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

Aurélien Deya, Fabien Panloup, Samy Tindel

The Annals of Probability
2019, Vol. 47, No. 1, 464–518
https://doi.org/10.1214/18-AOP1265

RATE OF CONVERGENCE TO EQUILIBRIUM OF FRACTIONAL DRIVEN STOCHASTIC DIFFERENTIAL EQUATIONS WITH ROUGH MULTIPLICATIVE NOISE

BY AURÉLIEN DEYA, FABIEN PANLOUP AND SAMY TINDEL

Université de Lorraine, Université d'Angers and Purdue University

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 [*Ann. Probab.* **33** (2005) 703–758] 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 [*Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 503–538], 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 [*Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 503–538] by deleting the gradient assumption on the noise coefficient σ. The main theorems include some existence and uniqueness results for the invariant distribution.

CONTENTS

Received October 2016; revised November 2017.
MSC2010 subject classifications. 60G22, 37A25.
Key words and phrases. Stochastic differential equations, fractional Brownian motion, multiplicative noise, ergodicity, rate of convergence to equilibrium, Lyapunov function, total variation distance.

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 $(Y_t)_{t\geq 0}$ 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

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

Under condition (1.2) (and the nondegeneracy 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:

(i) *Functional inequalities.* Starting from Poincaré-type inequalities (or further refinements) for the solution of (1.1), and invoking Dirichlet form techniques, exponential and subexponential rates of convergence are obtained, for example, in [3, 28].

(ii) *Lyapunov/Coupling techniques.* In these methods (see, e.g., [10]), 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 nondegeneracy 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 nondegeneracy condition on the noisy part of the equation. Finally, observe that the first approach generally leads to sharper exponents but may require 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 [20] handles for instance infinite dimensional situations where only asymptotic couplings of the process (starting from different initial conditions) are available. Let us also mention [19], 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 [2, 6, 15]. 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 nonstandard 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., [16, 23, 24, 27]), 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 $(X_t)_{t\geq 0}$ 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 [18], Hairer and Ohashi [21] and by Hairer and Pillai [22], respectively, in the additive noise, multiplicative noise with $H > 1/2$ and multiplicative hypoelliptic noise with $H \in (1/3, 1/2)$. Except [18] which also deals with rate of convergence to equilibrium, these papers mainly focus on a way to define stationary solutions, and on extending tools of the ergodic 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 [4, 5] for some results on approximations of stationary solutions. In all those articles, the Markovian formalism is based on the Mandelbrot–Van Ness representation of the fractional Brownian motion, namely

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

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 $(Y_t, (X_{s+t})_{s\leq 0})_{t\geq 0}$ can be realized through a Feller transformation $(\mathcal{Q}_t)_{t\geq 0}$ 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 $(Y_0, (X_s)_{s\leq 0})$ where $(X_s)_{s\leq 0}$ 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 $t \geq 0$. As mentioned above, the uniqueness of such an invariant distribution is investigated in [18, 21, 22].

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 [18], 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

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

The upper bound above is believed to be nonoptimal, though its subexponential 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 [18] 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 $(\mathcal{Q}_t)_{t\geq 0}$. The main step consists (classically) in finding a stopping time τ_∞ such that $(Y^{\mu_0}_{t+\tau_\infty})_{t\geq 0} = (Y^{\mu}_{t+\tau_\infty})_{t\geq 0}$. The rate of convergence in total variation is then obtained by means of an accurate bound on $\mathbb{P}(\tau_\infty > t)$, $t \geq 0$.

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 [11], 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 [11] in two different directions:

(i) We get rid of the gradient type hypothesis assumed in [11], 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.

One point should be made clear right now: the techniques displayed in this paper can cover both the case $H \in (1/3, 1/2)$ [as mentioned in point (ii) above] and the case $H > 1/2$ [thus extending the results of [11] 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 2.7 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 nontrivial coupling of the noises. 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 nonconstant coefficient σ, the implementation of the coupling strategy requires some nontrivial 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, which provides existence and uniqueness of an invariant distribution for the rough equation (1.3), as well as a rate of convergence toward this distribution. The Markov setting for this equation, as well as Lyapunov-type inequalities, are given in Section 3, leading to the proof of the existence statement. Our global strategy to get uniqueness and the convergence rate is explained at Section 4. The end of the proof of Theorem 2.5 can thus be found in Section 4.3, slightly anticipating the technical results of the subsequent sections. The setting and results related to the so-called singular rough equations are introduced in Section 5, and then applied at Section 6 to some specific *hitting* system. Eventually, the controls associated with Step 2 and Step 3 of the procedure are exhibited in Sections 7 and 8, respectively.

Throughout the paper, we will use the convenient notation $A \lesssim B$ for any estimate of the form $A \leq cB$, where $c > 0$ is a constant that does not depend on the parameters under consideration.

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 original formulation of rough paths theory, introduced in [17] (see also [13]). 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 $k \geq 1$, 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 $i \leq k - 1$. 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

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

for any $s < u < t \in I$. Besides, throughout the paper, we will use the notation $(\delta \mathcal{T})_{st} = t - s$ for any $s < t$.

Our analysis will rely on some regularity considerations 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

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

and

$$\mathcal{C}_2^\mu(I; V) = \big\{f \in \mathcal{C}_2(I; V); \mathcal{N}\big[f; \mathcal{C}_2^\mu(I; V)\big] < \infty\big\}.$$

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

$$\mathcal{N}\big[f; \mathcal{C}_1^\mu(I; V)\big] = \mathcal{N}\big[\delta f; \mathcal{C}_2^\mu(I; V)\big]$$

and

$$\mathcal{C}_1^\mu(I; V) = \big\{f \in \mathcal{C}_1(I; V); \mathcal{N}\big[f; \mathcal{C}_1^\mu(I; V)\big] < \infty\big\}.$$

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 $i, j \in \{1, \dots, d\}$,

$$\delta \mathbf{x}^{\mathbf{2};ij}_{sut} = \delta x^i_{su} \delta x^j_{ut} \quad \text{and} \quad \mathbf{x}^{\mathbf{2};ij}_{st} + \mathbf{x}^{\mathbf{2};ji}_{st} = \delta x^i_{st} \delta x^j_{st}. \tag{2.1}$$

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

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

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 $\mathbf{x} := (x, \mathbf{x}^{\mathbf{2}})$ above x such that the sequence $\mathbf{x}^n := (x^n, \mathbf{x}^{\mathbf{2},n})$ defined by

$$\mathbf{x}^{\mathbf{2},n}_{st} := \int_s^t (\delta x^n)_{su} \otimes dx^n_u \tag{2.2}$$

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

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

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 [7]). 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 $\mathbf{x} := (x, \mathbf{x}^{\mathbf{2}})$ be a γ-rough path. Then, for any continuous, respectively, differentiable, vector field

$$b : \mathbb{R}^d \to \mathbb{R}^d, \qquad \text{resp. } \sigma : \mathbb{R}^d \to \mathcal{L}\big(\mathbb{R}^d; \mathbb{R}^d\big),$$

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

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\mathbf{x}_t, \qquad y_{t_0} = a, \tag{2.3}$$

if the two-parameter path R^y defined as

$$R^y_{st} := (\delta y)_{st} - b(y_s)(\delta \mathcal{T})_{st} - \sigma_j(y_s)\big(\delta x^j\big)_{st} - (D\sigma_j \cdot \sigma_k)(y_s)\mathbf{x}^{\mathbf{2},jk}_{st}$$

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

$$(D\sigma_i \cdot \sigma_k)(v) := (D\sigma_j)(v)\big(\sigma_k(v)\big) \qquad \text{for every } v \in \mathbb{R}^d. \tag{2.4}$$

Applications of the abstract rough paths setting to a fractional Brownian motion X depends on a proper construction of the Lévy area $\mathbf{X}^{\mathbf{2}}$. The reader is referred to [14], 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 $0 \le s < t \le T$, the random variable $\mathbf{X}^{\mathbf{2}}_{st}$ satisfies the following inequality*:

$$\mathbf{E}\big[|\mathbf{X}^{\mathbf{2}}_{st}|^p\big] \le c_p (t-s)^{2Hp}, \qquad p \ge 1.$$

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). *$b : \mathbb{R}^d \to \mathbb{R}^d$, respectively, $\sigma : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$, is a $\mathcal{C}^3$, respectively, $\mathcal{C}^4$, vector field such that*

$$\begin{aligned} &\sup_{v\in\mathbb{R}^d} \big\|\big(D^{(\ell)}b\big)(v)\big\| < \infty \qquad \text{for } \ell \in \{1,2,3\}, \\ \text{resp. } &\sup_{v\in\mathbb{R}^d} \big\|\big(D^{(\ell)}\sigma\big)(v)\big\| < \infty \qquad \text{for } \ell \in \{0,\dots,4\}. \end{aligned} \tag{2.5}$$

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*

$$\big\langle v, b(v)\big\rangle \le C_1 - C_2\|v\|^2. \tag{2.6}$$

Finally, one needs a nondegeneracy assumption on σ.

