigma_j^i)(y,\jmath)((y_1,\jmath_1),\dots,(y_q,\jmath_q)) \\ &\qquad = \begin{cases} (\partial_{k_1}\cdots\partial_{k_q}\sigma_j^i)(y)\, y_1^{k_1}\cdots y_q^{k_q} \\ (\partial_k\partial_{k_1}\cdots\partial_{k_q}\sigma_j^i)(y)\, y_1^{k_1}\cdots y_q^{k_q}\,\jmath^k + (\partial_{k_1}\cdots\partial_{k_q}\sigma_j^i)(y)\sum_{r=1,\dots,q} y_1^{k_1}\cdots y_{r-1}^{k_{r-1}}\jmath_r^{k_r} y_{r+1}^{k_{r+1}}\cdots y_q^{k_q} \end{cases} . \end{aligned}$$

In particular,

$$(D\Sigma_j^i \cdot \Sigma_k)(y,\jmath) = \begin{cases} (\partial_\ell \sigma_j^i)(y)\sigma_k^\ell(y) \\ \big\{(\partial_m\partial_\ell\sigma_j^i)(y)\sigma_k^\ell(y) + (\partial_\ell\sigma_i^j)(y)(\partial_m\sigma_k^\ell)(y)\big\}\,\jmath^m \end{cases} .$$

Based on these formulas, the two conditions (5.52) and (5.53) for Σ are immediate. We can then exhibit a similar expression for $D^{(q)}B_0$.

□

Combining Lemma 6.1 with the well-posedness results of Theorems 5.4 and 5.5 yields the following statement:

Proposition 6.2. *Under Hypothesis* (**H1**)*, the following assertions hold true:*

(a) Let $V_1 := L^\infty([0,1];\mathbb{R}^d)^2$. Then for every $A \in V_1$ and every smooth function $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ bounded with bounded derivatives, the rough system

$$dy_t(\xi) = \Big[\varphi(b(y_t(\xi))) - \int_0^\xi \varphi(\jmath_t(\eta))\, d\eta\Big]\, dt + \sigma(y_t(\xi))\, dh_t + \sigma(y_t(\xi))\, d\mathbf{z}_t \ , \tag{6.72}$$

$$d\jmath_t(\xi) = (\partial_k b)(y_t(\xi))\varphi(\jmath_t(\xi))_k\, dt + (\partial_k\sigma)(y_t(\xi))\varphi(\jmath_t(\xi))_k\, dh_t + (\partial_k\sigma)(y_t(\xi))\varphi(\jmath_t(\xi))_k\, d\mathbf{z}_t \ , \tag{6.73}$$

with initial condition $(y_0,\jmath_0) = A$, admits a unique solution

$$(y,\jmath) =: \Psi_{V_1}(A,\varphi,(h,\mathbf{z})) \in \mathcal{C}_1^\gamma([0,1];V_1) \ ,$$

in the sense of Definition 5.2.

(b) Let $V_2 := \mathcal{W}^{1,\infty}([0,1];\mathbb{R}^d) \times L^\infty([0,1];\mathbb{R}^d)$ and pick $A \in V_2$. Assume that $\|A\|_{V_2} \leqslant K$ and $|||h|||_{1;\gamma} \leqslant K$, for some fixed $K \geqslant 1$. Then there exists a constant $M_K > 0$ such that if $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$, the conclusion of point (i) is still true for $\varphi \equiv Id$ and V_1 replaced with V_2, and one has

$$\mathcal{N}\big[\Psi_{V_2}(A,Id,(h,\mathbf{z}));\mathcal{C}_1^0([0,1];V_2)\big] \leqslant C(K) \ , \tag{6.74}$$

for some growing function $C : \mathbb{R}^+ \to \mathbb{R}^+$.

Proof. Point (b) is obtained through the combination of Lemma 6.1 and Theorem 5.5. As for point (a), it suffices to observe that for every fixed φ, the vector fields involved in (6.72)-(6.73) satisfy Hypothesis (VF1), and we can therefore appeal to Theorem 5.4 to conclude in this case.

□

Let us now rigourously check that when $\varphi \equiv \mathrm{Id}$, the hitting system (6.72)-(6.73) indeed satisfies the desired property, namely offering a way to see two rough solutions (with different initial conditions and drivers differing from a sole drift term) hit a time 1.

Proposition 6.3. *Let $V_2 := \mathcal{W}^{1,\infty}([0,1];\mathbb{R}^d) \times L^\infty([0,1];\mathbb{R}^d)$ and consider a V_2-valued solution $(y, \jmath)$ on $[0,1]$ (in the sense of Definition 5.2) of the rough system (6.72)-(6.73) with $\varphi \equiv Id$ and inital condition*

$$y_0(\xi) = (1-\xi)a_0 + \xi a_1 \quad , \quad \jmath_0(\xi) = a_1 - a_0 \ ,$$

for fixed $a_0, a_1 \in \mathbb{R}^d$. Then the following assertions hold true:

(a) The $\mathbb{R}^d$-valued path $y^{(0)} := y_\cdot(0)$, is the solution on $[0,1]$ (in the sense of Definition 2.3) of the rough equation

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d\mathbf{x}_t \quad , \quad y_0 = a_0 \ , \tag{6.75}$$

where $\mathbf{x}$ is the canonical rough path above $x := z + h$.

(b) The $\mathbb{R}^d$-valued path $y^{(1)} := y_\cdot(1)$ is the solution on $[0,1]$ (in the sense of Definition 2.3) of the rough equation

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d\widetilde{\mathbf{x}}_t \quad , \quad y_0 = a_1 \ , \tag{6.76}$$

where $\widetilde{\mathbf{x}}$ is the canonical rough path above $\widetilde{x} := z + (h + e)$, with

$$e_t := -\int_0^t ds\, \sigma(y_s(1))^{-1} \int_0^1 d\eta\, \jmath_s(\eta) \ .$$

(c) It holds that $y_1^{(0)} = y_1^{(1)}$.

Proof. (a) Let y be a solution of (6.75) in the sense of Definition 2.3. Then clearly it is also a solution in the sense of Definition 5.2, and by Corollary A.2, we can conclude that y is a solution of the equation

$$dy_t = \big[b(y_t)\, dt + \sigma(y_t)\, dh_t\big] + \sigma(y_t)\, d\mathbf{z}_t \quad , \quad y_0 = a_0 \ , \tag{6.77}$$

(that is, in Definition 5.2, we take $V := \mathbb{R}^d$, $m - 1 = n = d$, $B_0(y) := b(y)$, $B_i(y) := \sigma_i(y)$, $\Sigma_i(y) := \sigma_i(y)$, $h_t \leftrightarrow (t, h_t)$, $\mathbf{z} \leftrightarrow \mathbf{z}$). The conclusion then comes from the uniqueness statement contained in Theorem 5.5, since $y_\cdot(0)$ trivially satisfies Equation (6.77) as well.

(b) As above, observe that due to the regularity of the path e and thanks to Corollary A.2, the solution of (6.76) (in the sense of Definition 2.3) is also the solution of

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d(h+e)_t + \sigma(y_t)\, d\mathbf{z}_t \quad , \quad y_0 = a_1 \ , \tag{6.78}$$

in the sense of Definition 5.2 (that is, with $V := \mathbb{R}^d$, $m - 1 = n = d$, $B_0(y) := b(y)$, $B_i(y) := \sigma_i(y)$, $\Sigma_i(y) := \sigma_i(y)$, $h_t \leftrightarrow (t, h_t + e_t)$, $\mathbf{z} \leftrightarrow \mathbf{z}$). It turns out that the path $w := y_\cdot(1)$ satisfies Equation (6.78) as well. This can be easily derived from the fact that

$$\begin{aligned}
&\Big|\sigma(w_s)\,(\delta e)_{st} - \Big(-\int_0^1 d\eta\, \jmath_s(\eta)\Big)\,(t-s)\Big| \\
&= \quad |\sigma(w_s)|\,\Big|\int_s^t dr \int_0^1 d\eta \Big[\sigma(w_r)^{-1}\jmath_r(\eta) - \sigma(w_s)^{-1}\jmath_s(\eta)\Big]\Big| \\
&\lesssim \quad |t-s|^{1+\gamma}\left\{\mathcal{N}[\jmath; \mathcal{C}_1^0([0,1])]\,\mathcal{N}[w; \mathcal{C}_1^\gamma([0,1])] + \mathcal{N}[\jmath; \mathcal{C}_1^\gamma([0,1])]\right\} .
\end{aligned}$$

Therefore, just as for point (a), we can conclude with the help of the uniqueness property stated in Theorem 5.5.

(*c*) The assertion relies on the following identity: for every $t, \xi \in [0,1]$, one has

$$(\partial_\xi y_t)(\xi) = \jmath_t(\xi)(1-t) \ . \tag{6.79}$$

It can indeed be checked that, when seen as paths with values in $V := L^\infty([0,1];\mathbb{R}^d)^3$, the triplets $(y, \jmath, \partial_\xi y)$ and $(y, \jmath, g)$ (where we have set $g_t(\xi) := \jmath_t(\xi)(1-t)$) are both solution of the system obtained by adding to (6.72)-(6.73) the third equation

$$dg_t = [(\partial_k b)(y_t) g_t^k - \jmath_t]\, dt + (\partial_k \sigma)(y_t) g_t^k \, dh_t + (\partial_k \sigma)(y_t) g_t^k \, d\mathbf{z}_t \quad , \quad g_0(\xi) = a_1 - a_0 \ .$$

The conclusion is now immediate:

$$y_1^{(1)} - y_1^{(0)} = y_1(1) - y_1(0) = \int_0^1 g_1(\xi)\, d\xi = 0 \ .$$

□

Before we summarize the previous results into a single statement (Theorem 6.4 below), let us introduce an auxiliary system which will later serve us as an ingredient to 'invert' the hitting system. This system (or rather this family of systems) takes values in $V_1 := L^\infty([0,1])^2$, and is defined for every smooth compactly-supported $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ as follows:

$$\begin{aligned} d\bar{y}_t(\xi) &= \left[\varphi(b(\bar{y}_t(\xi))) - \int_0^\xi \varphi(\bar{\jmath}_t(\eta))\, d\eta + \sigma(\bar{y}_t(\xi))\sigma(\bar{y}_t(1))^{-1} \int_0^1 \varphi(\bar{\jmath}_t(\eta))\, d\eta \right] dt \\ &\qquad + \sigma(\bar{y}_t(\xi))\, dh_t + \sigma(\bar{y}_t(\xi))\, d\mathbf{z}_t \ , \end{aligned} \tag{6.80}$$

$$\begin{aligned} d\bar{\jmath}_t(\xi) &= \left[(\partial_k b)(\bar{y}_t(\xi))\varphi(\bar{\jmath}_t(\xi))_k + (\partial_k\sigma)(\bar{y}_t(\xi))\varphi(\bar{\jmath}_t(\xi))_k \sigma(\bar{y}_t(1))^{-1} \int_0^1 \varphi(\bar{\jmath}_t(\eta))\, d\eta \right] dt \\ &\qquad + (\partial_k\sigma)(\bar{y}_t(\xi))\varphi(\bar{\jmath}_t(\xi))_k \, dh_t + (\partial_k\sigma)(\bar{y}_t(\xi))\varphi(\bar{\jmath}_t(\xi))_k \, d\mathbf{z}_t \ . \end{aligned} \tag{6.81}$$

with initial condition $(y_0(\xi), \jmath_0(\xi)) = ((1-\xi)a_0 + \xi a_1, a_1 - a_0)$. It is not hard to see that for every such fixed φ, the vector fields involved in (6.80)-(6.81) satisfy the conditions of Theorem 5.4 (that is, Hypothesis (VF1)), and therefore the system admits a unique solution

$$(\bar{y}, \bar{\jmath}) =: \bar{\Psi}_{V_1}((a_0, a_1), \varphi, (h, \mathbf{z})) \in \mathcal{C}_1^\gamma([0,1]; V_1) \ . \tag{6.82}$$

Notation. For all $h : [0,1] \to \mathbb{R}^d$ and $g \in L^1([0,1];\mathbb{R}^d)$, we set, for all $t \in [0,1]$,

$$T(h,g)_t := h_t + \int_0^t g_s\, ds \ .$$

Theorem 6.4. *Fix $K \geqslant 1$ and $V_1 := L^\infty([0,1];\mathbb{R}^d)^2$. Then there exists a smooth compactly-supported function $\varphi_K : \mathbb{R}^d \to \mathbb{R}^d$ such that, for all $a_0, a_1 \in \mathbb{R}^d$, the following assertions hold true:*

(*i*) *The system*

$$\begin{aligned} dy_t(\xi) &= \left[\varphi_K(b(y_t(\xi))) - \int_0^\xi \varphi_K(\jmath_t(\eta))\, d\eta\right] dt + \sigma(y_t(\xi))\, dh_t + \sigma(y_t(\xi))\, d\mathbf{z}_t \ , \\ d\jmath_t(\xi) &= (\partial_k b)(y_t(\xi))\varphi_K(\jmath_t(\xi))_k \, dt + (\partial_k\sigma)(y_t(\xi))\varphi_K(\jmath_t(\xi))_k \, dh_t + (\partial_k\sigma)(y_t(\xi))\varphi_K(\jmath_t(\xi))_k \, d\mathbf{z}_t \ , \end{aligned}$$

with initial condition $(y_0(\xi), \jmath_0(\xi)) = ((1-\xi)a_0 + \xi a_1, a_1 - a_0)$, admits a unique solution

$$(y, \jmath) =: \Psi_{V_1}((a_0, a_1), \varphi_K, (h, \mathbf{z})) \in \mathcal{C}_1^\gamma([0,1]; V_1) \ , \tag{6.83}$$

in the sense of Definition 5.2.