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

$$\sup_{x\in\mathbb{R}^d} \big\|\sigma(x)^{-1}\big\| < +\infty. \tag{2.7}$$

We are now in a position to state our main result. One denotes by $\mathcal{L}((Y_t^{\mu_0})_{t\ge 0})$ the distribution of the process Y on $\mathcal{C}([0,+\infty),\mathbb{R}^d)$ starting from a (generalized) initial condition μ_0 (see Section 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 $\bar{\mu}_0(dx)$ stands for the first marginal of $\mu_0(dx,dw)$. Finally, the total variation norm is classically denoted by $\|\cdot\|_{\mathrm{TV}}$.

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*

$$\big\|\mathcal{L}\big((Y_{t+s}^{\mu_0})_{s\ge 0}\big) - \bar{\mathcal{Q}}\mu\big\|_{\mathrm{TV}} \le C_\varepsilon t^{-(\frac{1}{8}-\varepsilon)}. \tag{2.8}$$

In particular,

$$\|\mathcal{L}(Y_t^{\mu_0}) - \bar{\mu}\|_{\mathrm{TV}} \le C_\varepsilon t^{-(\frac{1}{8}-\varepsilon)}, \tag{2.9}$$

where $\bar{\mu}$ denotes the first marginal of μ.

REMARK 2.6. 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 [14] for the general theory of differential equations driven by a rough path. We prove existence of the invariant distribution below in Theorem 3.5 and Corollary 3.6. 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 2.7. As mentioned before, when $H > 1/2$, Theorem 2.5 has already be shown in [11]. 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 nonintegrability 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.

REMARK 2.8. The regularity assumptions contained in Hypothesis (H1) are relatively standard in the setting of rough paths theory. To be more specific, equation (2.3) is known to have a unique global solution provided $b \in \mathcal{C}^2$, $\sigma \in \mathcal{C}^3$, and both vector fields are bounded with bounded derivatives (see, e.g., [14], Theorem 12.10). We shall show on the one hand (Section 3) that the boundedness of b is actually not an essential condition in the latter well-posedness statement. On the other hand, assuming here that $b \in \mathcal{C}^3$, respectively, $\sigma \in \mathcal{C}^4$, instead of $b \in \mathcal{C}^2$, respectively, $\sigma \in \mathcal{C}^3$, is due to the specific expression of the rough coupling system we will consider in our procedure [see (4.43) or (6.1)–(6.2), and observe how the first-order derivatives of b and σ are involved in the second line of these rough systems].

REMARK 2.9. The exponent $-1/8$ in (2.9) may be surprising at first sight, although it is the very same exponent as in the additive case. Indeed, this value does not depend on the Hurst parameter H as one might expect. In fact, this exponent is the result of an optimization of t^{-p} with $p \in (0, \gamma(1-2\gamma))$ and $\gamma \in (0, H \wedge 1/2)$. More precisely, when $H \in (1/3, 1/2)$, $\frac{1}{8} = \max\{\gamma(1-2\gamma), \gamma \in (0, H)\}$ (see Section 4.3 for details). Since this maximum is attained for $\gamma = 1/4$, this explains the explicit dependency in H when $H < 1/4$ in [18] (in the additive setting). Observe

that the term t^{-p} alluded to above is obtained through a procedure which consists in gluing two solutions starting from different initial conditions after they meet (see Section 4.2 for a precise description of this step). This procedure is highly nontrivial in our non-Markovian setting, and the polynomial exponent $-1/8$ has thus to be interpreted as an effect of the memory of equation (1.2).

Furthermore, in view of the Markovian rate of convergence which is typically exponential, one can wonder about the optimal one in the fractional setting. To this aim, notice first that in the specific setting of additive fractional SDEs whose drift term satisfies, for all $(x, y) \in \mathbb{R}^2$,

$$\langle b(x) - b(y), x - y \rangle \leq \alpha \|x - y\|^2, \tag{2.10}$$

a simple Gronwall argument shows that the L^2-distance between solutions built with the same fBm decreases exponentially (see, e.g., [15]). Then it can be deduced that the rate of convergence to the stationary regime is exponential in Wasserstein distance (this property extends in fact to the total variation distance through a simple coalescent coupling between $t - 1$ and t). However, this setting is not representative of the fractional setting since the rate is entirely obtained with the help of the drift condition (2.10) without any probabilistic property. A more typical ergodic situation can be obtained by considering $X_t = [x + B_t^H]$, that is, the fractional part of a fBm starting from x. In this setting, the fact that $B_t^H \sim \mathcal{N}(0, t^H)$ allows us to deduce from a scaling and from the exponential convergence of the fractional part of the Brownian motion to the uniform distribution on $[0, 1]$ that $(\mathcal{L}(X_t))_{t \geq 0}$ also converges to the uniform distribution on $[0, 1]$ with a rate proportional to $\exp(-t^H)$. The interest of this example is to show that a dependency in H appears but is still very particular since it derives from a scaling property. It is also worth noting that this leads to a marginal result like (2.9) but not to (2.8). In the general fractional setting, whenever the drift satisfies the weak contracting relation (2.6) instead of (2.10), the question of the typical rate of convergence is still open. However, our opinion is that exponential or subexponential rates are out of reach with the coupling method recalled in the Introduction and inspired by [18]. The functional approach to convergence to equilibrium remains largely unexplored for fractional Brownian motion driven systems, and it might yield better results. Nevertheless, this method seems to rely too heavily on semi-group techniques to be implemented in a proper way.

3. Existence of invariant distribution. The main result of this section is Theorem 3.5 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 [22] (going back to [18] and [21]), the system (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 $\mathcal{C}_0^\infty(\mathbb{R}_-)$ be the space of $\mathcal{C}^\infty$-functions w, with compact support on $\mathbb{R}_-$ and 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 $\|\cdot\|_{\mathcal{W}_\gamma}$ defined by

$$\|x\|_{\mathcal{W}_\gamma} := \sup_{s,t\in\mathbb{R}_-} \frac{|x(t)-x(s)|}{|t-s|^\gamma (1+|t|+|s|)^{\frac{1}{2}}}. \tag{3.1}$$

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 $\mathcal{C}_0^\infty([0,1])$ with respect to the norm $\|\cdot\|_{\gamma;[0,1]}$. Then, with the help of operators related to the Mandelbrot representation (see [22] 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

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

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 [14] and Hypothesis (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: $\mathcal{Q}(z,x,\cdot) := \Psi_z^*\mathcal{P}(x,\cdot)$ where $\Psi_z(x) := (\Phi_1(z,\mathbf{x}_+),x)$.

REMARK 3.1. In [18], 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 $(X_t)_{t\geq0}$. However, an adaptation of the previous strategy leads to the construction of a Feller Markov semi-group $(\mathcal{Q}_t)_{t\geq0}$ above $(Y_t)_{t\geq0}$ (on $\mathbb{R}^d\times\mathcal{W}_\gamma$).

3.1.2. *Invariant distribution.* Following [21], 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 3.2. Let ν be a generalized initial condition. We say that ν is an invariant distribution for $(Y_t)_{t\geq 0}$ if for every $t \geq 0$, $\nu\mathcal{Q}_t = \nu$.

DEFINITION 3.3. 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 $t \in [0, 1]$,

$$\int V(x)(\mu\mathcal{Q}_t)(dx, dw) \leq C + \rho \int V(x)\mu(dx, dw) \tag{3.2}$$

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

We have the following (classical) result.

PROPOSITION 3.4. *The existence of a Lyapunov function V for the Feller semi-group $(\mathcal{Q}_t)_{t\geq 0}$ implies the existence of an invariant distribution for $(Y_t)_{t\geq 0}$. Furthermore, for any generalized initial condition μ such that*

$$\int V(x)\mu(dx, dw) < +\infty, \tag{3.3}$$

one has $\sup_{t\geq 0} \mathbb{E}_\mu[V(Y_t)] < +\infty$.

PROOF. Let μ denote an initial condition on $\mathbb{R}^d \times \mathcal{W}_-$ such that (3.3) is satisfied and denote by $(\mu_t)_{t\geq 1}$ the sequence defined by

$$\mu_t = \frac{1}{t}\int_0^t \mu\mathcal{Q}_s\, ds, \qquad t \geq 1. \tag{3.4}$$

By construction and by the Feller property, every weak limit of $(\mu_t)_{t\geq 1}$ is an invariant distribution for $(\mathcal{Q}_t)_{t\geq 0}$. It is thus enough to prove the tightness of $(\mu_t)_{t\geq 1}$: 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 $(\nu_t)_{t\geq 0}$, ν_t being the first marginal of μ_t. But the definition of the Lyapunov function implies (by an iteration) $\sup_{t\geq 1} \nu_t(V) < +\infty$ 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

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\mathbf{x}_t, \qquad t \in [0,1],\ y_0 = a \in \mathbb{R}^d, \tag{3.5}$$

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 3.5. *Under Hypothesis (H1) and for every initial condition $y_0 \in \mathbb{R}^d$, equation (3.5) 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 $b, \sigma, \gamma, C_1, C_2$, but not on $\mathbf{x}$) such that*

$$\|y_1\|^2 \le e^{-C_2/2}\|y_0\|^2 + C\{1 + \|\mathbf{x}\|_\gamma^\mu\} \qquad \text{with } \mu := \frac{8}{3\gamma - 1}. \tag{3.6}$$

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 (3.2) is immediate. It is indeed a well-known fact (see, for instance, [14], 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 3.6. *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 $p \ge 1$. As a consequence, there exists at least one invariant distribution ν for $(Y_t)_{t\ge 0}$, which additionally admits finite moments of any order.*

In the same spirit as in Davie's work [7] (see also [14], Chapter 10), our proof of Theorem 3.5 relies on a careful examination of the natural discrete scheme associated with the rough equation, that is the approximation y^n defined along the formula

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

for some partition $(t_i)_{0\le i\le n}$ of $[0,1]$. The idea (morally) is to apply the standard rough-paths control strategy to both y^n and the related quadratic path $z^n := \frac{1}{2}\|y^n\|^2$, and involve Hypothesis (H2) at some specific point of the procedure. As we do not want these technical estimates to compromise the overall clarity of the presentation, we have provided them in the Supplementary Material [9], Section 1, attached to this paper.

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 toward the stationary solution. This result is based on a coupling method first introduced in [18]. We thus begin by recalling the details of the strategy. To this end, we first introduce some notation 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 $(X_t)_{t\geq 0}$ [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

$$X_t = D_t + Z_t, \tag{4.1}$$

where the process D defined by

$$D_t := \alpha_H \int_{-\infty}^{0} \{(t-r)^{H-\frac{1}{2}} - (-r)^{H-\frac{1}{2}}\}\, dW_r$$

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 the smoothness of D was also already used in the past as a central ingredient while studying integration issues with respect to the fBm (see, e.g., [1, 29]), and similar ideas can be found in [22] as well. All along the procedure, we will thus be led to control the past of the process through the quantity

$$|||D|||_{1;\gamma} := \sup_{t\in(0,1]} t^{1-\gamma}\big|D'(t)\big| \tag{4.2}$$

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

NOTATION 4.1. For every $k \geq 1$ and every $\gamma \in (0,1)$, we denote by $\mathcal{E}_\gamma^k$ 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

$$|||f|||_{k;\gamma} := \max_{1\leq \ell\leq k} \sup_{t\in(0,1]} t^{\ell-\gamma}\big|f^{(\ell)}(t)\big| < \infty. \tag{4.3}$$

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 $H \in (1/3, 1/2)$. From now on and for the rest of the paper, we fix a parameter $\gamma \in (1/3, H)$ that will serve us throughout the reasoning. Then, denote by $(Y_t, \widetilde{Y}_t)$, a couple of solutions to (1.3):

$$\begin{cases} dY_t = b(Y_t)\, dt + \sigma(Y_t)\, d\mathbf{X}_t, \\ d\widetilde{Y}_t = b(\widetilde{Y}_t)\, dt + \sigma(\widetilde{Y}_t)\, d\widetilde{\mathbf{X}}_t \end{cases} \tag{4.4}$$

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

$$(X_t)_{t\le 0} = (\widetilde{X}_t)_{t\le 0} \tag{4.5}$$

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

$$\widetilde{\mu}(dx, d\widetilde{x}, dw) = \nu_0(w, dx)\nu(w, d\widetilde{x})\mathbb{P}_H(dw), \tag{4.6}$$

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

$$\mu_0(d\widetilde{x}, dw) := \nu_0(w, d\widetilde{x})\mathbb{P}_H(dw) \quad \text{and} \quad \mu(dx, dw) = \nu(w, d\widetilde{x})\mathbb{P}_H(dw). \tag{4.7}$$

In other words, $\widetilde{Y}$ is a stationary solution whereas Y starts with a given initial condition μ_0. At this point, let us remember that thanks to Corollary 3.6, we can choose μ in such a way that for every $r > 0$,

$$\int |x|^r \bar{\mu}(dx) < +\infty,$$

where, as usual, $\bar{\mu}$ stands for the first marginal of μ. In fact, for the rest of the paper and along the assumptions of Theorem 2.5, we fix $r > 0$ such that one has simultaneously

$$\int |x|^r \bar{\mu}_0(dx) < +\infty \quad \text{and} \quad \int |x|^r \bar{\mu}(dx) < +\infty. \tag{4.8}$$

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 $(\mathcal{F}_t)_{t\ge 0}$ is also generated by $(\sigma(W_s, \widetilde{W}_s, (Y_0, \widetilde{Y}_0))_{s\le t})_{t\ge 0}$.

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):

$$d\widetilde{W}_t = dW_t + g_W(t)\,dt \quad \text{and} \quad d\widetilde{X}_t = dX_t + g_X(t)\,dt. \tag{4.9}$$

In the continuity of (4.5), we assume that for every $t \in \mathbb{R}_-$,

$$W_t = \widetilde{W}_t \qquad \text{a.s.} \quad \text{or equivalently that} \quad g_W(t) = g_X(t) = 0 \qquad \text{a.e.}$$

Note that the functions g_W and g_X are linked by the following formulas, whenever the latter make sense (see [18], Lemma 4.2 for details):

$$g_W(t) = \alpha_H \frac{d}{dt}\int_{-\infty}^{t} (t-s)^{\frac{1}{2}-H} g_X(s)\,ds, \tag{4.10}$$

$$g_X(t) = \gamma_H \alpha_{1-H} \frac{d}{dt}\int_{-\infty}^{t} (t-s)^{H-\frac{1}{2}} g_W(s)\,ds. \tag{4.11}$$

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

$$\tau_\infty := \inf\{t \geq 0, Y_s = \widetilde{Y}_s \text{ for all } s \geq t\}. \tag{4.12}$$

Before going further, let us recall a classical relation between τ_∞ and the total variation distance. Denote by $\mathcal{B}$ the set of Borel functions $F : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}$. Then, owing to the stationarity of $\widetilde{Y}$, we have for any $t \geq 0$,

$$\begin{aligned} &\big\|\mathcal{L}(Y_{t+\cdot}) - \bar{\mathcal{Q}}\mu\big\|_{\mathrm{TV}} \\ &\quad= \sup_{F\in\mathcal{B}, \|F\|_\infty \leq 1} \mathbb{E}\big[\big(F(Y_{t+\cdot}) - F(\widetilde{Y}_{t+\cdot})\big)\big] \\ &\quad= \sup_{F\in\mathcal{B}, \|F\|_\infty \leq 1} \mathbb{E}\big[\big(F(Y_{t+\cdot}) - F(\widetilde{Y}_{t+\cdot})\big)1_{\{\tau_\infty > t\}}\big] \leq 2\mathbb{P}(\tau_\infty > t). \end{aligned} \tag{4.13}$$

As a consequence, in the sequel, we will focus on the quantity $t \mapsto \mathbb{P}(\tau_\infty > t)$ in order to prove the main theorem.

As usual, the construction of the coupling is achieved through 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 $\tau_0 \geq 0$ the beginning of the first trial and by τ_k, $k \geq 1$, the end of each trial. This also means that τ_{k-1} designates the beginning of the kth 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.