(*ii*) *There exists a constant $M_K > 0$ such that if $\max(|a_0|, |a_1|, |a_1 - a_0|) \leqslant K$, $|||h|||_{1;\gamma} \leqslant K$ and $\|\mathbf{z}\|_{\gamma;[0,1]} \leqslant M_K$, then, defining $(y, \jmath)$ through (6.83), one has: (ii-a) the $\mathbb{R}^d$-valued path $y^{(0)} = y_\cdot(0)$ is the solution of*

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d\mathbf{x}_t \ , \ y_0 = a_0 \ ,$$

where $\mathbf{x}$ *is the canonical rough path above* $x := z + h$*;* $(ii\text{-}b)$ *the* $\mathbb{R}^d$*-valued path* $y^{(1)} = y_\cdot(1)$ *is the solution of*

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\widetilde{\mathbf{x}}_t \ , \ y_0 = a_1 \ ,$$

where $\widetilde{\mathbf{x}}$ *is the canonical rough path above* $\widetilde{x} := z + T(h,g)$*, with*

$$g(t) := -\sigma(y_t(1))^{-1} \int_0^1 d\eta\, \varphi_K(j_t(\eta)) \quad , \quad (y,j) := \Psi_{V_1}((a_0,a_1),\varphi_K,(h,\mathbf{z})) \ ; \tag{6.84}$$

$(ii\text{-}c)$ $y_1^{(0)} = y_1^{(1)}$.

(iii) *With notations (6.82) and (6.83) in mind, we have the following identities:*

$$\Psi_{V_1}((a_0,a_1),\varphi_K,(h,\mathbf{z})) = \bar{\Psi}_{V_1}((a_0,a_1),\varphi_K,(T(h,g),\mathbf{z})) \ , \tag{6.85}$$

with g *defined just as in (6.84), and*

$$(\bar{y},\bar{j}) := \bar{\Psi}_{V_1}((a_0,a_1),\varphi_K,(h,\mathbf{z})) = \Psi_{V_1}((a_0,a_1),\varphi_K,(T(h,\bar{g}),\mathbf{z})) \ , \tag{6.86}$$

with

$$\bar{g}(t) := \sigma(\bar{y}_t(1))^{-1} \int_0^1 d\eta\, \varphi_K(\bar{j}_t(\eta)) \ . \tag{6.87}$$

Proof. With the notation $C(K)$ used in Proposition 6.2 point (b), consider any smooth function $\varphi_K : \mathbb{R}^d \to \mathbb{R}^d$ such that $\varphi_K \equiv 1$ on $[-C(K),C(K)]^d$ and $\varphi_K(x) = 0$ for every $|x| \geqslant 2\,C(K)$. Then (i) follows immediately from Proposition 6.2 point (a). Besides, owing to (6.74), it is clear that by defining M_K just as in Proposition 6.2 point (b), and under the assumptions of the above point (ii), one has the identity

$$\Psi_{V_1}((a_0,a_1),\varphi_K,(h,\mathbf{z})) = \Psi_{V_2}(A,\mathrm{Id},(h,\mathbf{z})) \quad \text{with } A(\xi) := ((1-\xi)a_0 + \xi a_1, a_1 - a_0) \ .$$

Therefore, the three points $(ii\text{-}a)$-$(ii\text{-}b)$-$(ii\text{-}c)$ can be readily deduced from Proposition 6.3.

In order to prove (iii), observe first that with the notations in (6.84), one has, at least at a formal level,

$$\begin{aligned} dy_t(\xi) \quad = \quad & \Big[\varphi(b(y_t(\xi))) - \int_0^\xi \varphi(j_t(\eta))\,d\eta + \sigma(y_t(\xi))\sigma(y_t(1))^{-1} \int_0^1 \varphi(j_t(\eta))\,d\eta\Big] \\ & + \sigma(y_t(\xi))\,d\big(T(h,g)\big)_t + \sigma(y_t(\xi))\,d\mathbf{z}_t \ , \end{aligned}$$

with a similar transformation for the equation involving $\jmath$. Given the regularity of g, the latter transformations can actually be justified in a rigourous way, that is in the framework settled through Definition 5.2: one can for instance mimick the arguments of the proof of Proposition 6.3 point (b). Identity (6.85) now follows from the uniqueness of the solution to the system (6.80)-(6.81) (with fixed $\varphi := \varphi_K$). Identity (6.86) can then be derived from a similar transformation of (6.80)-(6.81), which completes the proof of our statement. □

6.2 Toward a Girsanov transformation

Let us go back to our stochastic setting, where $x = X$ stands for a H-fBm. The aim now is to translate the previous results at the level of the underlying Wiener paths, so as to construct the expected coupling $(W,\widetilde{W})$ on $[\tau_k,\tau_k+1]$ via a Girsanov-type argument. To this end, we will deduce from Theorem 6.4 how to build an appropriate drift function $g_{_W}$ for the hitting objective to be achieved. This property is the topic of Theorem 6.7 below, that we write (without loss of generality) with $\tau_k = 0$. Just before we state this result, we need to specify, through the following technical lemma, how the Wiener space can be somehow 'decomposed' in accordance with the past-innovation splitting (4.26).

Let us recall that we have fixed $H \in (\frac{1}{3},\frac{1}{2})$ and $\gamma \in (\frac{1}{3},H)$ for the whole section. Besides, in the sequel, we will indifferently denote by $\mathbb{P}_W$, and call the Wiener measure, the Wiener measure on $\mathcal{C}((-\infty,0];\mathbb{R}^d)$ (reversed Brownian motion), the Wiener measure on $\mathcal{C}([0,1];\mathbb{R}^d)$, as well as the law of a two-sided Brownian motion on $\mathcal{C}((-\infty,1];\mathbb{R}^d)$.

Lemma 6.5. *There exist two spaces $\Omega_- \subset \mathcal{C}(]-\infty,0];\mathbb{R}^d)$ and $\Omega_+ \subset \mathcal{C}([0,1];\mathbb{R}^d)$ of full Wiener measure such that the following properties are satisfied:*

(i) Let $\mathcal{D}_X^-$ be defined for every smooth compactly-supported $\varphi :]-\infty,0] \to \mathbb{R}^d$ vanishing at 0 by

$$\mathcal{D}_X^-\varphi(t) := \alpha_H \int_{-\infty}^0 \Big((t-r)^{H-\frac{1}{2}} - (-r)^{H-\frac{1}{2}} \Big)\, d\varphi(r) \quad \text{if } t \in (0,1]$$

and $\mathcal{D}_X^-\varphi(0) = 0$. Then $\mathcal{D}_X^-$ extends to Ω_- as an application with values in $\mathcal{E}_\gamma^2$. Besides, for every $w_- \in \Omega_-$ and every bounded function $g_W^- :]-\infty,0] \to \mathbb{R}^d$ with compact support, the path $\widetilde{w}_- := w_- + \int_{-\infty}^{\cdot} g_W^-(s)\, ds$ still belongs to Ω_-.

(ii) Let $\mathcal{D}_X^+$ be defined for every smooth compactly-supported $\varphi : [0,1] \to \mathbb{R}^d$ vanishing at 0 by

$$\mathcal{D}_X^+\varphi(t) = \alpha_H \int_0^t (t-r)^{H-\frac{1}{2}} d\varphi(r)\ ,\ t \in [0,1]\ .$$

Then $\mathcal{D}_X^+$ extends to Ω_+ as an application with values in $\mathcal{C}^\gamma([0,1];\mathbb{R}^d)$, and for every $w_+ \in \Omega_+$, $\mathcal{D}_X^+ w_+$ can be canonically lifted into a rough path $\mathfrak{L}(\mathcal{D}_X^+ w_+)$, in the sense of Definition 2.2. Besides, for every $w_+ \in \Omega_+$ and every continuous $g_W^+ : [0,1] \to \mathbb{R}^d$, the path $\widetilde{w}_+ := w_+ + \int_0^{\cdot} g_W^+(s)\, ds$ still belongs to Ω_+.

(iii) For every $\varepsilon > 0$, it holds that

$$\mathbb{P}_W(w_+ \in \Omega_+ :\ \|\mathfrak{L}(\mathcal{D}_X^+ w_+)\|_{\gamma;[0,1]} \leqslant \varepsilon) > 0\ .$$

(iv) Set $\Omega := \{w_- \sqcup w_+ :\ w_- \in \Omega_-, w_+ \in \Omega_+\} \subset \mathcal{C}(]-\infty,1];\mathbb{R}^d)$ and for every $w = w_- \sqcup w_+ \in \Omega$,

$$\mathcal{D}_X w := \mathcal{D}_X^- w_- + \mathcal{D}_X^+ w_+\ .$$

Then $\mathcal{D}_X w$ belongs to $\mathcal{C}^\gamma([0,1];\mathbb{R}^d)$ and can be canonically lifted as a rough path, in the sense of Definition 2.2. Furthermore, as a random variable on $(\Omega, \mathbb{P}_W)$, $\mathcal{D}_X$ has the law of a fBm of Hurst index H.

(v) It holds that $\Theta_{-1}(\Omega) \subset \Omega_-$, where Θ stands for the shift operator, that is $\Theta_\tau(w)_t = w_{t+\tau}$.

Proof. Let us fix $\varepsilon \in (0, H-\gamma)$.

(i) Note first that, using an elementary integration-by-parts formula, $\mathcal{D}_X^-$ can be equivalently defined as

$$\mathcal{D}_X^-\varphi(t) := \alpha_H \left(H - \frac{1}{2}\right) \int_{-\infty}^0 \Big((t-r)^{H-\frac{3}{2}} - (-r)^{H-\frac{3}{2}} \Big)\, \varphi(r)\, dr \quad \text{if } t \in (0,1] \tag{6.88}$$

and $\mathcal{D}_X^-\varphi(0) = 0$, for every test-function φ vanishing at 0. Then fix $\varepsilon' > 0$ and consider the subspace $\Omega_- := \mathcal{C}_{\varepsilon,\varepsilon'}$ of paths $w_- : (-\infty,0] \to \mathbb{R}^d$ which are $(\frac{1}{2}-\varepsilon)$-Hölder continuous on compacts intervals, vanish at 0, and satisfy

$$\lim_{t \to -\infty} \frac{w_-(t)}{|t|^{\frac{1}{2}+\varepsilon'}} = 0\ .$$

Owing to some classical properties on the Wiener process, this subspace is of Wiener measure 1 for any fixed $\varepsilon' > 0$, and from (6.88), it is easy to check that $\mathcal{D}_X^-$ continuously extends to Ω_-, as an application with values in $\mathcal{E}_\gamma^2$. Indeed, for every test-function φ vanishing at 0, every $k \geqslant 1$ and every $t \in (0,1]$, it holds that, for some constant $c_k \geqslant 0$,

$$\begin{aligned}
t^{k-\gamma}\big|(\mathcal{D}_X^-\varphi)^{(k)}(t)\big| &= c_k\, t^{k-\gamma} \left| \int_{-\infty}^0 (t-r)^{H-\frac{3}{2}-k} \varphi(r)\, dr \right| \\
&\leqslant c_k \left\{ \int_{-\infty}^{-1} (t-r)^{H-\frac{3}{2}-k} |\varphi(r)|\, dr + t^{k-\gamma} \int_{-1}^0 (t-r)^{H-\frac{3}{2}-k} |\varphi(r)|\, dr \right\} \\
&\lesssim \left(\sup_{r \leqslant -1} \frac{|\varphi(r)|}{|r|^{\frac{1}{2}+\varepsilon'}} \right) \int_{-\infty}^{-1} \frac{dr}{|r|^{1+k-H-\varepsilon'}} + \|\varphi\|_{\gamma;[-1,0]} t^{k-\gamma} \int_{-1}^0 (t-r)^{H-1-k-\varepsilon}\, dr \\
&\lesssim \sup_{r \leqslant -1} \frac{|\varphi(r)|}{|r|^{\frac{1}{2}+\varepsilon'}} + \|\varphi\|_{\gamma;[-1,0]}\ .
\end{aligned}$$

(*ii*) Just as above, note that $\mathcal{D}_X^+$ can be equivalently defined as

$$\mathcal{D}_X^+\varphi(t) := \alpha_H \varphi(t)\, t^{H-\frac{1}{2}} + \alpha_H \left(H - \frac{1}{2}\right) \int_0^t (t-r)^{H-\frac{3}{2}} \big(\varphi(t) - \varphi(r)\big)\, dr \quad \text{if } t \in (0,1] \tag{6.89}$$

and $\mathcal{D}_X^-\varphi(0) = 0$, for every test-function φ on $[0,1]$ vanishing at 0. From this expression, it is easy to check that, as a map with values in $\mathcal{C}^\gamma([0,1])$, $\mathcal{D}_X^+$ continuously extends to the (full-Wiener-measure) space $\widetilde{\Omega}_+$ of $(\frac{1}{2}-\varepsilon)$-Hölder paths on $[0,1]$ which vanish at 0. We can also check that the covariance function of the Gaussian process $\mathcal{D}_X^+ : (\widetilde{\Omega}_+, \mathbb{P}_W) \to \mathcal{C}^\gamma([0,1])$ so defined satisfies the conditions of [12, Theorem 15.33], which allows us to assert that the subspace

$$\Omega_+ := \big\{ w_+ \in \widetilde{\Omega}_+ : \mathcal{D}_X^+ w_+ \text{ can be canonically lifted as a } \gamma\text{-rough path} \big\}$$

is of full Wiener measure. Finally, the stability of Ω_+ through the transformation $w_+ \mapsto w_+ + \int_0^\cdot g_w^+(s)\, ds$ follows from the definition of $\widetilde{\Omega}_+$ and the result of Proposition A.1.

(*iii*) By [12, Theorem 15.60] (and using the terminology therein introduced), the assertion reduces to showing that the Gaussian process $\mathcal{D}_X^+ : (\Omega_+, \mathbb{P}_W) \to \mathcal{C}^\gamma([0,1])$ satisfies the complementary Young regularity condition. It turns out that this specific result has been proved in [10, Example 2.11], which immediately yields the conclusion.

(*iv*) It holds that $\mathcal{E}_\gamma^1 \subset \mathcal{C}^\gamma([0,1])$, and so $\mathcal{D}_X w$ does belong to $\mathcal{C}^\gamma([0,1])$, for every $w \in \Omega$. The fact that it can be canonically lifted as a γ-rough path follows from point (*ii*) (that is, we can lift $\mathcal{D}_X^+ w_+$) and Proposition A.1 (due to $\mathcal{D}_X^- w_- \in \mathcal{E}_\gamma^2$). Finally, when dealing with a two-sided Brownian motion W on $(-\infty, 1]$, and starting from the explicit formulas (6.88)-(6.89) for $\mathcal{D}_X^-$ and $\mathcal{D}_X^+$, we can apply Itô formula to identify $\mathcal{D}_X W$ with the Mandelbrot-Van Ness transformation of W.