$\triangleright$ 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, one begins Step 3 (see below) at time $\tau_{k-1} + 1$. Otherwise, one begins Step 2.

$\triangleright$ Step 2 is in fact a series of trials on some intervals $I_{k,\ell}$ with length

$$|I_{k,\ell}| = c_2 2^\ell, \tag{4.14}$$

independently of k, 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,

$$g_X(t)=0 \qquad \text{a.s.}$$

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, that is, 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 [18] that we choose to recall here in a slightly different way.

LEMMA 4.2. *Let τ, t_0 be fixed positive numbers, and assume that $g_W(t)=0$ on $(-\infty,0]$. Then $g_X=0$ on $(\tau,\tau+t_0]$ if and only if for all $t\in(0,t_0]$,*

$$g_W^{\tau}(t)=\big(\mathcal{R}_0 g_W^{\tau}\big)(t), \tag{4.15}$$

where $g_W^{\tau}(\cdot)=g_W(\tau+\cdot)$ and where the operator $\mathcal{R}_0$ is defined as

$$(\mathcal{R}_0 g)(t)=C_H\int_{-\infty}^{0}\frac{t^{\frac{1}{2}-H}(-s)^{H-\frac{1}{2}}}{t+T-s}g(s)\,ds, \qquad t\in(0,+\infty) \tag{4.16}$$

for some appropriate constant C_H (depending only on H).

REMARK 4.3. Since $g_W=0$ on $(-\infty,0]$, observe that $(\mathcal{R}_0 g_W^{\tau}(t))_{t\geq 0}$ only involves the values of g_W on $[0,\tau]$ and, therefore, relation (4.15) provides an explicit description of the values of g_W on $(\tau,\tau+t_0]$ in terms of $(g_W(t))_{0\leq t\leq\tau}$, as expected.

REMARK 4.4. For our further computations, we will also use the following generalization of the operator $\mathcal{R}_0$, defined for a parameter $T\geq 0$:

$$(\mathcal{R}_T g)(t)=C_H\int_{-\infty}^{0}\frac{t^{\frac{1}{2}-H}(T-s)^{H-\frac{1}{2}}}{t+T-s}g(s)\,ds, \qquad t\in(0,+\infty). \tag{4.17}$$

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

In the continuity of [18] and [11], we thus introduce an admissibility condition [we recall that we have fixed a parameter $\gamma\in(1/3,H)$ for the whole study].

DEFINITION 4.5. Let K and α be some positive constants and fix a time4 $\tau\geq 0$. Then we will say that a state $\pi:=(a,\tilde{a},w,\tilde{w})\in(\mathbb{R}^d)^2\times\mathcal{C}((-\infty,\tau];\mathbb{R}^d)^2$ is (K,α,γ)-admissible at time τ and if the following conditions are satisfied:

(i) It holds that $\tilde{w}=w+\int_{-\infty}^{\cdot}g_W(s)\,ds$, with g_W satisfying:

$$\sup_{T\geq 0}\int_0^{+\infty}(1+t)^{2\alpha}\big|\big(\mathcal{R}_T|g_W^{\tau}|\big)(t)\big|^2\,dt\leq 1. \tag{4.18}$$

(ii) It holds that

$$|a| + |\tilde{a}| + \|\!| D^{(\tau)}(w) |\!\|_{1;\gamma} + \|\!| D^{(\tau)}(\tilde{w}) |\!\|_{1;\gamma} \leq K, \tag{4.19}$$

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

$$D^{(\tau)}(w)_t := \int_{-\infty}^{\tau} \big((t+\tau-r)_+^{H-\frac{1}{2}} - (-r)_+^{H-\frac{1}{2}} \big)\, dw_r. \tag{4.20}$$

REMARK 4.6. We are aware that, following the subsequent Lemma 6.5, the above transformation $D^{(\tau)}$ is only defined on a subspace

$$\Omega_-^\tau \subset \mathcal{C}\big((-\infty, \tau]; \mathbb{R}^d\big)$$

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

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

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

Regarding Lemma 4.2, one can remark that Condition (4.18) 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.

Item (ii) in Definition 4.5 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

$$A_k(K, \alpha, \gamma) := \{\tau_{k-1} < \infty\} \cap \big\{ (K, \alpha, \gamma)\text{-admissibility at time } \tau_{k-1} \big\}. \tag{4.21}$$

$\triangleright$ If the coupling attempt fails, one begins Step 3, that is, one waits sufficiently before another attempt. This waiting time will be chosen exponentially proportional to the length of the failing attempt. More precisely, let $\ell \geq 1$ 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, for $k \geq 1$ and $\ell \geq 0$,

$$F_{k,\ell} := \{\tau_{k-1} < \infty\} \cap \{\text{at trial } k, \text{ Step 2 fails exactly at the } \ell\text{th attempt}\} \tag{4.22}$$

and

$$\begin{aligned} \mathcal{B}_{k,\ell} := {} & \{\tau_{k-1} < \infty\} \\ & \cap \big\{\text{at trial } k, \text{ Step 2 succeeds} \\ & \qquad \text{at least up to the } \ell\text{th attempt (included)}\big\}. \end{aligned} \tag{4.23}$$

Note that with our convention, $\mathcal{B}_{k,0}$ thus corresponds to the event where Step 1 succeeds. Let us finally label the following family of sets, defined inductively, for further use:

$$\mathcal{E}_k := (\tau_k < \infty) = (\tau_{k-1} < \infty) \cap \Big(\bigcup_{\ell \geq 0} F_{k,\ell} \Big). \tag{4.24}$$

In fact, we will assume in the sequel (see Section 8) that for every $\omega \in F_{k,\ell}$,

$$\tau_k - \tau_{k-1}^3 = \Delta_3(\ell, k) \qquad \text{with } \Delta_3(\ell, k) := c_3 \varsigma^k 2^{\beta \ell}, \tag{4.25}$$

where τ_{k-1}^3 denotes the beginning of Step 3 (in the kth attempt), and c_3, ς and β are deterministic parameters that will be calibrated in the course of the reasoning. In particular, conditionally to $F_{k,\ell}$, the length of each step will be assumed to be deterministic. During this waiting time, we simply set

$$g_W(t) = 0 \qquad \text{on } \big[\tau_{k-1}^3, \tau_k\big], \quad \text{that is,} \quad \widetilde{W}_t - \widetilde{W}_{\tau_{k-1}^3} = W_t - W_{\tau_{k-1}^3}. \tag{4.26}$$

4.3. *Proof of Theorem 2.5: Uniqueness and rate of convergence.* Denoting by $(Y, \widetilde{Y})$ the coupling of solutions derived from the above described construction, and with notation (4.12) in mind, our aim is to prove the following assertion: for every $\alpha \in (0, H)$ and every $p \in (0, \alpha(1 - 2\alpha))$, there exists a constant $C_{\alpha,p} > 0$ such that, for some appropriate calibration of the 3-step scheme, one has

$$\mathbb{P}(\tau_\infty > t) \leq C_{\alpha,p} t^{-p}. \tag{4.27}$$

Optimizing the latter bound with respect to p and then using (4.13), the uniqueness assertion of Theorem 2.5 follows immediately, as well as the convergence rate (2.8), thus completing the proof of our main result.

The strategy toward (4.27) is based on a combination of the successive controls that will be obtained at each stage of our 3-step scheme, namely the results of Proposition 6.9, Proposition 7.1 and Proposition 8.1. Although the controls in question will only be shown in the next sections, we have found it important to anticipate these results so as to provide the proof of (4.27) right now, which will allow us to both give the reader a general overview of our arguments and also to motivate the forthcoming technical considerations.