(*v*) It can be immediately checked from the previous constructions. □

We will also rely on the following inversion formula, borrowed from [17, p. 741]:

Lemma 6.6. *Let $H < 1/2$. Let $w_-, \widetilde{w}_- \in \Omega_-$, $w_+, \widetilde{w}_+ \in \Omega_+$ be such that*

$$\widetilde{w}_- = w_- + \int_{-\infty}^\cdot g_w^-(s)\, ds \quad , \quad \widetilde{w}_+ = w_+ + \int_0^\cdot g_w^+(s)\, ds$$

and

$$\mathcal{D}_X(\widetilde{w}_- \sqcup \widetilde{w}_+)_{|[0,1]} = \mathcal{D}_X(w_- \sqcup w_+)_{|[0,1]} + \int_0^\cdot g_X^+(s)\, ds \ ,$$

where $g_w^- : (-\infty, 0] \to \mathbb{R}^d$ is a bounded function with compact support on $(-\infty, 0)$ and $g_w^+, g_X^+ : [0,1] \to \mathbb{R}^d$ is also bounded. Then it holds that $g_w^+ = \mathcal{H}(g_w^-, g_X^+)$ where

$$\begin{aligned}\mathcal{H}(g_1, g_2)_t &:= C_1 \int_{-\infty}^0 \frac{t^{\frac{1}{2}-H}(-s)^{H-\frac{1}{2}}}{t-s} g_1(s)\, ds + \alpha_H \frac{d}{dt}\left(\int_0^t (t-s)^{\frac{1}{2}-H} g_2(s)\, ds \right) \\ &= C_1 \mathcal{R}_0 g_1(t) + C_2 \int_0^t (t-s)^{-\frac{1}{2}-H} g_2(s)\, ds \ .\end{aligned}$$

The above formula can be interpreted as follows: in the spirit of (4.34), the second term of $\mathcal{H}(g_w^-, g_X^+)$ corresponds to the the drift on the Wiener component induced by the "fractional drift" g_X^+ whereas by Lemma 4.1, the first term is the drift on the Wiener component on $[0,1]$ which ensures that, given a past g_w^-, the corresponding fractional drift is equal to 0.

In the next statement, we denote by $\Phi(a; \mathbf{x})$ the unique solution on $[0,1]$ of the rough equation

$$dy_t = b(y_t)\, dt + \sigma(y_t)\, d\mathbf{x}_t \quad , \quad y_0 = a \ ,$$

understood in the sense of Definition 2.3. Let us recall that b and σ are assumed to satisfy Hypothesis (**H1**), so that the above equation indeed admits a unique solution on $[0,1]$.

Theorem 6.7. *Fix two parameters $K, \alpha > 0$, as well as a (K, α, γ)-admissible state $\pi = (a, \widetilde{a}, w_-, \widetilde{w}_-) \in (\mathbb{R}^d)^2 \times \Omega_-^2$. Also, consider $g_W^- :]-\infty, 0] \to \mathbb{R}^d$ such that $\widetilde{w}_- = w_- + \int_{-\infty}^{\cdot} g_W^-(s)\, ds$. Then there exists an application $\Lambda = \Lambda_\pi : \Omega_+ \to \Omega_+$ such that:*

(i) For every $w_+ \in \Omega_+$ and every $t \in [0,1]$, $\Lambda(w_+)_t = w_+(t) + \int_0^t g_W^+(w_+, s) ds$, for some $(\mathcal{G}_t)_{t\in[0,1]}$-adapted function g_W^+,

(ii) There exists a positive constant δ_K depending only on K such that

$$\mathbb{P}_W\big(w_+ \in \Omega_+ : \Phi(a; \mathfrak{L}(\mathcal{D}_X(w_- \sqcup w_+)))_1 = \Phi(\tilde{a}; \mathfrak{L}(\mathcal{D}_X(\widetilde{w}_- \sqcup \Lambda(w_+)))_1\big) \geqslant \delta_K \ .$$

(iii) Λ is bijective with inverse Λ^{-1} satisfying $\Lambda^{-1}(w_+)_t = w_+(t) + \int_0^t \bar{g}_W^+(w_+, s) ds$, for some $(\mathcal{G}_t)_{t\in[0,1]}$-adapted function $\bar{g}_W^+$,

(iv) There exists $C_K > 0$ depending only on K such that for $\mathbb{P}_W$-every $w_+ \in \Omega_+$,

$$\int_0^1 (|g_W^+|^2 + |\bar{g}_W^+|^2)(w_+, s)\, ds \leqslant C_K \ .$$

(v) For every fixed $w_+ \in \Omega_+$, consider the function $g_W(w_+, .) :]-\infty, 1] \to \mathbb{R}^d$ defined as $g_W(w_+, t) = g_W^-(t)$ if $t \leqslant 0$ and $g_W(w_+, t) = g_W^+(w_+, t)$ if $t \in (0, 1]$, and denote by $g_X(w_+, .)$ the image of $g_W(w_+, .)$ through the transformation (4.35). Then there exists $C_K > 0$ depending only on K such that for $\mathbb{P}_W$-every $w_+ \in \Omega_+$,

$$\sup_{t\in[0,1]} |g_X(w_+, t)| \leqslant C_K \ .$$

Proof. Set $h_{w_-} := \mathcal{D}_X^- w_-$ and write, for every $w_+ \in \Omega_+$, $\mathbf{z}_{w_+} := \mathfrak{L}(\mathcal{D}_X^+ w_+)$ (see Lemma 6.5(*ii*)).

(*i*) With the notations of Theorem 6.4, consider the function $g_X : \Omega_+ \times [0,1] \to \mathbb{R}^d$ given by

$$g_X(w_+, t) := -\sigma(y_t(1))^{-1} \int_0^1 d\eta\, \varphi_K(j_t(\eta)), \quad \text{with} \quad (y, j) := \Psi_{V_1}((a, \widetilde{a}), \varphi_K, (h_{w_-}, \mathbf{z}_{w_+})) \ . \tag{6.90}$$

Then, with the notations of Lemma 6.6, we define Λ as

$$\Lambda(w_+)_t := w_+(t) + \int_0^t \mathcal{H}(g_W^-, g_X(w_+, .))_s\, ds = T\big(w_+, \mathcal{H}(g_W^-, g_X(w_+, .))\big)_t \ , \quad t \in [0,1].$$

(*ii*) By the very definition of $\mathcal{H}$, we have, for every $w_+ \in \Omega_+$,

$$\mathcal{D}_X(\widetilde{w}_- \sqcup \Lambda(w_+))_{|0,1]} = \mathcal{D}_X(w_- \sqcup w_+)_{|[0,1]} + \int_0^{\cdot} g_X(w_+, s)\, ds \ .$$

Besides, by the admissibility condition, we know that $|||h_{w_-}|||_{1;\gamma} \leqslant K$ and $\max(|a|, |\widetilde{a}|, |\widetilde{a} - a|) \leqslant 2K$. Therefore, we are exactly in a position to apply Theorem 6.4 and deduce the existence of a positive constant $M_K > 0$ such that

$$\mathbb{P}_W\big(w_+ \in \Omega_+ : \Phi(a; \mathfrak{L}(\mathcal{D}_X(w_- \sqcup w_+)))_1 = \Phi(\tilde{a}; \mathfrak{L}(\mathcal{D}_X(\widetilde{w}_- \sqcup \Lambda(w_+)))_1\big)$$

$$\geqslant \mathbb{P}_W(w_+ \in \Omega_+ : \|\mathbf{z}_{w_+}\|_{\gamma;[0,1]} \leqslant M_K) \ .$$

The conclusion now comes from Lemma 6.5, point (*iii*).

(*iii*) Set $h_{\widetilde{w}_-} := \mathcal{D}_X^- \widetilde{w}_-$ and with the notations of Theorem 6.4(*iii*), define, for every $w_+ \in \Omega_+$,

$$\bar{g}_X(w_+, t) := \sigma(\bar{y}_t(1))^{-1} \int_0^1 d\eta\, \varphi_K(\bar{j}_t(\eta)), \quad \text{with} \quad (\bar{y}, \bar{j}) := \bar{\Psi}_{V_1}((a, \widetilde{a}), \varphi_K, (h_{\widetilde{w}_-}, \mathbf{z}_{w_+})) \ .$$

Then consider the application $\bar{\Lambda} : \Omega_+ \to \Omega_+$ given by

$$\bar{\Lambda}(w_+)_t = w_+(t) + \int_0^t \mathcal{H}(-g_W^-, \bar{g}_X(w_+,.))_s \, ds = T_t(w_+, \mathcal{H}(-g_W^-, \bar{g}_X(w_+,.))) \; .$$

Let us check that $\bar{\Lambda}$ is actually the inverse of Λ, by showing first that $\bar{\Lambda} \circ \Lambda = \mathrm{Id}$. To this end, fix $w_+ \in \Omega_+$ and set $\bar{w}_+ := \Lambda(w_+) = T(w_+, g_W^+(w_+,.))$, where, according to (i), $g_W^+(w_+,.) := \mathcal{H}(g_W^-, g_X(w_+,.))$, with $g_X(w_+,.)$ given by (6.90). Then consider the functions $G_W^- = \int_{-\infty}^{\cdot} g_W^-(s)\, ds$ and $G_W^+(w_+,.) := \int_0^{\cdot} g_W^+(w_+,s)\, ds$ (defined respectively on $(-\infty, 0]$ and $[0,1]$), so that, by construction,

$$z_{\bar{w}_+} = z_{w_+} + \mathcal{D}_X^+(G_W^+(w_+,.)) \; .$$

Thus, using Corollary A.2, we get

$$\begin{aligned} \bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, (h_{\widetilde{w}_-}, \mathbf{z}_{\bar{w}_+})) &= \bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, \big(h_{\widetilde{w}_-} + \mathcal{D}_X^+(G_W^+(w_+,.)), \mathbf{z}_{w_+})) \\ &= \bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, (h_{w_-} + \mathcal{D}_X^- G_W^- + \mathcal{D}_X^+(G_W^+(w_+,.)), \mathbf{z}_{w_+}) \; . \end{aligned} \tag{6.91}$$

At this point, observe that by the inversion formula (4.35), one has for any $t \in (0,1]$,

$$\begin{aligned} &\frac{d}{dt}\left(\mathcal{D}_X^- G_W^- + \mathcal{D}_X^+(G_W^+(w_+,.))\right)(t) \\ &= \alpha_H \frac{d}{dt}\left(\int_{-\infty}^0 \left((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}\right) g_W^-(s)\, ds + \int_0^t (t-s)^{H-\frac{1}{2}} \mathcal{H}(g_W^-, g_X(w_+,.))_s \, ds\right) \\ &= g_X(w_+,t) \; . \end{aligned}$$

Therefore, $h_{w_-} + \mathcal{D}_1 G_W^- + \mathcal{D}_2 G_W^+ = T(h_{w_-}, g_X(w_+,.))$, which, going back to (6.91), gives us

$$\bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, (h_{\widetilde{w}_-}, \mathbf{z}_{\bar{w}_+})) = \bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, T(h_{w_-}, g_X(w_+,.)), \mathbf{z}_{w_+})) \; .$$

We can now apply identity (6.85) to assert that

$$\bar{\Psi}_{V_1}((a,\widetilde{a}), \varphi_K, (h_{\widetilde{w}_-}, \mathbf{z}_{\bar{w}_+})) = \Psi_{V_1}((a,\widetilde{a}), \varphi_K, (h_{w_-}, \mathbf{z}_{w_+})) \; ,$$

which readily entails that $\bar{g}_X(\bar{w}_+,.) = -g_X(w_+,.)$. The conclusion is immediate:

$$\begin{aligned} \bar{\Lambda}(\bar{w}_+) &= T(\bar{w}_+, \mathcal{H}(-g_W^-, \bar{g}_X(\bar{w}_+,.))) \\ &= T(w_+, \mathcal{H}(g_W^-, g_X(w_+,.)) + \mathcal{H}(-g_W^-, \bar{g}_X(\bar{w}_+,.))) = T(w_+, 0) = w_+ \; . \end{aligned}$$

The fact that $\Lambda \circ \bar{\Lambda} = \mathrm{Id}$ follows from symmetric arguments (by using (6.86) instead of (6.85)).

(iv) Let us recall that

$$g_W^+(w_+,t) = \mathcal{H}(g_W^-, g_X(w_+,.))_t = C_1(\mathcal{R}_0 g_W^-)(t) + C_2 \int_0^t (t-s)^{-\frac{1}{2}-H} g_X(w_+,s)\, ds \; .$$

The desired uniform bound on the L^2-norm of $g_W^+(w_+,.)$ then follows from the admissibility condition, Hypothesis (**H3**) and the definition of φ_K. The same arguments can be used for the bound on the L^2-norm of $\bar{g}_W^+(w_+,.)$.

(v) Just as above, it is an immediate consequence of Hypothesis (**H3**) and the definition of φ_K.

□

6.3 Achievement of Step 1

As a conclusion of this section, we obtain the following result.

Proposition 6.8. *Assume* (**H1**) *and* (**H3**). *Let α be a positive constant. Then, for every $K > 0$, there exists a positive constant C_K such that for each $k \geqslant 1$, $(W, \widetilde{W})$ can be built on $[\tau_{k-1}, \tau_{k-1} + 1]$ in such a way that the following properties hold:*

(a) There exists $\widetilde{\delta}_0 > 0$ depending only on K, α and $\gamma \in (1/3, H)$ such that for all $k \geqslant 0$,

$$\mathbb{P}(Y_{\tau_{k-1}+1} = \widetilde{Y}_{\tau_{k-1}+1} | A_k(K, \alpha, \gamma)) \geqslant \widetilde{\delta}_0.$$

(b) If $\omega \in A_k(K, \alpha, \gamma)$, $\sup_{t\in[0,1]} |g_X(t + \tau_{k-1})| \leqslant C_K$.