Let us fix $\alpha \in (0, H)$, $p \in (0, \alpha(1-2\alpha))$ and $\beta \in ((1-2\alpha)^{-1}, \alpha/p)$. In order to suitably calibrate the scheme, we first consider the parameter $K > 0$ given by Proposition 8.1 for the particular choice $\varepsilon = 1/2$. In other words, with the notation of Proposition 8.1, we set

$$K := \mathbf{K}(1/2, \alpha). \tag{4.28}$$

Then we denote by $\delta = \boldsymbol{\delta}(K, \alpha) > 0$ the constant provided by Proposition 6.9, and we fix a constant $\varsigma = \varsigma(\delta, p, \alpha) > 1$ such that $\varsigma^p < (1 - \frac{\mathbf{c}_{\alpha,K} \cdot \delta}{2})^{-1}$, where $\mathbf{c}_{\alpha,K} := (1 - \rho^2_{\alpha,K}) \prod_{\ell=1}^{\infty} (1 - 2^{-\alpha\ell})$, $\rho^2_{\alpha,K} \in (0,1)$ being here the constant given by Proposition 7.1. Finally, with the notation of Proposition 8.1, we define c_3 by

$$c_3 := \mathbf{c_3}(1/2, \alpha, \beta, \varsigma). \tag{4.29}$$

With these parameters in hand, we will assume in the sequel that Step 2 and Step 3 of the above-described coupling scheme are respectively calibrated along the formulas

$$c_2 := \big(\mathbf{C}^{\mathbf{2}}_{\alpha,K}\big)^{\frac{1}{2\alpha}} \quad \text{and} \quad \Delta_3(k,\ell) := c_3 \varsigma^k 2^{\beta\ell}, \tag{4.30}$$

where $\mathbf{C}^{\mathbf{2}}_{\alpha,K} \geq 1$ stands for the constant provided by Proposition 7.1. We can now turn to the reasoning toward (4.27). In what follows, we denote by C_p, resp. $C_{\alpha,p,\beta}$, any generic constant that depends only on p, respectively, (α, p, β).

Set $k^* := \inf\{k \geq 1, \Delta\tau_k = +\infty\}$, where $\Delta\tau_k := \tau_k - \tau_{k-1}$. Using that $\tau_0 = 0$ and $\tau_\infty \leq \tau_{k^*-1} + 1$, we have

$$\mathbb{P}(\tau_\infty > t) \leq \mathbb{P}\bigg(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t - 1\bigg). \tag{4.31}$$

By Markov's inequality and the fact that $|u+v|^p \leq |u|^p + |v|^p$ (because $p \in (0,1]$), we then deduce, provided $t > 1$,

$$\mathbb{P}\bigg(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t - 1\bigg) \leq \frac{C_p}{t^p} \sum_{k=1}^{+\infty} \mathbb{E}\big[|\Delta\tau_k|^p 1_{\{k^*>k\}}\big],$$

and as $(k^* > k) = \bigcup_{\ell \geq 0} F_{k,\ell}$, this yields

$$\mathbb{P}\bigg(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t - 1\bigg) \leq \frac{C_p}{t^p} \sum_{k=1}^{+\infty} \sum_{\ell=0}^{\infty} \mathbb{E}\big[|\Delta\tau_k|^p 1_{F_{k,\ell}}\big]. \tag{4.32}$$

On $F_{k,0}$ (that is the case where Step 1 fails), we have set $\Delta\tau_k = 1 + c_3\varsigma^k$ according to (4.25). For $\ell \geq 1$, we have by definition, on the event $F_{k,\ell}$,

$$\begin{aligned}
\Delta\tau_k &= 1 + c_2(1 + \cdots + 2^\ell) + c_3\varsigma^k 2^{\beta\ell} \\
&\leq c_2 + c_2(1 + \cdots + 2^\ell) + c_3\varsigma^k 2^{\beta\ell} \\
&\leq \frac{c_3}{2} + \frac{c_3}{2}(2^{\ell+1} - 1) + c_3\varsigma^k 2^{\beta\ell} \leq c_3\varsigma^k 2^{\beta\ell+1},
\end{aligned}$$

where we have also used the fact that, according to Proposition 8.1, one has $c_3 \geq 2c_2$. Thus we can recast relation (4.32) as

$$\mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t-1\right) \leq \frac{C_{\alpha,p,\beta}}{t^p} \sum_{k=1}^{+\infty} \varsigma^{kp} \sum_{\ell=0}^{\infty} 2^{\beta\ell p} \mathbb{P}(F_{k,\ell}). \tag{4.33}$$

In addition, owing to our definitions (4.22) and (4.23), it is readily checked that $F_{k,\ell} \subset \mathcal{B}_{k,\ell-1} \cap \mathcal{B}_{k,\ell}^c$ for every $\ell \geq 1$, and so by Proposition 7.1 [remember that we are working under the calibration (4.30) for c_2] we have for every $\ell \geq 1$,

$$\mathbb{P}(F_{k,\ell}|\mathcal{E}_{k-1}) \leq \mathbb{P}\big(\mathcal{B}_{k,\ell}^c|\mathcal{B}_{k,\ell-1}\big) \leq 2^{-\alpha\ell-1}. \tag{4.34}$$

Plugging this inequality into (4.33), we end up with

$$\begin{aligned} \mathbb{P}\left(\sum_{k=1}^{+\infty} \Delta\tau_k 1_{k^*>k} > t-1\right) &\leq \frac{C_{\alpha,p,\beta}}{t^p} \sum_{k=1}^{+\infty} \varsigma^{kp} \left(\sum_{\ell=0}^{+\infty} 2^{(\beta p-\alpha)\ell}\right) \mathbb{P}(\mathcal{E}_{k-1}) \\ &\leq \frac{C_{\alpha,p,\beta}}{t^p} \sum_{k=1}^{+\infty} \varsigma^{kp} \mathbb{P}(\mathcal{E}_{k-1}), \end{aligned} \tag{4.35}$$

where we have used the fact that $\beta p - \alpha < 0$. Now

$$\mathbb{P}(\mathcal{E}_{k-1}) = \prod_{m=1}^{k-1} \mathbb{P}(\mathcal{E}_m|\mathcal{E}_{m-1}) = \prod_{m=1}^{k-1} \big(1 - \mathbb{P}\big(\mathcal{E}_m^c|\mathcal{E}_{m-1}\big)\big), \tag{4.36}$$

and observe that for every $m \geq 1$,

$$\mathbb{P}\big(\mathcal{E}_m^c|\mathcal{E}_{m-1}\big) \geq \mathbb{P}\big(\Delta\tau_m = +\infty|\mathcal{E}_{m-1} \cap A_m(K,\alpha,\gamma)\big)\mathbb{P}\big(A_m(K,\alpha,\gamma)|\mathcal{E}_{m-1}\big). \tag{4.37}$$

At this point, we can first apply Proposition 8.1 [remember that K and c_3 are defined by (4.28) and (4.29)] to derive that

$$\mathbb{P}\big(A_m(K,\alpha,\gamma)|\mathcal{E}_{m-1}\big) \geq \frac{1}{2}. \tag{4.38}$$

On the other hand, using the decomposition

$$\begin{aligned} &\mathbb{P}\big(\Delta\tau_m = +\infty|\mathcal{E}_{m-1} \cap A_m(K,\alpha,\gamma)\big) \\ &\quad = \mathbb{P}\big(Y_{\tau_{m-1}+1} = \widetilde{Y}_{\tau_{m-1}+1}|\mathcal{E}_{m-1} \cap A_m(K,\alpha,\gamma)\big) \prod_{\ell=1}^{+\infty} \mathbb{P}(\mathcal{B}_{m,\ell}|\mathcal{B}_{m,\ell-1}), \end{aligned} \tag{4.39}$$

we can easily combine the results of Proposition 6.9 and Proposition 7.1 to assert that

$$\mathbb{P}\big(\Delta\tau_m = +\infty|\mathcal{E}_{m-1} \cap A_m(K,\alpha,\gamma)\big) \geq \mathbf{c}_{\alpha,K}\delta, \tag{4.40}$$

where $\mathbf{c}_{\alpha,K} > 0$ and $\delta > 0$ have been introduced at the beginning of the proof.

Injecting (4.38)–(4.40) into (4.36) and going back to (4.37), we have thus shown that

$$\sum_{k=1}^{+\infty} \varsigma^{kp} \mathbb{P}(\tau_{k-1} < +\infty) \leq \sum_{k=1}^{+\infty} \varsigma^{pk} \left(1 - \frac{\mathbf{c}_{\alpha,K}\,\delta}{2}\right)^{k-1}, \tag{4.41}$$

and the latter quantity is known to be finite due to our choice of ς [i.e., ς such that such that $\varsigma^p < (1 - (\mathbf{c}_{\alpha,K}\,\delta)/2)^{-1}$]. The expected bound (4.27) then follows from the combination of (4.31), (4.35) and (4.41), and this achieves the proof of Theorem 2.5.

The remainder of the paper is now devoted to the proof of the intermediate results at the core of the above arguments, that is, the results of Proposition 6.9, Proposition 7.1 and Proposition 8.1, which actually correspond to controls at Step 1, Step 2 and Step 3, respectively. As we mentioned in the Introduction, the main difficulties of this procedure (more precisely, the most innovative part with respect to the analysis in [11]) essentially lie in the *hitting* Step 1. We thus propose in the next subsection to give a heuristic description of the coupling construction during this first stage, before we go into the technical details of Sections 5 and 6.

4.4. *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 Section 4.2, 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 $(y_t^0, y_t^1, g_t)_{t\in[0,1]}$ satisfying the system,

$$\begin{cases} dy_t^0 = b(y_t^0)\,dt + \sigma(y_t^0)\,dx_t, \\ dy_t^1 = b(y_t^1)\,dt + \sigma(y_t^1)(dx_t + g_t\,dt), \end{cases}$$

as well as the constraints $y_0^0 = a_0$, $y_0^1 = a_1$ and $y_1^0 = y_1^1$. Using our invertibility assumptions on σ [i.e., Hypothesis (H3)], this amounts to finding $(y_t^0, y_t^1, h_t)_{t\in[0,1]}$ such that

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

and $y_0^0 = a_0$, $y_0^1 = a_1$, $y_1^0 = y_1^1$. In fact, let us consider the slightly more general issue of exhibiting a family of paths $(y_t^\xi, h_t^\xi)_{t\in[0,1],\xi\in[0,1]}$ that satisfy the equation

$$dy_t^\xi = b(y_t^\xi)\,dt + \sigma(y_t^\xi)\,dx_t + h_t^\xi\,dt, \tag{4.42}$$

as well as the constraints $y_0^\xi = a_0 + \xi(a_1 - a_0)$, $y_1^0 = y_1^1$ and $h_\cdot^0 \equiv 0$. Then, using the basic identity $y_1^0 - y_1^1 = \int_0^1 d\xi\, \partial_\xi y_1^\xi$, we are led to the following sufficient formulation of the problem: finding a family $(y_t^\xi, h_t^\xi)_{t\in[0,1],\xi\in[0,1]}$ that satisfies both (4.42) and the constraints $y_0^\xi = a_0 + \xi(a_1 - a_0)$, $\partial_\xi y_1^\cdot \equiv 0$, $h_\cdot^0 \equiv 0$.

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 $h_t^\xi := -\int_0^\xi d\eta\, J_t^\eta$, where for each ξ, J^ξ stands for the solution of the equation

$$dJ_t^\xi = b'(y_t^\xi) J_t^\xi\, dt + \sigma'(y_t^\xi) J_t^\xi\, dx_t, \qquad J_0^\xi = a_1 - a_0.$$

With this specific choice of h_t^ξ, it is readily checked that the two paths $t \mapsto \partial_\xi y_t^\xi$ and $t \mapsto J_t^\xi(1-t)$ satisfy the very same equation

$$dz_t^\xi = \bigl[b'(y_t^\xi) z_t^\xi - J_t^\xi\bigr]dt + \sigma'(y_t^\xi) z_t^\xi\, dx_t.$$

Accordingly, $\partial_\xi y_t^\xi = J_t^\xi(1-t)$ and the above constraints $\partial_\xi y_1^\cdot \equiv 0$, $h_\cdot^0 \equiv 0$ are indeed satisfied.

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

$$\begin{cases} dy_t^\xi = \left[b(y_t^\xi) - \displaystyle\int_0^\xi d\eta\, J_t^\eta\right] dt + \sigma(y_t^\xi)\, dx_t, \\ dJ_t^\xi = b'(y_t^\xi) J_t^\xi\, dt + \sigma'(y_t^\xi) J_t^\xi\, dx_t, \end{cases} \tag{4.43}$$

with initial conditions $y_0^\xi = a_0 + \xi(a_1 - a_0)$, $J_0^\xi = a_1 - a_0$ 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_t^ξ, J_t^ξ) understood as a rough solution of (4.43), in the sense of Definition 2.3. Just as above, once endowed with such a solution (y, J), defining g_X as $g_X(t) := -\sigma(y_t^1)^{-1} \int_0^1 d\eta\, J_t^\eta$ would then close the procedure.

Unfortunately, as the reader can easily check it, the vector fields involved in (4.43) 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 [7], Theorem 6.1, or in [14], 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 introduce 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, because the required smallness condition on $\|\mathbf{X}\|_{\gamma;[0,1]}$ implicitly involves the past trajectory $(W_t)_{t\le 0}$, 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 (4.1) of X, and rewrite the *hitting* system as

$$\begin{cases} dY_t^\xi = \left[b(Y_t^\xi)\, dt - \int_0^\xi d\eta J_t^\eta\, dt + \sigma(Y_t^\xi)\, dD_t \right] + \sigma(Y_t^\xi)\, dZ_t, \\ dJ_t^\xi = [b'(Y_t^\xi) J_t^\xi\, dt + \sigma'(Y_t^\xi) J_t^\xi\, dD_t] + \sigma'(Y_t^\xi) J_t^\xi\, dZ_t, \end{cases} \tag{4.44}$$

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 (4.44) 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 (4.44) on $[0,1]$ provided the norm of $\|\mathbf{Z}\|_{\gamma;[0,1]}$ [which no longer depends on $(W_t)_{t\le 0}$] 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 (4.44) (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 (4.44).

The effective application of these general (pathwise) considerations to the particular fractional system (4.44) 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).

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 $\alpha \in (0,1]$, $\mu \geq \alpha$, set, for any map $f : I^2 \to V$, resp. $f : I^3 \to V$,

$$\mathcal{N}\big[f; \mathcal{C}^{\alpha,\mu}_{2;\beta}(I;V)\big] := \max\Big(\sup_{s<t\in I} \frac{\|f_{st}\|_V}{|t-s|^\alpha}, \sup_{0<s<t\in I} \frac{\|f_{st}\|_V}{|t-s|^\mu s^{\beta-1}} \Big), \tag{5.1}$$

respectively,

$$\mathcal{N}\big[f; \mathcal{C}^{\alpha,\mu}_{3;\beta}(I;V)\big] := \max\Big(\sup_{s<u<t\in I} \frac{\|f_{sut}\|_V}{|t-s|^\alpha}, \sup_{0<s<u<t\in I} \frac{\|f_{sut}\|_V}{|t-s|^\mu s^{\beta-1}} \Big), \tag{5.2}$$

and then

$$\mathcal{C}^{\alpha,\mu}_{1;\beta}(I;V) := \big\{ f \in \mathcal{C}_1(I;V) : \delta f \in \mathcal{C}^{\alpha,\mu}_{2;\beta}(I;V) \big\}. \tag{5.3}$$

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 (relatively sharp) inclusion.

LEMMA 5.1. *Let $\mathcal{E}^1_\gamma([0,1];\mathbb{R}^d)$ be the space introduced in Notation 4.1. 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 $g \in \mathcal{E}^1_\gamma([0,1];\mathbb{R}^d)$,*

$$\mathcal{N}\big[\delta g; \mathcal{C}^{\gamma,1}_{2;\gamma}([0,1];\mathbb{R}^d)\big] \leq c_\gamma |\!|\!| g |\!|\!|_{1;\gamma}. \tag{5.4}$$

PROOF. It suffices to observe that for all $0 < s < t \leq 1$,

$$\begin{aligned} |\delta g_{st}| = \Big| \int_s^t g'_u \, du \Big| &\leq |\!|\!| g |\!|\!|_{1;\gamma} \int_s^t u^{\gamma-1} \, du \\ &\lesssim |\!|\!| g |\!|\!|_{1;\gamma} \min\big(s^{\gamma-1}|t-s|, \{t^\gamma - s^\gamma\} \big) \\ &\lesssim |\!|\!| g |\!|\!|_{1;\gamma} \min\big(s^{\gamma-1}|t-s|, |t-s|^\gamma \big). \end{aligned}$$

□

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 ℓth 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 $\mathbf{z} = (z, \mathbf{z^2})$, 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 any continuous, respectively, differentiable, vector field

$$B : V \to \mathcal{L}\big(\mathbb{R}^m; V\big), \qquad \text{resp. } \Sigma : V \to \mathcal{L}\big(\mathbb{R}^n; V\big),$$

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

$$dy_t = B(y_t)\, dh_t + \Sigma(y_t)\, d\mathbf{z}_t, \qquad y_{t_0} = v_0, \tag{5.5}$$

if the two-parameter path R^y defined as

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

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

$$(D\Sigma_j \cdot \Sigma_k)(v) := (D\Sigma_j)(v)\big(\Sigma_k(v)\big) \qquad \text{for every } v \in V.$$

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 [7, 8, 14], 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 [7], Definition 3.1). This is not the case in the above formulation, where the solution is still expected to belong to $\mathcal{C}^\gamma_1$. Another (equivalent) possibility would here have consisted in keeping track of the specific controls (in the sense of Lyons' theory) arising throughout the procedure.