(c) $\int_{\tau_{k-1}}^{\tau_{k-1}+1} |g_W(s)|^2 ds \leqslant C_K \quad a.s.$

Proof. At the price of a change of variable, we can assume that $\tau_{k-1} = 0$. Then, with the notations of Theorem 6.7, let π denote the current state and let Λ denote the related coupling function. Let $\mathbb{P}_W$ denote the Wiener measure on $[0, 1]$ and $\Lambda^*\mathbb{P}_W$ be the image measure of $\mathbb{P}_W$ by the mapping Λ. By Girsanov Theorem, $\Lambda^*\mathbb{P}_W(dw) = D_\Lambda(w)\mathbb{P}_W(dw)$ where, with the notations of Theorem 6.7,

$$D_\Lambda(w) = \exp\left(\int_0^1 g_W^+(w, s)dw(s) - \frac{1}{2}\int_0^1 |g_W^+(w, s)|^2 ds\right). \tag{6.92}$$

The construction of the coupling then follows the lines of [17] and [9]. For the sake of completeness, one however recalls the principle below.

First, if $\omega \notin A_k(K, \alpha, \gamma)$, one does not attempt Step 1. In other words, in this case, $g_W = 0$ on $[\tau_{k-1}, \tau_{k-1}+1]$.

Second, assume that $\omega \in A_k(K, \alpha, \gamma)$. For positive measures μ_1 and μ_2 with densities D_1 and D_2 with respect to another measure μ, denote by $\mu_1 \wedge \mu_2$ the measure defined by $(\mu_1 \wedge \mu_2)(dw) = D_1(w) \wedge D_2(w)\mu(dw)$. With the help of the invertibility of Λ proved in Theorem 6.7, we define a non-negative measure $\mathbf{P}_1$ on $\mathcal{C}([0, 1], \mathbb{R}^d)^2$ by

$$\mathbf{P}_1 = \frac{1}{2}\Lambda_1^*\mathbb{P}_W \wedge \Lambda_2^*\mathbb{P}_W$$

where Λ_1 and Λ_2 are the functions *a.s.* defined on $\mathcal{C}([0, 1], \mathbb{R}^d)$ by

$$\Lambda_1(w) = (w, \Lambda(w)) \quad \text{and} \quad \Lambda_2(w) = (\Lambda^{-1}(w), w).$$

Let us recall here that, even though this is not suggested by the notation, $\mathbf{P}_1$ strongly depends on the current state π (via Λ). By construction,

$$\Lambda_1^*\mathbb{P}_W(dw_1, dw_2) = \mathbf{1}_{\{(\Lambda^{-1}(w),w)\}}(w_1, w_2)D_\Lambda(w)\mathbb{P}_W(dw),$$

where D_Λ is defined by (6.92). This implies that $\mathbf{P}_1$ satisfies

$$\mathbf{P}_1(dw_1, dw_2) = \frac{1}{2}\mathbf{1}_{\{(\Lambda^{-1}(w),w)\}}(w_1, w_2)(D_\Lambda(w) \wedge 1)\mathbb{P}_W(dw). \tag{6.93}$$

Write $S(w_1, w_2) = (w_2, w_1)$ and denote by $\widetilde{\mathbf{P}}_1$ the "symmetrized" non-negative measure induced by $\mathbf{P}_1$, $\widetilde{\mathbf{P}}_1 := \mathbf{P}_1 + S^*\mathbf{P}_1$. We then define the coupling $(W_t^{\tau_{k-1}}, \widetilde{W}_t^{\tau_{k-1}}) = (W_{t+\tau_{k-1}} - W_{\tau_{k-1}}, \widetilde{W}_{t+\tau_{k-1}} - \widetilde{W}_{\tau_{k-1}})$ as follows:

$$\mathcal{L}((W_t^{\tau_{k-1}}, \widetilde{W}_t^{\tau_{k-1}})_{t\in[0,1]}) = \widetilde{\mathbf{P}}_1 + \Delta^*(\mathbb{P}_W - \Pi_1^*\widetilde{\mathbf{P}}_1) = \mathbf{P}_1 + \mathbf{P}_2$$

with $\Delta(w) = (w, w)$, $\Pi_1(w_1, w_2) = w_1$ and $\mathbf{P}_2 = S^*\mathbf{P}_1 + \Delta^*(\mathbb{P}_W - \Pi_1^*\widetilde{\mathbf{P}}_1)$. Using (6.93), we check that for nonnegative functions f,

$$\Pi_1^*\widetilde{\mathbf{P}}_1(f) \leqslant \frac{1}{2}\int \left(f(\Lambda^{-1}(w))D_\Lambda(w) + f(w)\right)\mathbb{P}_W(dw) \leqslant \mathbb{P}_W(f),$$

hence $\mathbf{P}_2$ is the sum of two non-negative measures. Thanks to the symmetry property of $\widetilde{\mathbf{P}}_1$ and to the fact that $\Pi_1 \circ \Delta$ is the identity, one can also check that the marginals of $\mathbf{P}_1 + \mathbf{P}_2$ are both equal to $\mathbb{P}_W$.

Statement (b) is a direct consequence of the last statement of Theorem 6.7. For (c), the result is obvious if $\omega \notin A_k(K, \alpha, \gamma)$. Otherwise, the construction of $\mathbf{P}_1 + \mathbf{P}_2$ implies that one has only three possibilities: $w_2 = \Lambda(w_1)$ or $w_2 = \Lambda^{-1}(w_1)$ or $w_1 = w_2$ (so that $g_W = 0$ in this case). Thus, the property follows from the fourth statement of Theorem 6.7.

Lower-bound for the probability of coupling: We deduce from Theorem 6.7 that

$$\mathbb{P}(Y_{\tau_{k-1}+1} = \widetilde{Y}_{\tau_{k-1}+1} | A_k(K,\alpha,\gamma)) \geqslant \|\mathbf{1}_{\mathcal{R}\times\Lambda(\mathcal{R})}\mathbf{P}_1\|_{TV}$$

with $\mathcal{R} := \{(w, (y^{\pi,w}(1) = \tilde{y}^{\pi,\Lambda(w)}(1)\}$, where $(y^{\pi,w}, y^{\pi,\Lambda(w)})$ stands for a coupled solution to the SDE on $[0,1]$ with initial condition π and couple of Wiener innovations $(w, \Lambda(w))$. By Theorem 6.7 , there exists a positive δ depending only on K such that $\mathbb{P}_W(\mathcal{R}) \geqslant \delta$.

Now, by (6.93) and Lemma C.1. of [23] (applied to $p = 2$, $\mu_1 = \Lambda^*\mathbb{P}_W$, $\mu_2 = \mathbb{P}_W$ and $X = \mathcal{R}$) we have

$$\|\mathbf{1}_{\mathcal{R}\times\Lambda(\mathcal{R})}\mathbf{P}_1\|_{TV} \geqslant \frac{\left[\int_{\Lambda(\mathcal{R})} D_\Lambda(w)\mathbb{P}_W(dw)\right]^2}{4\int_{\Lambda(\mathcal{R})} D_\Lambda(w)^3\mathbb{P}_W(dw)}.$$

Following the lines of the proof of [9] (Lemma 3.1) and using that $w \mapsto g(w,s)$ is bounded by a constant depending only on K, one deduces that

$$\|\mathbf{1}_{\mathcal{R}\times\Lambda(\mathcal{R})}\mathbf{P}_1\|_{TV} \geqslant C\left[\mathbb{P}_W(\mathcal{R})\right]^2 \geqslant C\delta > 0.$$

□

7 About Step 2

As explained before, Step 2 is a series of trials on some intervals I_ℓ of length $c_2 2^\ell$ (the first one of length $2c_2$, the second one of length $4c_2$,...). We denote by $s_{k,\ell}$ the left extreme of each interval I_ℓ. More precisely, for every $k \geqslant 1$, we define $(s_{k,\ell})_{\ell\geqslant 0}$ by

$$s_{k,0} = s_{k,1} = \tau_{k-1} + 1 \quad \text{and for every } \ell \geqslant 1 \quad s_{k,\ell+1} = s_{k,\ell} + c_2 2^\ell. \tag{7.94}$$

Also denote by ℓ_k^* the (first) trial after time τ_{k-1} where Step 2 fails. The case $\ell_k^* = 0$ and $\ell_k^* = +\infty$ correspond respectively to the failure of Step 1 and to the success of Step 2. For some given positive α and K, we set

$$\mathcal{B}_{k,\ell} := A_k(K,\alpha,\gamma) \cap \{\ell_k^* > \ell\}, \quad k \geqslant 1,\ \ell \geqslant 0. \tag{7.95}$$

With this definition,

$$\mathbb{P}\big(\tau_k = +\infty \,|\, A_k(K,\alpha,\gamma)\big) = \mathbb{P}\big(Y_{\tau_{k-1}+1} = \widetilde{Y}_{\tau_{k-1}+1} \,|\, A_k(K,\alpha,\gamma)\big) \prod_{\ell=1}^{+\infty} \mathbb{P}(\mathcal{B}_{k,\ell}|\mathcal{B}_{k,\ell-1}). \tag{7.96}$$

Consequently, the aim is now to lower-bound $\mathbb{P}(\mathcal{B}_{k,\ell}|\mathcal{B}_{k,\ell-1})$. The scheme of coupling being identical to [9], we only state the result. To this end, let us introduce some notations. Owing to the one-to-one correspondence between g_W and g_X, there is a unique choice for function g_W in $[\tau_{k-1}+1, \infty)$ which ensures that $g_X(t) = 0$ after $\tau_{k-1}+1$ (or equivalently that $Y_t = \widetilde{Y}_t$ after $\tau_{k-1}+1$). We denote it by g_S:

Lemma 7.2. Let $K > 0$ and assume that $\alpha \in (0, H)$. There exists a constant $C_K \geqslant 1$ which does not depend on k such that,

$$\int_0^{+\infty} (1+t)^{2\alpha} |g_S(\tau_{k-1}+1+t)|^2 dt \leqslant C_K.$$

Then, $(W, \widetilde{W})$ can be constructed during Step 2 in such a way that for all $k \geqslant 0$ and $\ell \geqslant 2$,

$$\rho_K^1 \leqslant \mathbb{P}(\mathcal{B}_{k,1}|\mathcal{B}_{k,0}) \leqslant \rho_K^2 \quad \text{and} \quad (1-\rho_K^3 2^{-\alpha\ell}) \leqslant \mathbb{P}(\mathcal{B}_{k,\ell}|\mathcal{B}_{k,\ell-1}) \leqslant (1-\rho_K^3 2^{-\alpha\ell-1}) \tag{7.97}$$

where $\rho_K^1, \rho_K^2 \in (0,1)$ do not depend on k and $\rho_K^3 = c_2^{-\alpha}\sqrt{C_K}$. In particular, if $c_2 = C_K^{\frac{1}{2\alpha}}$, $\rho_K^3 = 1$ and in this case, if $2 \leqslant \ell_k^* < +\infty$ one has

$$\int_{s_{k,\ell_k^*}}^{s_{k,\ell_k^*+1}} |g_W(t)|^2 dt \leqslant (2(\ell_k^*+3))^2 \quad \text{and for all } \ell \in \{2,\dots,\ell_k^*\}, \quad \int_{s_{k,\ell-1}}^{s_{k,\ell}} |g_W(t)|^2 dt \leqslant 2^{-2\alpha\ell},$$

while if $\ell_k^* = 1$, $\int_{s_{k,1}}^{s_{k,2}} |g_W(t)|^2 dt \leqslant C'_K$, for some finite constant C'_K.

Remark 7.5. Going back to (7.96), the lower-bounds in (7.97), combined with the result of Proposition 6.8, ensure that

$$\mathbb{P}\big(\tau_k = +\infty \,|\, A_k(K,\alpha,\gamma)\big) > 0\ .$$

The other properties will be needed for the sequel. Note that c_2 can be chosen in such a way that the involved quantities do not depend on K except if $\ell = 1$.

8 (K,α,γ)-admissibility condition

In this section, we assume that Steps 1 and 2 are carried out as described previously, and the aim is to ensure that the system is (K,α,γ)-admissible with positive probability at every time τ_k. This is the purpose of the next proposition:

Proposition 8.1. *Assume* (**H1**), (**H2**) *and* (**H3**). *Let* $(Y_t,\widetilde{Y}_t)_{t\geqslant 0}$ *denote a solution to* (4.30) *with initial condition* $\widetilde{\mu}$ *satisfying* (4.31) *and* (4.32) *and such that there exists* $r>0$ *for which* $\int |x|^r \mu_0(dx,dw) < +\infty$. *Let* $\alpha \in (0,1/2)$. *Assume that for each* $K>0$, c_2 *defined in* (4.36) *satisfies* $c_2 = C_K^{\frac{1}{2\alpha}}$ *(where* C_K *is a constant greater than* 1 *defined in Lemma 7.2) and that for every* $k\geqslant 1$ *and* $\ell \geqslant 0$, $\Delta_3(\ell,k)$ *introduced in* (4.42) *is defined by* $\Delta_3(\ell,k) = c_3 a_k 2^{\beta\ell}$ *with* $\beta > (1-2\alpha)^{-1}$, $a_k = \varsigma^k$ *for some (arbitrary) fixed* $\varsigma>1$, *and* c_3 *an appropriate constant depending on the previous parameters. Then, for every* $\varepsilon>0$ *and* $\gamma\in(1/3,H)$, *there exists* $K_\varepsilon>0$ *such that for every* $k\geqslant 0$,

$$\mathbb{P}\big(A_{k+1}(K_\varepsilon,\alpha,\gamma)\,\big|\,\tau_k<+\infty\big)\geqslant 1-\varepsilon.$$

where $A_{k+1}(K,\alpha,\gamma)$ *is defined by* (4.40). *More precisely, for every* $k\geqslant 0$,

$$\mathbb{P}\left(\sup_{T\geqslant 0}\int_0^{+\infty}(1+t)^{2\alpha}((\mathcal{R}_T|g_W^{\tau_k}|)(t))^2dt\leqslant 1\,\middle|\,\tau_k<+\infty\right)=1, \tag{8.98}$$

and

$$\mathbb{P}\left(|Y_{\tau_k}(\omega)|+|\widetilde{Y}_{\tau_k}(\omega)|+|||D^{(\tau_k)}(W)|||_{1;\gamma}+|||D^{(\tau_k)}(\widetilde{W})|||_{1;\gamma}\leqslant K_\varepsilon\,\middle|\,\tau_k<+\infty\right)\geqslant 1-\varepsilon. \tag{8.99}$$

The last two statements refer to the two admissibility conditions introduced before.