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 (5.5). 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 B, respectively, Σ, is $\mathcal{C}^2$, respectively, $\mathcal{C}^3$, and*

$$\sup_{v\in V} \big\|D^{(\ell)}B(v)\big\| < \infty \qquad \textit{for } \ell \in \{1,2\},$$

$$\textit{resp. } \sup_{v\in V} \big\|D^{(\ell)}\Sigma(v)\big\| < \infty \qquad \textit{for } \ell \in \{0,\dots,3\}.$$

HYPOTHESIS (VF2). *The vector field B, respectively, Σ, is $\mathcal{C}^2$, respectively, $\mathcal{C}^3$, and the following bounds hold true: for all $\ell \in \{0,1,2\}$, $m \in \{0,\dots,3\}$,*

$$\big\|\big(D^{(\ell)}B\big)(v)\big\| \lesssim 1 + \|v\|, \qquad \big\|\big(D^{(m)}\Sigma\big)(v)\big\| \lesssim 1 + \|v\|, \tag{5.6}$$

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

$$\begin{aligned} \big\|(D\Sigma\cdot\Sigma)(v)\big\| &\lesssim 1 + \|v\|, \\ \big\|(D\Sigma\cdot\Sigma)(v) - (D\Sigma\cdot\Sigma)(w)\big\| &\lesssim \|v-w\|\big\{1+\|v\|\big\}. \end{aligned} \tag{5.7}$$

THEOREM 5.4 [(VF1)-situation]. *Under Hypothesis (VF1), and for any $v_0 \in V$, equation (5.5) 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 (5.5) admits at most one solution on $[0,1]$ with initial condition v_0, in the sense of Definition 5.2.*

(ii) *For every $K \geq 1$, there exists $M_K > 0$ such that if*

$$\|v_0\| \leq K, \qquad \mathcal{N}\big[\delta h; \mathcal{C}^{\gamma,1}_{2;\beta}([0,1])\big] \leq K \quad \text{and} \quad \|\mathbf{z}\|_{\gamma;[0,1]} \leq M_K,$$

then equation (5.5) admits a unique solution y on $[0,1]$ with initial condition v_0, in the sense of Definition 5.2. Besides,

$$\mathcal{N}\big[y; \mathcal{C}^0_1([0,1];V)\big] + \mathcal{N}\big[y; \mathcal{C}^\gamma_1([0,1];V)\big] \leq C(K) \tag{5.8}$$

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

For the sake of conciseness, we have provided the details of the proof of these technical results in the companion paper [9], Section 2.

6. Hitting step. Keeping in mind the strategy sketched out in Section 4.4, 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 (4.44) 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.

We recall that we have fixed $H \in (1/3, 1/2)$, $\gamma \in (1/3, H)$, 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 recall that the space $\mathcal{E}^2_\gamma([0,1];\mathbb{R}^d)$ has been introduced through Notation 4.1, and let us fix two paths $h \in \mathcal{E}^2_\gamma([0,1];\mathbb{R}^d)$, $z \in \mathcal{C}^\gamma([0,1];\mathbb{R}^d)$, assuming in addition 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:

$$B^i_0 \begin{pmatrix} y \\ J \end{pmatrix}(\xi) := \begin{pmatrix} b^i(y(\xi)) - \displaystyle\int_0^\xi d\eta J^i(\eta) \\ (\partial_k b^i)(y(\xi)) J^k(\xi) \end{pmatrix},$$

$$B^i_j \begin{pmatrix} y \\ J \end{pmatrix}(\xi) = \Sigma^i_j \begin{pmatrix} y \\ J \end{pmatrix}(\xi) := \begin{pmatrix} \sigma^i_j(y(\xi)) \\ (\partial_k \sigma^i_j)(y(\xi)) J^k(\xi) \end{pmatrix}$$

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

PROOF. We have the following explicit expressions:

$$(D\Sigma_j^i)\begin{pmatrix} y \\ J \end{pmatrix}\begin{pmatrix} y_1 \\ J_1 \end{pmatrix} = \begin{pmatrix} (\partial_k \sigma_i^j)(y) y_1^k \\ (\partial_k \partial_\ell \sigma_j^i)(y) y_1^\ell J^k + (\partial_k \sigma_i^j)(y) J_1^k \end{pmatrix},$$

$$(D^{(2)}\Sigma_j^i)\begin{pmatrix} y \\ J \end{pmatrix}\left(\begin{pmatrix} y_1 \\ J_1 \end{pmatrix}, \begin{pmatrix} y_2 \\ J_2 \end{pmatrix}\right)$$

$$= \begin{pmatrix} (\partial_{k_1} \partial_{k_2} \sigma_j^i)(y) y_1^{k_1} y_2^{k_2} \\ (\partial_k \partial_{k_1} \partial_{k_2} \sigma_j^i)(y) y_1^{k_1} y_q^{k_2} J^k + (\partial_{k_1} \partial_{k_2} \sigma_j^i)(y)\{J_1^{k_1} y_2^{k_2} + y_1^{k_1} J_2^{k_2}\} \end{pmatrix},$$

$$(D^{(3)}\Sigma_j^i)\begin{pmatrix} y \\ J \end{pmatrix}\left(\begin{pmatrix} y_1 \\ J_1 \end{pmatrix}, \begin{pmatrix} y_2 \\ J_2 \end{pmatrix}, \begin{pmatrix} y_3 \\ J_3 \end{pmatrix}\right)$$

$$= \begin{pmatrix} (\partial_{k_1} \partial_{k_2} \partial_{k_3} \sigma_j^i)(y) y_1^{k_1} y_2^{k_2} y_3^{k_3} \\ (\partial_k \partial_{k_1} \partial_{k_2} \partial_{k_3} \sigma_j^i)(y) y_1^{k_1} y_2^{k_2} y_3^{k_3} J^k + (\partial_{k_1} \partial_{k_2} \partial_{k_3} \sigma_j^i)(y)\{J_1^{k_1} y_2^{k_2} y_3^{k_3} + y_1^{k_1} J_2^{k_2} y_3^{k_3} + y_1^{k_1} y_2^{k_2} J_3^{k_3}\} \end{pmatrix}.$$

In particular,

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

Based on these formulas, the two conditions (5.6) and (5.7) for Σ are immediate. We can then exhibit a similar expression for $D^{(q)} B_0$, $q \in \{0, 1, 2\}$. □

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*

$$\begin{aligned} dy_t(\xi) &= \left[\varphi\big(b(y_t(\xi))\big) - \int_0^\xi \varphi\big(J_t(\eta)\big)\, d\eta\right] dt \\ &\quad + \sigma\big(y_t(\xi)\big)\, dh_t + \sigma\big(y_t(\xi)\big)\, d\mathbf{z}_t, \end{aligned} \tag{6.1}$$

$$\begin{aligned} dJ_t(\xi) &= (\partial_k b)\big(y_t(\xi)\big)\varphi\big(J_t(\xi)\big)_k\, dt + (\partial_k \sigma)\big(y_t(\xi)\big)\varphi\big(J_t(\xi)\big)_k\, dh_t \\ &\quad + (\partial_k \sigma)\big(y_t(\xi)\big)\varphi\big(J_t(\xi)\big)_k\, d\mathbf{z}_t, \end{aligned} \tag{6.2}$$

with initial condition $(y_0, J_0) = A$, admits a unique solution

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

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* $\|A\|_{V_2} \le K$ *and* $|||h|||_{1;\gamma} \le K$, *for some fixed* $K \ge 1$. *Then there exists a constant* $M_K > 0$ *such that if* $\|\mathbf{z}\|_{\gamma;[0,1]} \le M_K$, *the conclusion of point* (a) *is still true for* $\varphi \equiv \mathrm{Id}$ *and* V_1 *replaced with* V_2, *and one has*

$$\mathcal{N}\big[\Psi_{V_2}(A, \mathrm{Id}, (h,\mathbf{z})); \mathcal{C}_1^0([0,1]; V_2)\big] \le C(K) \tag{6.3}$$

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 (6.1)–(6.2) 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 (6.1)–(6.2) 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* (6.1)–(6.2) *with* $\varphi \equiv \mathrm{Id}$ *and initial condition*

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

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*

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\mathbf{x}_t, \qquad y_0 = a_0, \tag{6.4}$$

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*

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\widetilde{\mathbf{x}}_t, \qquad y_0 = a_1, \tag{6.5}$$

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

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

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

PROOF. (a) Let y be a solution of (6.4) in the sense of Definition 2.3. Then clearly it is also a solution in the sense of Definition 5.2, and by the identification result [9], Corollary 3.2, we can conclude that y is a solution of the equation

$$dy_t = \big[b(y_t)\,dt + \sigma(y_t)\,dh_t\big] + \sigma(y_t)\,d\mathbf{z}_t, \qquad y_0 = a_0 \tag{6.6}$$

[i.e., 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 (6.6) as well.