The proof of (8.98) is identical to the one of Proposition 4.6 of [9]. Thus, we do not reprove it in this paper. However, it worth noting that the condition $\beta>(1-2\alpha)^{-1}$, which is of first importance in the upper-bound on the exponent of the rate, appears in this proof.

We thus focus on the second admissibility condition and thus, on the property (8.99). Using that for every events A_1, A_2, A_3 and A_4, $\mathbb{P}(\cap_{i=1}^4 A_i)\geqslant \sum_{i=1}^4\mathbb{P}(A_i)-3$ and the Markov inequality, one can easily check that it is enough to prove that there exists $C>0$ independent of K such that,

$$\sup_{k\geqslant 0}\mathbb{E}[|||D^{\tau_k}(W)|||_{1;\gamma}|\tau_k<+\infty]\leqslant C,\quad \sup_{k\geqslant 0}\mathbb{E}[|||D^{\tau_k}(\widetilde{W})|||_{1;\gamma}|\tau_k<+\infty]\leqslant C \tag{8.100}$$

and that there exists $p>0$ such that

$$\sup_{k\geqslant 0}\mathbb{E}[|Y_{\tau_k}|^p|\tau_k<+\infty]\leqslant C\quad\text{and}\quad \sup_{k\geqslant 0}\mathbb{E}[|\widetilde{Y}_{\tau_k}|^p|\tau_k<+\infty]\leqslant C. \tag{8.101}$$

We focus successively on these two properties.

8.1 Admissibility and Lyapunov bounds

In this subsection, the aim is to prove (8.101). This property will be obtained as a consequence of the Lyapunov control established in Theorem 3.4 combined with the control of the law of the noise conditionally to the failure of the previous attempts. The result is given below:

Proposition 8.2. *Let $\gamma \in (1/3, H)$ and suppose that the assumptions of Proposition 8.1 hold true. Then, there exist some positive constants C and $p > 0$ such that for every $k \in \mathbb{N} \cup \{0\}$ and $K > 0$,*

$$\mathbb{E}\big[|Y_{\tau_k}|^p \big| \mathcal{E}_k\big] \leqslant C \quad \text{and} \quad \mathbb{E}\big[|\widetilde{Y}_{\tau_k}|^p \big| \mathcal{E}_k\big] \leqslant C. \tag{8.102}$$

where $\mathcal{E}_k = \{\tau_k < +\infty\} = \{\tau_\ell < +\infty, \ell \leqslant k\}$.

Proof. Since the arguments are identical for Y and $\widetilde{Y}$, we only prove the statement for Y. By Theorem 3.4, for every $t \in \mathbb{R}_+$,

$$|Y_{t+1}|^2 \leqslant e^{-C_2/2}|Y_t|^2 + C\big\{1 + \|\mathbf{X}^{(t)}\|_\gamma^\mu\big\} \quad , \quad \text{with } \mu := \frac{8}{3\gamma - 1} \tag{8.103}$$

where $X^{(t)}$ is defined by $X_s^{(t)} = X_{t+s} - X_t$, $s \geqslant 0$. By the elementary inequalities $|u+v|^{\bar{p}} \leqslant |u|^{\bar{p}} + |v|^{\bar{p}}$ for $\bar{p} \in (0,1]$, applied with $\bar{p} = \frac{p}{2}$, we deduce that there exist $\rho \in (0,1)$ and $C > 0$ such that for every $t \in \mathbb{R}_+$

$$|Y_{t+1}|^p \leqslant \rho\,|Y_t|^p + C\big\{1 + \|\mathbf{X}^{(t)}\|_\gamma^{\frac{\mu p}{2}}\big\}, \tag{8.104}$$

An induction yields

$$|Y_{\tau_k}|^p \leqslant \rho^{\Delta\tau_k}|Y_{\tau_{k-1}}|^p + C\sum_{\ell=0}^{\Delta\tau_k - 1} \rho^{\Delta\tau_k - \ell}(1 + \|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{\frac{\mu p}{2}}).$$

First, since $\Delta\tau_k \geqslant \frac{\log(\delta_1/2)}{\log\rho}$, we deduce that $\rho^{\Delta\tau_k} \leqslant \frac{\delta_1}{2}$. Thus,

$$\mathbb{E}[|Y_{\tau_k}|^p|\mathcal{E}_k] \leqslant \frac{\delta_1}{2}\mathbb{E}[\Psi(Y^j_{\tau_{k-1}})|\mathcal{E}_k] + C\sum_{u=0}^{+\infty}\rho^u + C\mathbb{E}\left[\sum_{\ell=0}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{\frac{\mu p}{2}}|\mathcal{E}_k\right].$$

Since $\mathcal{E}_k \subset \mathcal{E}_{k-1}$ and $\mathbb{P}(\mathcal{E}_k|\mathcal{E}_{k-1}) \geqslant \delta_1$, $\mathbb{E}[|X_{\tau_{k-1}}|^p|\mathcal{E}_k] \leqslant \delta_1^{-1}\mathbb{E}[|X_{\tau_{k-1}}|^p|\mathcal{E}_{k-1}]$. It follows that

$$\mathbb{E}[|Y_{\tau_k}|^p|\mathcal{E}_k] \leqslant \frac{1}{2}\mathbb{E}[|Y_{\tau_{k-1}}|^p|\mathcal{E}_{k-1}] + \frac{C}{1-\rho} + C\mathbb{E}[\sum_{\ell=1}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{\frac{\mu p}{2}}|\mathcal{E}_k]$$

and by induction,

$$\sup_{k\geqslant 0}\mathbb{E}\big[|Y_{\tau_k}|^p\big|\mathcal{E}_k\big] \leqslant \mathbb{E}\big[|Y_{\tau_0}|^p\big|\mathcal{E}_0\big] + C_\rho = \mathbb{E}[|Y_0|^p] + C_\rho,$$

where

$$C_\rho := \sup_{k\geqslant 1}\mathbb{E}\left[\sum_{\ell=0}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{\frac{\mu p}{2}}|\mathcal{E}_k\right]. \tag{8.105}$$

On the one hand, if $p < r$, $\mathbb{E}[|Y_0|^p] = \int|x|^p d\mu_0 < +\infty$ (For $\widetilde{Y}_0$ one knows that $\mathbb{E}[|\widetilde{Y}_0|^p] < +\infty$ for every $p > 0$ since the invariant measure has moments of any order). It thus remains to show that there exists a choice for p such that C_ρ can be bounded independently of K. This result is achieved in Proposition 8.4 below. □

8.2 Bounds for the memory component of the fBm

The aim of this section is to prove (8.100), *i.e.* the part of the admissibility condition involving the components D and $\widetilde{D}$ of the fBms X and $\widetilde{X}$. Let us remark that $(D^{(\tau)})'$ is well-defined and satisfies:

$$(D^{(\tau)})'_t = \alpha_H\left(H - \frac{1}{2}\right)\int_{-\infty}^{\tau}(t+\tau-r)^{H-\frac{3}{2}}dW_r.$$

Set

$$\mathcal{D}^\tau_{u,v}(t) = \int_u^v (t+\tau-r)^{H-\frac{3}{2}}dW_r.$$

For $k \geqslant 1$, we thus decompose $D^{(\tau_k)}$ in a series of terms depending on the sequence $(\tau_m)_{m=0}^k$:

$$(D^{(\tau_k)})'_t = \alpha_H \left(H - \frac{1}{2}\right) \left(\mathcal{D}^{\tau_k}_{-\infty,0}(t) + \sum_{m=1}^{k-1} \mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)\right). \tag{8.106}$$

The idea of the sequel is to control each term of the right-hand side separately. We begin by a simple lemma:

Lemma 8.3. For every $u < v < \tau$, for every $t > 0$,

$$|\mathcal{D}^{\tau}_{u,v}(t)| \leqslant C \left((t+\tau-u)^{H-\frac{3}{2}} |W_v - W_u| + \int_u^v (t+\tau-r)^{H-\frac{5}{2}} |W_v - W_r| \, dr\right).$$

Proof. This result is a straightforward consequence of an integration by parts. □

We now state some controls related to the decomposition (8.106).

Lemma 8.4. Assume that there exists $\delta_1 > 0$ and $\alpha > 0$ such that for all $k, \ell \in \mathbb{N}^*$ and $K > 0$

$$\mathbb{P}(\mathcal{E}_k | \mathcal{E}_{k-1}) \geqslant \delta_1 > 0 \quad \mathbb{P}(F_{k,\ell} | \mathcal{E}_{k-1}) \leqslant 2^{-\alpha \ell} \quad \text{and} \quad \Delta\tau_k \geqslant a_k$$

where $(a_k)_{k \geqslant 1}$ is a deterministic sequence such that $a_k \geqslant 1$ for every $k \geqslant 1$. Then, for every $q > 1$ $\gamma \in (0, H)$ and $\varepsilon \in (0, 1-H)$, there exists $C \in \mathbb{R}_+^*$ such that for every $k \geqslant 1$ and every $m \in \{1, \dots, k\}$,

$$\mathbb{E}\big[\big(\|\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}\|_{1;\gamma}\big)^q \big| \mathcal{E}_k\big] \leqslant C \frac{(\sum_{v=m+1}^k a_v \vee 1)^{H-1+\varepsilon}}{\delta_1^{(k-m)(1-\frac{1}{q})}}.$$

Remark 8.6. The assumption on $\mathbb{P}(F_{k,\ell}|\mathcal{E}_{k-1})$ is ensured by Lemma 7.2 (when c_2 is chosen correctly). Actually,

$$\mathbb{P}(F_{k,\ell}|\mathcal{E}_{k-1}) \leqslant \mathbb{P}(\mathcal{B}^c_{k,\ell}|\mathcal{B}_{k,\ell-1}) \leqslant 2^{-\alpha\ell}.$$

Proof. **Step 1.** The aim of this step is to "make deterministic" the duration of Attempt m and to go back to a conditioning by $\mathcal{E}_{m-1}$. First, remark that on the event $F_{m,\ell}$, $\Delta\tau_m$ is deterministic. Denote it by $\Delta(m,\ell)$. One can thus check that for a given non-negative random variable depending on τ_m denoted by R_{τ_m}

$$\mathbb{E}[R_{\tau_m}|\mathcal{E}_m] = \sum_{\ell \geqslant 0} \mathbb{E}[R_{\tau_{m-1}+\Delta(m,\ell)}|F_{m,\ell}]\mathbb{P}(F_{m,\ell}|\mathcal{E}_m).$$

By the Cauchy-Schwarz inequality and the fact that $F_{m,\ell} \subset \mathcal{E}_{m-1}$, one can check that

$$\begin{aligned}\mathbb{E}[R_{\tau_{m-1}+\Delta(m,\ell)}|F_{m,\ell}] &= \frac{\mathbb{E}[R_{\tau_{m-1}+\Delta(m,\ell)}1_{F_{m,\ell}}|\mathcal{E}_{m-1}]}{\mathbb{P}(F_{m,\ell}|\mathcal{E}_{m-1})} \\ &\leqslant \mathbb{E}[R^2_{\tau_{m-1}+\Delta(m,\ell)}|\mathcal{E}_{m-1}]^{\frac{1}{2}} \mathbb{P}(F_{m,\ell}|\mathcal{E}_{m-1})^{-\frac{1}{2}}.\end{aligned}$$

Since $F_{m,\ell} \subset \mathcal{E}_m \subset \mathcal{E}_{m-1}$, one checks that

$$\mathbb{P}(F_{m,\ell}|\mathcal{E}_m) = \frac{\mathbb{P}(F_{m,\ell}|\mathcal{E}_{m-1})}{\mathbb{P}(\mathcal{E}_m|\mathcal{E}_{m-1})}$$

so that

$$\mathbb{E}[R_{\tau_m}|\mathcal{E}_m] \leqslant \delta_1^{-\frac{1}{2}} \sum_{\ell \geqslant 0} \mathbb{E}[R^2_{\tau_{m-1}+\Delta(m,\ell)}|\mathcal{E}_{m-1}]^{\frac{1}{2}} \mathbb{P}(F_{m,\ell}|\mathcal{E}_{m-1})^{\frac{1}{2}}.$$

Under the assumptions of the lemma, we deduce that

$$\mathbb{E}[R_{\tau_m}|\mathcal{E}_m] \leqslant \delta_1^{-\frac{1}{2}} \sum_{\ell \geqslant 0} 2^{-\frac{\alpha\ell}{2}} \mathbb{E}[R^2_{\tau_{m-1}+\Delta(m,\ell)}|\mathcal{E}_{m-1}]^{\frac{1}{2}}$$

so that

$$\mathbb{E}[R_{\tau_m}|\mathcal{E}_m] \leqslant C \sup_{\ell \geqslant 0} \mathbb{E}[R^2_{\tau_{m-1}+\Delta(m,\ell)}|\mathcal{E}_{m-1}]^{\frac{1}{2}}. \tag{8.107}$$

Finally, since $\mathcal{E}_k \subset \mathcal{E}_m$, a similar argument as before implies that for a given random variable S, for every p and $q > 1$ such that $p^{-1} + q^{-1} = 1$,

$$|\mathbb{E}[S|\mathcal{E}_k]| \leqslant \frac{\mathbb{E}[|S|^p|\mathcal{E}_m]^{\frac{1}{p}}}{\mathbb{P}(\mathcal{E}_k|\mathcal{E}_m)^{1-\frac{1}{q}}}. \tag{8.108}$$

Using that $\mathbb{P}(\mathcal{E}_k|\mathcal{E}_m) \geqslant \delta_1^{k-m}$, we can conclude this step with the following control:

$$\mathbb{E}[R_{\tau_m}|\mathcal{E}_k] \leqslant C\delta_1^{(k-m)(1-\frac{1}{q})} \sup_{\ell \geqslant 0} \mathbb{E}[|R_{\tau_{m-1}+\Delta(m,\ell)}|^{2p}|\mathcal{E}_{m-1}]^{\frac{1}{2p}}. \tag{8.109}$$

where C depends on α only.

Step 2. Case $m < k$. Let $\varepsilon \in (0, 1-H)$ and $\gamma \in (0, H)$. By Lemma 8.3 and the fact that $\Delta\tau_v \geqslant a_v$ for every $v \in \mathbb{N}$, we have for every $t \in [0,1]$,

$$\sup_{t\in[0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)| \leqslant C\Big(\sum_{v=m+1}^{k} a_v\Big)^{H-1+\varepsilon} R_{\tau_m} \tag{8.110}$$

with

$$R_{\tau_m} := (\tau_m - \tau_{m-1})^{-\frac{1}{2}-\varepsilon}\left|W_{\tau_m} - W_{\tau_{m-1}}\right| + \int_{\tau_{m-1}}^{\tau_m} (1+\tau_m - r)^{-\frac{3}{2}-\varepsilon}\,|W_{\tau_m} - W_r|\,dr.$$

Since $\int_{\tau_{m-1}}^{\tau_m}(1+\tau_m - r)^{-1-\varepsilon} \leqslant c_\varepsilon = \frac{1}{1+\varepsilon}$, one can first check by the Jensen inequality that

$$\left(\int_{\tau_{m-1}}^{\tau_m} (1+\tau_m - r)^{-\frac{3}{2}-\varepsilon}\,|W_{\tau_m} - W_r|\,dr\right)^p \leqslant c_{\varepsilon,p}\int_{\tau_{m-1}}^{\tau_m} (1+\tau_m - r)^{-1-\varepsilon}\left((1+\tau_m - r)^{\frac{1}{2}}\,|W_{\tau_m} - W_r|\right)^p dr.$$

Then, it follows from the scaling property of the Brownian motion that

$$\mathbb{E}[|R_{\tau_{m-1}+\Delta(m,\ell)}|^{2p}|\mathcal{E}_{m-1}] \leqslant C$$

where C depends on ε and p only. The result then follows (in this case) from (8.110) and (8.109).

Step 3. Case $m = k$. In this case, we cut the interval $[\tau_{k-1}, \tau_k]$ in two parts $[\tau_{k-1}, \tau_k - 1]$ and $[\tau_k - 1, \tau_k]$. On the one hand, the fact that

$$\mathbb{E}\left[\left(|||\mathcal{D}^{\tau_k}_{\tau_{k-1},\tau_k-1}|||_{1;\gamma}\right)^p|\mathcal{E}_k\right] \leqslant C$$

independent of K, follows from similar arguments as in the previous step. On the other hand, for $[\tau_k - 1, \tau_k]$, applying (8.107) leads to the inequality:

$$\mathbb{E}\left[\left(|||\mathcal{D}^{\tau_k}_{\tau_k-1,\tau_k}|||_{1;\gamma}\right)^p|\mathcal{E}_k\right] \leqslant C\sup_{\ell\geqslant 0}\mathbb{E}[R^2_{\tau_{k-1}+\Delta(m,\ell)}|\mathcal{E}_{k-1}]$$

with $R_{\tau_k} := \left(|||\mathcal{D}^{\tau_k}_{\tau_k-1,\tau_k}|||_{1;\gamma}\right)^p$. But, owing to the independence of the increments of the Brownian motion, one can check that for every $\ell \geqslant 0$,

$$\mathbb{E}[R^2_{\tau_{k-1}+\Delta(m,\ell)}|\mathcal{E}_{k-1}] = \mathbb{E}\left[\left(|||\mathcal{D}^0_{-1,0}|||_{1;\gamma}\right)^p\right] \tag{8.111}$$

$$= \mathbb{E}\left[\sup_{t\in[0,1]}\left|t^{1-\gamma}\int_{-1}^{0}(t-r)^{H-\frac{3}{2}}dW_r\right|^p\right]. \tag{8.112}$$

But by Lemma 8.3 and the fact that for every $r \in [-1, 0)$, $t^{1-\gamma} \leqslant (t-r)^{1-\gamma}$, one checks that

$$\left|t^{1-\gamma}\int_{-1}^{0}(t-r)^{H-\frac{3}{2}}dW_r\right| \leqslant C\left(|W_{-1}| + \int_{-1}^{0}(t-r)^{H-\frac{3}{2}-\gamma}\,|W_r|\,dr\right).$$

Since $H - \frac{3}{2} - \gamma \leqslant 0$, we have for every $t \in (0,1]$ $(t-r)^{H-\frac{3}{2}-\gamma} \leqslant (-r)^{H-\frac{3}{2}-\gamma}$. This implies that

$$\int_{-1}^{0}(t-r)^{H-\frac{3}{2}-\gamma}\,|W_r|\,dr \leqslant \int_{-1}^{0}(-r)^{H-1-\gamma}\left((-r)^{-\frac{1}{2}}\,|W_r|\right)dr.$$

Since $\gamma < H$, one can proceed as at the end of Step 2 by making use of the Jensen inequality with respect to the the finite measure $(-r)^{H-1-\gamma}1_{(0,1)}(-r)dr$ to obtain

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|t^{1-\gamma}\int_{-1}^{0}(t-r)^{H-\frac{3}{2}}dW_r\right|^p\right] \leqslant C\left(1+\int_{-1}^{0}(-r)^{H-1-\gamma}\mathbb{E}[(-r)^{-\frac{p}{2}}|W_r|^p]dr\right) \leqslant \tilde{C} < +\infty.$$

This concludes the proof. □

Proposition 8.3. *Suppose that the assumptions of Lemma 8.4 hold with $a_k = a_0\varsigma^k$ for some $\varsigma > 1$. Then, for any $\gamma \in (0, H)$, there exists a constant $C > 0$ such that for every $k \geqslant 1$,*

$$\sup_{k\geqslant 0}\mathbb{E}[|||D^{\tau_k}(W)|||_{1;\gamma}|\tau_k < +\infty] \leqslant C, \quad \sup_{k\geqslant 0}\mathbb{E}[|||D^{\tau_k}(\widetilde{W})|||_{1;\gamma}|\tau_k < +\infty] \leqslant C.$$

Proof. Once again, we only prove the result for W. First, assume that $k \geqslant 1$. By the decomposition (8.106),

$$\mathbb{E}[|||D^{\tau_k}(W)\,|||_{1;\gamma}\,|\mathcal{E}_k] \leqslant C\left(\mathbb{E}[|||\mathcal{D}^{\tau_k}_{-\infty,0}\,|||_{1;\gamma}\,|\mathcal{E}_k] + \sum_{\ell=1}^{k}\mathbb{E}[|||\mathcal{D}^{\tau_k}_{\tau_{\ell-1},\tau_\ell}\,|||_{1;\gamma}\,|\mathcal{E}_k]\right).$$

On the one hand, by the assumption on a_k, one remarks that for every $m \in \{1,\dots,k-1\}$, $\sum_{v=m+1}^{k} a_v \geqslant a_0\varsigma^k$. By Lemma 8.4, one deduces that

$$\sum_{m=1}^{k}\mathbb{E}[|||\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}\,|||_{1;\gamma}\,|\mathcal{E}_k] \leqslant C\left(1+\varsigma^{k(H-1+\varepsilon)}\sum_{m=1}^{k}\left(\frac{1}{\delta_1}\right)^{(k-m)(1-\frac{1}{q})}\right) \leqslant C\left(1+\left(\frac{\varsigma^{(H-1+\varepsilon)}}{\delta_1^{1-\frac{1}{q}}}\right)^k\right).$$

Using that $H-1+\varepsilon < 0$, one then checks that for q large enough,

$$\frac{\varsigma^{(H-1+\varepsilon)}}{\delta_1^{1-\frac{1}{q}}} \leqslant 1$$

so that there exists a constant $C > 0$ such that

$$\sum_{m=1}^{k}\mathbb{E}[|||\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}\,|||_{1;\gamma}\,|\mathcal{E}_k] \leqslant C.$$

Let us now focus on $|||\mathcal{D}^{\tau_k}_{-\infty,0}|||_{1;\gamma}$ with $k \geqslant 0$. In this case, one deduces from (8.108) and from a variant of Lemma 8.3 with $u = -\infty$ and $v = 0$ (involving the fact that $\lim_{u\to-\infty} u^{H-\frac{3}{2}}|W_u| = 0$ *a.s.*) that for every $k \geqslant 0$,

$$\begin{aligned}\mathbb{E}[|||\mathcal{D}^{\tau_k}_{-\infty,0}\,|||_{1;\gamma}\,|\mathcal{E}_k] &\leqslant \delta_1^{k(1-\frac{1}{q})}\mathbb{E}\left[\sup_{t\in(0,1]}t^{1-\gamma}\left(\int_{-\infty}^{0}(t-r)^{H-\frac{5}{2}}|W_{-r}|dr\right)^p\right]^{\frac{1}{p}}.\\ &\leqslant \delta_1^{k(1-\frac{1}{q})}\left(\mathbb{E}\left[\left(\int_{-\infty}^{-1}(-r)^{H-\frac{5}{2}}|W_{-r}|dr\right)^p\right]^{\frac{1}{p}} + \mathbb{E}\left[\left(\int_{-1}^{0}(-r)^{H-1-\gamma}|(-r)^{-\frac{1}{2}}W_{-r}|dr\right)^p\right]^{\frac{1}{p}}\right).\end{aligned}$$

The control of the two above terms follows from similar arguments as in Steps 2 and 3 of the proof of Lemma 8.4, involving Jensen inequality and the scaling property of the Brownian motion. The result follows. □

In the next proposition, we now states a result which achieves the proof of Proposition 8.2 (see (8.105)).

Proposition 8.4. *Let $\rho \in (0,1)$. Under the assumptions of Proposition 8.3, there exists $q > 0$ such that*

$$\sup_{k\geqslant 1}\mathbb{E}\left[\sum_{\ell=0}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\left(\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^q + \|\tilde{\mathbf{X}}^{(\tau_{k-1}+\ell)}\|_\gamma^q\right)|\mathcal{E}_k\right] < +\infty$$

Remark 8.7. In fact, the result could be proved for any $q > 0$. For the sake of simplicity, we only state the result for a given $q > 0$, which is sufficient for our purpose.

Proof. As in the previous results, the proof is identical for $\mathbf{X}$ and $\tilde{\mathbf{X}}$. We thus only focus on $\mathbf{X}$. By (8.107), one first obtains that

$$\mathbb{E}\left[\sum_{\ell=1}^{\Delta\tau_k} \rho^{\Delta\tau_k-\ell}\left(\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma\right)^q |\mathcal{E}_k\right] \leqslant C \sup_{\ell\geqslant 1}\sum_{\ell=1}^{\Delta(\ell,k)} \rho^{\Delta(\ell,k)-\ell}\mathbb{E}\left[\left(\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma\right)^{2q}|\mathcal{E}_{k-1}\right]^{\frac{1}{2}}$$
$$\leqslant \frac{C}{1-\rho}\sup_{\ell,k}\mathbb{E}\left[\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{2q}|\mathcal{E}_{k-1}\right]^{\frac{1}{2}}.$$

Second, by Corollary A.3,

$$\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma \leqslant C(1+\|\mathbf{Z}^{(\tau_{k-1}+\ell)}\|_\gamma^2 + |||D^{(\tau_{k-1}+\ell)}|||_{1;\gamma}^2)\ , \tag{8.113}$$

so that taking $q = 1/4$ and using the elementary inequality $|u+v|^p \leqslant |u|^p + |v|^p$ for $p \in (0,1]$ and $u, v \in \mathbb{R}$, we obtain that

$$\sup_{\ell,k}\mathbb{E}\left[\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{2q}|\mathcal{E}_{k-1}\right] \leqslant \sup_{\ell,k}\mathbb{E}[\|\mathbf{Z}^{(\tau_{k-1}+\ell)}\|_\gamma|\mathcal{E}_{k-1}] + \sup_{\ell,k}\mathbb{E}[|||D^{(\tau_{k-1}+\ell)}|||_{1;\gamma}\,|\mathcal{E}_{k-1}].$$

Owing to the stationarity and the independence of the Brownian increments, it is clear that

$$\mathbb{E}[\|\mathbf{Z}^{(\tau_{k-1}+\ell)}\|_\gamma|\mathcal{E}_{k-1}] = \mathbb{E}[\|\mathbf{Z}\|_\gamma]\ ,$$

where the latter expectation is known to be finite (see [12, Theorem 15.33]).

Second, similarly to (8.106), one remarks that

$$(D^{(\tau_{k-1}+\ell)}(t))' = \alpha_H\left(H-\frac{1}{2}\right)\left(\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t) + \mathcal{D}^{\tau_{k-1}+\ell}_{\tau_{k-1},\tau_{k-1}+\ell}(t)\right),$$

the first and second terms being respectively $\mathcal{F}_{\tau_{k-1}}$-measurable and independent of $\mathcal{F}_{\tau_{k-1}}$. Then, on the one hand, the fact that

$$\mathbb{E}\left[\sup_{t\in[0,1]} t^{1-\gamma}\left|\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t)\right||\mathcal{E}_{k-1}\right]$$

is uniformly bounded in ℓ and k can be proved by following the lines of the proof of Lemma 8.4 and of Proposition 8.3. More precisely, the idea is to use a similar decomposition as that of (8.106) and to control each term with the help of Lemma 8.3. But since the right-hand term of this lemma decreases with τ, the dependency in ℓ can be managed as follows: for every ℓ, for every $u < v$,

$$\left|\mathcal{D}^{\tau_{k-1}+\ell}_{u,v}(t)\right| \leqslant C\left((t+\tau_{k-1}-u)^{H-\frac{3}{2}}|W_v - W_u| + \int_u^v (t+\tau_{k-1}-r)^{H-\frac{5}{2}}|W_v - W_r|\,dr\right),$$

which implies that the sequel of the proof is almost identical to that of Lemma 8.4 and of Proposition 8.3.

Let us finally focus on $\mathcal{D}^{\tau_{k-1}+\ell}_{\tau_{k-1},\tau_{k-1}+\ell}(t)$. Using independency and stationary properties of the Brownian motion and Lemma 8.3, we obtain that for every $\ell \geqslant 1$,

$$\mathbb{E}\left[\sup_{t\in[0,1]} t^{1-\gamma}\left|\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t)\right||\mathcal{E}_{k-1}\right] \leqslant C\mathbb{E}\left[\sup_{t\in[0,1]} t^{1-\gamma}\left((t+\ell)^{H-\frac{3}{2}}|W_\ell| + \int_0^\ell (t+\ell-r)^{H-\frac{5}{2}}|W_r|dr\right)\right].$$

For the first term, it is enough to remark that $t^{1-\gamma}(t+\ell)^{H-\frac{3}{2}}|W_\ell| \leqslant (\ell)^{-\frac{1}{2}}|W_\ell|$ whereas for the second one, one cuts the integral into two parts and the following inequality holds:

$$t^{1-\gamma}\int_0^\ell (t+\ell-r)^{H-\frac{5}{2}}|W_r - W_\ell|dr \leqslant \int_0^{\ell-1}(\ell-r)^{H-2}|(\ell-r)^{-\frac{1}{2}}|W_r - W_\ell|dr$$
$$+\int_{\ell-1}^{\ell}(\ell-r)^{H-1-\gamma}(\ell-r)^{-\frac{1}{2}}|W_r - W_\ell|dr.$$

Note that for the second term, we used that for every $r \in [\ell-1, \ell]$, for every $t \in (0,1]$,

$$t^{1-\gamma}(t+\ell-r)^{H-2} \leqslant (t+\ell-r)^{H-1-\gamma} \leqslant (\ell-r)^{H-1-\gamma}.$$

The (uniform) boundedness of $\mathbb{E}\left[\sup_{t\in[0,1]} t^{1-\gamma}\left|\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t)\right| |\mathcal{E}_{k-1}\right]$ follows. □

9 Proof of Theorem 2.5

Let $\alpha \in (0,1/2)$. We enforce the assumptions of Proposition 8.1. Assume that Y and $\widetilde{Y}$ have initial distributions μ_0 and μ respectively, where μ denotes an invariant distribution. First, denoting by $\bar{\mu}$ its first marginal, we recall that $\int |x|^r \bar{\mu}(dx) < +\infty$ for any positive r. It is therefore enough to show that for any initial condition $\tilde{\mu}$ of $(Y, \widetilde{Y})$ satisfying (4.31) and (4.32) and such that $\mu_0(|\widetilde{x}|^r) < +\infty$ for some $r > 0$, we have for each $\varepsilon > 0$ the existence of a $C_\varepsilon > 0$ such that $\mathbb{P}(\tau_\infty > t) \leqslant C_\varepsilon t^{-(\frac{1}{8}-\varepsilon)}$.

Set $k^* := \inf\{k \geqslant 1, \Delta\tau_k = +\infty\}$. Using that $\tau_0 = 0$, we have

$$\mathbb{P}(\tau_\infty > t) = \mathbb{P}(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t)$$

Let $p \in (0, \alpha/\beta) \subset (0,1)$. By the Markov inequality and the elementary inequality $|u+v|^p \leqslant |u|^p + |v|^p$,

$$\begin{aligned}\mathbb{P}(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t) &\leqslant \frac{1}{t^p}\sum_{k=1}^{+\infty} \mathbb{E}[|\Delta\tau_k|^p 1_{\{k^*>k\}}] \\ &\leqslant \frac{1}{t^p}\sum_{k=1}^{+\infty} \mathbb{E}[\mathbb{E}[|\Delta\tau_k|^p 1_{\{\Delta\tau_k<+\infty\}}|\mathcal{F}_{\tau_{k-1}}]1_{\tau_{k-1}<+\infty}].\end{aligned}$$

Let us bound the above conditional expectation. On the one hand, if Step 1 fails (including the case where $\omega \in A^c_{K,\alpha,\gamma,\tau_{k-1}}$, *i.e.* where Step 1 is not attempted), $\Delta\tau_k = 1 + c_3\varsigma^k$ where $\varsigma > 1$ can be chosen arbitrarily. On the other hand, by Lemma 7.2, we have for every $\ell \geqslant 2$,

$$\mathbb{P}(F_{k,\ell}|\mathcal{F}_{\tau_{k-1}} \cap \{\tau_{k-1} < +\infty\}) \leqslant 2^{-\alpha\ell}.$$

Since by construction, $\Delta\tau_k \leqslant C\varsigma^k 2^{\beta\ell}$ (with $\beta > (1-2\alpha)^{-1}$) on $F_{k,\ell}$, this yields

$$\mathbb{E}[|\Delta\tau_k|^p|1_{\{\Delta\tau_k<+\infty\}}|\mathcal{F}_{\tau_{k-1}} \cap \{\tau_{k-1} < +\infty\}] \leqslant C\varsigma^{kp}\left(\sum_{\ell=1}^{+\infty} 2^{(\beta p-\alpha)\ell}\right) \leqslant C\varsigma^{kp}.$$

Thus, for every $p \in (0, \alpha(1-2\alpha))$,

$$\mathbb{P}(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t) \leqslant \frac{C}{t^p}\sum_{k=1}^{+\infty} \varsigma^{kp}\mathbb{P}(k^* > k-1).$$

But

$$\mathbb{P}(k^* > k-1) = \prod_{m=1}^{k-1} \mathbb{P}(\mathcal{E}_m|\mathcal{E}_{m-1}) = \prod_{m=1}^{k-1}(1 - \mathbb{P}(\mathcal{E}^c_m|\mathcal{E}_{m-1}))$$

and by Proposition 8.1 the latter applied with (for instance) $\varepsilon = 1/2$, we have for every $m \geqslant 1$,

$$\mathbb{P}(\mathcal{E}^c_m|\mathcal{E}_{m-1}) \geqslant \mathbb{P}(\Delta\tau_m = +\infty|A_{K,\alpha,\gamma,\tau_{m-1}})\mathbb{P}(A_{K,\alpha,\gamma,\tau_{m-1}}|\mathcal{E}_{m-1}) \geqslant \frac{\delta_0}{2}$$

where δ_0 is a positive number depending on $K_{\frac{1}{2}}$. It follows that

$$\mathbb{P}(k^* > k-1) \leqslant (1 - \frac{\delta_0}{2})^{k-1}.$$

As a consequence,

$$\sum_{k=1}^{+\infty} \varsigma^{kp} \mathbb{P}(k^* > k-1) \leqslant \sum_{k=1}^{+\infty} \varsigma^{pk} (1 - \frac{\delta_0}{2})^{k-1} < +\infty$$

if ς is chosen in such a way that $\varsigma^p < (1 - \frac{\delta_0}{2})^{-1}$ (This is possible since ς is an arbitrary number greater than 1). Finally, for every $\alpha \in (0, 1/2)$, for every $p \in \alpha(1-2\alpha)$, there exists $C > 0$ such that $\mathbb{P}(\tau_\infty > t) \leqslant C t^{-p}$. To conclude the proof, it remains to optimize in α.

Acknowledgements: We are grateful to Martin Hairer for fruitful advices, and to Peter K. Friz for bringing our attention to the references behind Lemma 6.5(*iii*).

A Singular paths and canonical lift

Let us recall that the space $\mathcal{E}^2_\gamma([0,1];\mathbb{R}^d)$, as well as the notation $|||f|||_{1;\gamma}$, have been introduced in Section 4.1.

Proposition A.1. *Let $z \in \mathcal{C}^\gamma_1([0,1];\mathbb{R}^n)$ be a path that can be canonically lifted into a rough path $\mathfrak{L}(z)$, in the sense of Definition 2.2, and let $g \in \mathcal{E}^2_\gamma([0,1];\mathbb{R}^d)$. Then $z+g$ can be canonically lifted into a rough path $\mathfrak{L}(z+g)$, and it holds that*

$$\mathcal{N}[\mathfrak{L}(z+g)^{\mathbf{2}} - \mathfrak{L}(z)^{\mathbf{2}}; \mathcal{C}^{2\gamma,1+\gamma}_{2,\gamma}([0,1];\mathbb{R}^{n,n})] \lesssim 1 + |||g|||^2_{1;\gamma} + \mathcal{N}[z; \mathcal{C}^\gamma_1([0,1];\mathbb{R}^d)]^2 \ . \tag{A.114}$$

Corollary A.2. *In the setting of Definition 5.2 (fix $\beta := \gamma$), a path $y : [0,1] \to V$ is a solution of*

$$dy_t = B(y_t)\, dh_t + \Sigma(y_t)\, d\mathfrak{L}(z+g)_t \quad , \quad y_0 = v_0 \ ,$$

if and only if y is a solution of

$$dy_t = \big[B(y_t)\, dh_t + \Sigma(y_t)\, dg_t\big] + \Sigma(y_t)\, d\mathfrak{L}(z)_t \quad , \quad y_0 = v_0 \ .$$

Corollary A.3. *It holds that*

$$\|\mathfrak{L}(z+g)\|_{\gamma;[0,1]} \lesssim 1 + \|\mathfrak{L}(z)\|^2_{\gamma;[0,1]} + |||g|||^2_{1;\gamma} \ . \tag{A.115}$$

Lemma A.4. *Let $g \in \mathcal{E}^2_\gamma([0,1];\mathbb{R}^d)$ and denote by g^n the linear interpolation of g along the dyadic partition $\mathcal{P}_n$ of $[0,1]$. Then it holds that*

$$\sup_n \sup_{t \in (0,1] \setminus \mathcal{P}_n} t^{1-\gamma} |(g^n)'_t| \lesssim |||g|||_{1;\gamma} \tag{A.116}$$

and for every $0 < \gamma' < \gamma$,

$$\sup_{t \in (0,1] \setminus \mathcal{P}_n} t^{1-\gamma'} |(g^n - g)'_t| \lesssim |||g|||_{2;\gamma}\, 2^{-n(\gamma-\gamma')} \ . \tag{A.117}$$

Proof. Pick $t \in (t^n_i, t^n_{i+1})$, for some $i = 0, \dots, 2^n$. One has

$$t^{1-\gamma} |(g^n)'_t| = \frac{t^{1-\gamma}}{t^n_{i+1} - t^n_i} |g_{t^n_{i+1}} - g_{t^n_i}| \leqslant |||g|||_{1;\gamma}\, t^{1-\gamma} \int_0^1 \frac{dr}{(t^n_i + r(t^n_{i+1} - t^n_i))^{1-\gamma}} \ .$$

If $i = 0$, then $t \leqslant t^n_{i+1} - t^n_i$ and so $t^{1-\gamma}|(g^n)'_t| \leqslant |||g|||_{1;\gamma} \int_0^1 \frac{dr}{r^{1-\gamma}}$. If $i \geqslant 1$, then $\frac{t}{2} \leqslant \frac{i+1}{2^{n+1}} \leqslant \frac{i}{2^n} = t^n_i$, and so $t^{1-\gamma}|(g^n)'_t| \leqslant |||g|||_{1;\gamma}\, (t/t^n_i)^{1-\gamma} \leqslant |||g|||_{1;\gamma}\, 2^{1-\gamma}$, which completes the proof of (A.116).

For (A.117), note first that if $i = 0$, then $t \leqslant 2^{-n}$ and so by (A.116) we get in this case

$$t^{1-\gamma'}|(g^n - g)'_t| \leqslant 2^{-n(\gamma-\gamma')}\big\{t^{1-\gamma}|(g^n)'_t| + t^{1-\gamma}|g'_t|\big\} \lesssim |||g|||_{1;\gamma}\, 2^{-n(\gamma-\gamma')} \ .$$

If $i \geqslant 1$, then

$$\begin{aligned} t^{1-\gamma'}|(g^n - g)'_t| &= \frac{t^{1-\gamma'}}{t^n_{i+1} - t^n_i}\Big|\int_{t^n_i}^{t^n_{i+1}} \{g'_r - g'_t\}\, dr\Big| \\ &\lesssim |||g|||_{2;\gamma}\, \frac{t^{1-\gamma'}}{t^n_{i+1} - t^n_i}\int_{t^n_i}^{t^n_{i+1}} \frac{|t-u|}{(t^n_i)^{2-\gamma}}\, du \\ &\lesssim |||g|||_{2;\gamma}\left(\frac{t}{t^n_i}\right)^{1-\gamma'}\left(\frac{t^n_{i+1} - t^n_i}{t^n_i}\right)^{1-(\gamma-\gamma')} 2^{-n(\gamma-\gamma')} \ . \end{aligned}$$

As above, we can conclude by using the fact that in this case, one has $\max\big(\frac{t}{2}, t^n_{i+1} - t^n_i\big) \leqslant t^n_i$. □

Proof of Proposition A.1. Denote by z^n, resp. g^n, the linear interpolation of z, resp. g, along the dyadic partition $\mathcal{P}_n$. By (A.117), the convergence of g^n to g (and accordingly the convergence of $z^n + g^n$ to $z + g$) in $\mathcal{C}_1^{\gamma'}([0,1];\mathbb{R}^d)$ is immediate, since

$$|\delta(g^n - g)_{st}| \leqslant \int_s^t |(g^n - g)'_u|\, du \lesssim 2^{-n(\gamma-\gamma')}(t-s)\int_0^1 \frac{dr}{(s + r(t-s))^{1-\gamma'}} \lesssim 2^{-n(\gamma-\gamma')}(t-s)^{\gamma'} \ .$$

Then, by setting $x := z + g$ and using the notation (2.7), we have the following readily-checked decomposition

$$\mathbf{x}^{\mathbf{2},n}_{st} - \mathbf{z}^{\mathbf{2},n}_{st} = \int_s^t (\delta z^n)_{su} \otimes dg^n_u + \left(\int_s^t (\delta z^n)_{ut} \otimes dg^n_u\right)^* + \int_s^t (\delta g^n)_{su} \otimes dg^n_u \ . \tag{A.118}$$

Now consider the integral $\int_s^t (\delta z)_{su} \otimes dg_u$, which, due to the regularity of g, can be interpreted in the classical Lebesgue sense, and use (A.116)-(A.117) to assert that

$$\begin{aligned} &\Big|\int_s^t (\delta z^n)_{su} \otimes dg^n_u - \int_s^t (\delta z)_{su} \otimes dg_u\Big| \\ &\leqslant \int_s^t |\delta(z^n - z)_{su}| \otimes |dg^n_u| + \int_s^t |(\delta z)_{su}| \otimes |d(g^n - g)_u| \\ &\lesssim \mathcal{N}[z^n - z; \mathcal{C}_1^{\gamma'}([0,1];\mathbb{R}^d)] \int_s^t \frac{|u-s|^{\gamma'}}{u^{1-\gamma}}\, du + \mathcal{N}[z; \mathcal{C}_1^{\gamma}([0,1];\mathbb{R}^d)] 2^{-n(\gamma-\gamma')} \int_s^t \frac{|u-s|^{\gamma}}{u^{1-\gamma'}}\, du \\ &\lesssim |t-s|^{\gamma+\gamma'}\big\{\mathcal{N}[z^n - z; \mathcal{C}_1^{\gamma'}([0,1];\mathbb{R}^d)] + 2^{-n(\gamma-\gamma')}\big\} \int_0^1 \frac{dr}{r^{1-(\gamma+\gamma')}} \ . \end{aligned}$$

We can treat the two other summands in (A.118) along the same lines, which leads us to the desired conclusion, namely $\mathcal{N}[\mathbf{x}^{\mathbf{2},n} - \mathbf{z}^{\mathbf{2},n}; \mathcal{C}_2^{2\gamma'}([0,1];\mathbb{R}^d)] \to 0$ as $n \to \infty$. We even get the explicit description

$$\mathfrak{L}(z+g)^{\mathbf{2}}_{st} - \mathfrak{L}(z)^{\mathbf{2}}_{st} = \int_s^t (\delta z)_{su} \otimes dg_u + \left(\int_s^t (\delta z)_{ut} \otimes dg_u\right)^* + \int_s^t (\delta g)_{su} \otimes dg_u \ .$$

With this decomposition in hand, it is now easy to exhibit the bound (A.114): for instance, for every

$0 < s < t$,

$$\begin{aligned}
&\left| \int_s^t (\delta z)_{su} \otimes dg_u \right| \\
&\quad \leqslant \quad |||g|||_{1;\gamma} \, \mathcal{N}[z; \mathcal{C}_1^\gamma([0,1];\mathbb{R}^d)] \int_s^t \frac{|u-s|^\gamma}{u^{1-\gamma}} \, du \\
&\quad \leqslant \quad |||g|||_{1;\gamma} \, \mathcal{N}[z; \mathcal{C}_1^\gamma([0,1];\mathbb{R}^d)] \, |t-s|^{1+\gamma} \int_0^1 \frac{r^\gamma}{(s + r(t-s))^{1-\gamma}} \, dr \\
&\quad \leqslant \quad |||g|||_{1;\gamma} \, \mathcal{N}[z; \mathcal{C}_1^\gamma([0,1];\mathbb{R}^d)] \, \min\left(|t-s|^{2\gamma} \int_0^1 \frac{dr}{r^{1-2\gamma}}, s^{\gamma-1}|t-s|^{1+\gamma} \int_0^1 r^\gamma \, dr \right) .
\end{aligned}$$

□

B Proof of Lemma 5.6

The argument relies on the algorithm introduced in [7, Section 6] and which aims at "removing the points one by one" between t_p and t_{q+1} in a tricky way. First, just as in [7, Section 3.1], and given any (not necessarily uniform) subpartition Π of $\mathcal{P}_n$, we define the path G^Π as follows: for every $s \leqslant t \in \mathcal{P}_n$,

$$G^\Pi_{st} := \begin{cases} 0 & \text{if } (s,t) \cap \Pi = \varnothing \\ (\delta G)_{sut} & \text{if } (s,t) \cap \Pi = u \\ G_{st} - G_{s\tilde{t}_1} - \sum_{k=1}^{\ell-1} G_{\tilde{t}_k \tilde{t}_{k+1}} - G_{\tilde{t}_\ell t} & \text{if } (s,t) \cap \Pi = \{\tilde{t}_1, ..., \tilde{t}_\ell\} \end{cases} .$$

With this notation, if $s = t_p$ and $t = t_{q+1}$, one has in particular

$$G_{st} = G^{[\![s,t]\!]}_{st} + \sum_{i=p}^{q} G_{t_i t_{i+1}} \, . \tag{B.119}$$

As far as the sum is concerned, we have on the one hand, since $\mu_1 \geqslant 1$,

$$s^{1-\lambda} \| \sum_{i=p}^{q} G_{t_i t_{i+1}} \| \leqslant \mathcal{M}_\lambda^{\alpha,\mu_1}\big[G; [\![s,t]\!]\big] \cdot \sum_{i=p}^{q} |t_{i+1} - t_i|^{\mu_1} \leqslant \mathcal{M}_\lambda^{\alpha,\mu_1}\big[G; [\![s,t]\!]\big] \cdot |t-s|^{\mu_1} \, ,$$

and on the other hand

$$\| \sum_{i=p}^{q} G_{t_i t_{i+1}} \| \leqslant \mathcal{M}_\lambda^{\alpha,\mu_1}\big[G; [\![s,t]\!]\big] \cdot \left\{ |t_{p+1} - s|^\alpha + \sum_{i=p+1}^{q} t_i^{\lambda-1} |t_{i+1} - t_i|^{\mu_1} \right\} ,$$

with

$$\sum_{i=p+1}^{q} t_i^{\lambda-1} |t_{i+1} - t_i|^{\mu_1} = \frac{1}{2^{n(\lambda+\mu_1-1)}} \sum_{i=p+1}^{q} \frac{1}{i^{1-\lambda}} \lesssim \frac{1}{2^{n(\lambda+\mu_1-1)}} |q+1-p|^\lambda \lesssim |t-s|^\lambda \, .$$

Going back to (B.119), it remains us to bound $\|G^{[\![s,t]\!]}_{st}\|$. For the sake of clarity, let us temporarily change the notation by setting, for s, t fixed as above,

$$t_k := t - \frac{k}{2^n} \, , \; k = 0, \dots, N \, , \quad \text{where } N := 2^n(t-s) \;\; (= q+1-p) \, . \tag{B.120}$$

We make this (unnatural) choice to "reverse" the time, that is to consider a decreasing function $k \mapsto t_k$, in a such a way that the below notations will be consistent with those of [7, Section 6] (and especially those of [7, Proposition 6.2]). Consider indeed the algorithm described in [7, Section 6] to remove one by one the points between 0 and N, and accordingly the points of $\mathcal{P}_n$ between s and t (just use the transformation (B.120) to

connect one with the other). Denote by $(\Pi^m)_{m=0,\dots,N-1}$ the decreasing sequence of partitions of $[\![s,t]\!]$ that is associated with this algorithm. With the notations of [7, Section 6], it is readily checked that

$$G_{st}^{\Pi^m} - G_{st}^{\Pi^{m+1}} = (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} \quad , \quad G_{st}^{\Pi^0} = G_{st}^{[\![s,t]\!]} \quad , \quad G_{st}^{\Pi^{N-1}} = 0 \ ,$$

and so

$$G_{st}^{[\![s,t]\!]} = \sum_{m=1}^{N-1} (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} \ . \tag{B.121}$$

Now, still with the notations of [7, Section 6] in mind, write

$$\sum_{m=0}^{N} (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} = \sum_{r=1}^{M-1} \Big\{ (\delta G)_{s t_{k_{A_{r-1}+1}} t_{k_{A_{r-1}+1}^-}} + \sum_{m=A_{r-1}+2}^{A_r} (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} \Big\}$$

and so

$$\begin{aligned} &\| \sum_{m=0}^{N} (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} \| \\ &\leqslant \ \mathcal{N}[\delta G; \mathcal{C}_{3;\lambda}^{\alpha,\mu_2}([\![s,t]\!])] \cdot \sum_{r=1}^{M-1} \Big\{ |t_{k_{A_{r-1}+1}^-} - s|^\alpha + \sum_{m=A_{r-1}+2}^{A_r} t_{k_m^+}^{\lambda-1} |t_{k_m^-} - t_{k_m^+}|^{\mu_2} \Big\} \ . \end{aligned} \tag{B.122}$$

Observe at this point that

$$|t_{k_{A_{r-1}+1}^-} - s|^\alpha = |t-s|^\alpha \cdot \Big| 1 - \frac{k_{A_{r-1}+1}^-}{N} \Big|^\alpha$$

and

$$t_{k_m^+}^{\lambda-1} |t_{k_m^-} - t_{k_m^+}|^{\mu_2} \leqslant |t-s|^{\lambda+\mu_2-1} \cdot \frac{1}{N^{\mu_2}} \Big| 1 - \frac{k_m^+}{N} \Big|^{\lambda-1} |k_m^+ - k_m^-|^{\mu_2} \ .$$

Going back to (B.122), we get that

$$\| \sum_{m=0}^{N} (\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}} \| \leqslant |t-s|^\alpha \, \mathcal{N}[\delta G; \mathcal{C}_{3;\lambda}^{\alpha,\mu_2}([\![s,t]\!])] \cdot Q_{\alpha,\lambda,\mu_2}^N \ , \tag{B.123}$$

where we have set

$$Q_{\alpha,\lambda,\mu_2}^N := \sum_{r=1}^{M-1} \Big\{ \Big| 1 - \frac{k_{A_{r-1}+1}^-}{N} \Big|^\alpha + \frac{1}{N^{\mu_2}} \sum_{m=A_{r-1}+2}^{A_r} \Big| 1 - \frac{k_m^+}{N} \Big|^{\lambda-1} |k_m^+ - k_m^-|^{\mu_2} \Big\} \ .$$

Therefore, we are exactly in a position to apply [7, Proposition 6.2] and assert that $\sup_{N \geqslant 1} Q_{\alpha,\lambda,\mu_2}^N < \infty$. The combination of (B.121) and (B.123) then gives us the desired estimate, namely

$$\|G_{st}^{[\![s,t]\!]}\| \lesssim |t-s|^\alpha \, \mathcal{N}[\delta G; \mathcal{C}_{3;\lambda}^{\alpha,\mu_2}([\![s,t]\!])] \ .$$

The estimation of $s^{1-\lambda} \|G_{st}^{[\![s,t]\!]}\|$ is easier. Indeed, with decomposition (B.121) in mind, we simply use the fact that the above algorithm also satisfies

$$|k_m^+ - k_m^-| \leqslant \frac{2N}{(N-m+1)} \quad \text{for every } m = 1, \dots, N-1 \ ,$$

and consequently

$$\begin{aligned} s^{1-\lambda} \|G_{st}^{[\![s,t]\!]}\| \ &\leqslant \ \sum_{m=1}^{N-1} t_{k_m^+}^{1-\lambda} \|(\delta G)_{t_{k_m^+} t_{k_m} t_{k_m^-}}\| \\ &\leqslant \ \mathcal{N}[\delta G; \mathcal{C}_{3;\lambda}^{\alpha,\mu_2}([\![s,t]\!])] \cdot \sum_{m=1}^{N-1} |t_{k_m^-} - t_{k_m^+}|^{\mu_2} \ \lesssim \ |t-s|^{\mu_2} \, \mathcal{N}[\delta G; \mathcal{C}_{3;\lambda}^{\alpha,\mu_2}([\![s,t]\!])] \ . \end{aligned}$$

References

[1] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[2] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[3] Serge Cohen and Fabien Panloup. Approximation of stationary solutions of Gaussian driven stochastic differential equations. *Stochastic Process. Appl.*, 121(12):2776–2801, 2011.

[4] Serge Cohen, Fabien Panloup, and Samy Tindel. Approximation of stationary solutions to SDEs driven by multiplicative fractional noise. *Stochastic Process. Appl.*, 124(3):1197–1225, 2014.

[5] Hans Crauel. Non-Markovian invariant measures are hyperbolic. *Stochastic Process. Appl.*, 45(1):13–28, 1993.

[6] A. M. Davie. Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express. AMRX*, (2):Art. ID abm009, 40, 2007.

[7] Aurélien Deya. A discrete approach to rough parabolic equations. *Electron. J. Probab.*, 16:no. 54, 1489–1518, 2011.

[8] D. Down, S.P. Meyn, and R.L. Tweedie. Exponential and uniform ergodicity of markov processes. *The Annals of Probability*, 23:1671–1691, 1995.

[9] Joaquin Fontbona and Fabien Panloup. Rate of convergence to equilibrium of fractional driven stochastic differential equations with some multiplicative noise. *Preprint available at http://hal.archives-ouvertes.fr/hal-00989414*, pages 1–34, 2014.

[10] Peter K. Friz, Benjamin Gess, Archil Gulisashvili, and Sebastian Riedel. The Jain–Monrad criterion for rough paths and applications to random Fourier series and non-Markovian Hörmander theory. *Ann. Probab.*, 44(1):684–738, 2016.

[11] Peter K. Friz and Martin Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. With an introduction to regularity structures.

[12] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.

[13] María J. Garrido-Atienza, Peter E. Kloeden, and Andreas Neuenkirch. Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.*, 60(2):151–172, 2009.

[14] Paolo Guasoni. No arbitrage under transaction costs, with fractional Brownian motion and beyond. *Math. Finance*, 16(3):569–582, 2006.

[15] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.

[16] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.

[17] Martin Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.*, 33(2):703–758, 2005.

[18] Martin Hairer and Jonathan C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[19] Martin Hairer and Alberto Ohashi. Ergodic theory for SDEs with extrinsic memory. *Ann. Probab.*, 35(5):1950–1977, 2007.

[20] Martin Hairer and Natesh S. Pillai. Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. *Ann. Probab.*, 41(4):2544–2598, 2013.

[21] Jae-Hyung Jeon, Vincent Tejedor, Stas Burov, Eli Barkai, Christine Selhuber-Unkel, Kirstine Berg-Sørensen, Lene Oddershede, and Ralf Metzler. *In Vivo* anomalous diffusion and weak ergodicity breaking of lipid granules. *Phys. Rev. Lett.*, 106:048103, Jan 2011.

[22] S. C. Kou. Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *Ann. Appl. Stat.*, 2(2):501–535, 2008.

[23] Jonathan C. Mattingly. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Comm. Math. Phys.*, 230(3):421–462, 2002.

[24] David J. Odde, Elly M. Tanaka, Stacy S. Hawkins, and Helen M. Buettner. Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. *Biotechnology and Bioengineering*, 50(4):452–461, 1996.

[25] Michael Röckner and Feng-Yu Wang. Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. *J. Funct. Anal.*, 185(2):564–603, 2001.