(b) As above, observe that due to the regularity of the path e and thanks to [9], Corollary 3.2, the solution of (6.5) (in the sense of Definition 2.3) is also the solution of

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d(h+e)_t + \sigma(y_t)\,d\mathbf{z}_t, \qquad y_0 = a_1, \tag{6.7}$$

in the sense of Definition 5.2 [i.e., 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 (6.7) as well. This can be easily derived from the fact that

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

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

$$(\partial_\xi y_t)(\xi) = J_t(\xi)(1-t). \tag{6.8}$$

It can indeed be checked that, when seen as paths with values in $V := L^\infty([0,1]; \mathbb{R}^d)^3$, the triplets $(y, J, \partial_\xi y)$ and (y, J, g) [where we have set $g_t(\xi) := J_t(\xi)(1-t)$] are both solution of the system obtained by adding to (6.1)–(6.2) the third equation

$$\begin{aligned}
dg_t &= \big[(\partial_k b)(y_t) g_t^k - J_t\big]dt + (\partial_k \sigma)(y_t) g_t^k\, dh_t + (\partial_k \sigma)(y_t) g_t^k\, d\mathbf{z}_t, \\
g_0(\xi) &= a_1 - a_0.
\end{aligned}$$

The conclusion is now immediate:

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

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:

$$\begin{aligned} d\bar{y}_t(\xi) &= \Big[\varphi\big(b(\bar{y}_t(\xi))\big) - \int_0^\xi \varphi\big(\bar{j}_t(\eta)\big)\,d\eta \\ &\quad + \sigma\big(\bar{y}_t(\xi)\big)\sigma\big(\bar{y}_t(1)\big)^{-1}\int_0^1 \varphi\big(\bar{j}_t(\eta)\big)\,d\eta\Big]dt \\ &\quad + \sigma\big(\bar{y}_t(\xi)\big)\,dh_t + \sigma\big(\bar{y}_t(\xi)\big)\,d\mathbf{z}_t, \end{aligned} \tag{6.9}$$

$$\begin{aligned} d\bar{j}_t(\xi) &= \Big[(\partial_k b)\big(\bar{y}_t(\xi)\big)\varphi\big(\bar{j}_t(\xi)\big)_k \\ &\quad + (\partial_k\sigma)\big(\bar{y}_t(\xi)\big)\varphi\big(\bar{j}_t(\xi)\big)_k\sigma\big(\bar{y}_t(1)\big)^{-1}\int_0^1 \varphi\big(\bar{j}_t(\eta)\big)\,d\eta\Big]dt \\ &\quad + (\partial_k\sigma)\big(\bar{y}_t(\xi)\big)\varphi\big(\bar{j}_t(\xi)\big)_k\,dh_t + (\partial_k\sigma)\big(\bar{y}_t(\xi)\big)\varphi\big(\bar{j}_t(\xi)\big)_k\,d\mathbf{z}_t, \end{aligned} \tag{6.10}$$

with initial condition $(y_0(\xi), j_0(\xi)) = ((1-\xi)a_0 + \xi a_1, a_1 - a_0)$. It is not hard to see that for every such fixed φ, the vector fields involved in (6.9)–(6.10) satisfy the conditions of Theorem 5.4 [i.e., Hypothesis (VF1)] and, therefore, the system admits a unique solution

$$(\bar{y}, \bar{j}) =: \bar{\Psi}_{V_1}\big((a_0,a_1), \varphi, (h,\mathbf{z})\big) \in \mathcal{C}_1^\gamma\big([0,1]; V_1\big). \tag{6.11}$$

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

$$T(h,g)_t := h_t + \int_0^t g_s\,ds. \tag{6.12}$$

THEOREM 6.4. *Fix $K \geq 1$ 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*

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

with initial condition $(y_0(\xi), j_0(\xi)) = ((1-\xi)a_0 + \xi a_1, a_1 - a_0)$, admits a unique solution

$$(y, j) =: \Psi_{V_1}\big((a_0,a_1), \varphi_K, (h,\mathbf{z})\big) \in \mathcal{C}_1^\gamma\big([0,1]; V_1\big), \tag{6.13}$$

in the sense of Definition 5.2.

(ii) *There exists a constant* $M_K > 0$ *such that if* $\max(|a_0|, |a_1|, |a_1 - a_0|) \le K$, $|||h|||_{1;\gamma} \le K$ *and* $\|\mathbf{z}\|_{\gamma;[0,1]} \le M_K$, *then, defining* $(y, \jmath)$ *through* (6.13), *one has*: (ii)(a) *the* $\mathbb{R}^d$*-valued path* $y^{(0)} = y.(0)$ *is the solution of*

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

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

$$dy_t = b(y_t)\,dt + \sigma(y_t)\,d\widetilde{\mathbf{x}}_t, \qquad y_0 = a_1, \tag{6.14}$$

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

$$\begin{aligned} g_t &:= -\sigma\big(y_t(1)\big)^{-1} \int_0^1 d\eta \varphi_K\big(j_t(\eta)\big), \\ (y, j) &:= \Psi_{V_1}\big((a_0, a_1), \varphi_K, (h, \mathbf{z})\big); \end{aligned} \tag{6.15}$$

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

(iii) *With notation* (6.11) *and* (6.13) *in mind, we have the following identities*:

$$(y, j) = \Psi_{V_1}\big((a_0, a_1), \varphi_K, (h, \mathbf{z})\big) = \bar{\Psi}_{V_1}\big((a_0, a_1), \varphi_K, \big(T(h, g), \mathbf{z}\big)\big), \tag{6.16}$$

with g *defined just as in* (6.15), *and*

$$(\bar{y}, \bar{\jmath}) := \bar{\Psi}_{V_1}\big((a_0, a_1), \varphi_K, (h, \mathbf{z})\big) = \Psi_{V_1}\big((a_0, a_1), \varphi_K, \big(T(h, \bar{g}), \mathbf{z}\big)\big), \tag{6.17}$$

with

$$\bar{g}(t) := \sigma\big(\bar{y}_t(1)\big)^{-1} \int_0^1 d\eta \varphi_K\big(\bar{\jmath}_t(\eta)\big). \tag{6.18}$$

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

$$\varphi_K \equiv \mathrm{Id} \qquad \text{on } \big[-C(K), C(K)\big]^d$$

and $\varphi_K(x) = 0$ for every $|x| \ge 2C(K)$. Then (i) follows immediately from Proposition 6.2 point (a). Besides, owing to (6.3), 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

$$\Psi_{V_1}\big((a_0, a_1), \varphi_K, (h, \mathbf{z})\big) = \Psi_{V_2}\big(A, \mathrm{Id}, (h, \mathbf{z})\big),$$

with $A(\xi) := ((1 - \xi)a_0 + \xi a_1, a_1 - a_0)$. Therefore, the three points (ii)(a)–(ii)(b)–(ii)(c) can be readily deduced from Proposition 6.3.

In order to prove (iii), observe first that with the notation in (6.15), one has, at least at a formal level,

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

with a similar transformation for the equation involving J. Given the regularity of g, the latter transformations can actually be justified in a rigorous way, that is in the framework settled through Definition 5.2: one can for instance mimic the arguments of the proof of Proposition 6.3 point (b). Identity (6.16) now follows from the uniqueness of the solution to the system (6.9)–(6.10) (with fixed $\varphi := \varphi_K$). Identity (6.17) can then be derived from a similar transformation of (6.9)–(6.10), 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 (4.1).

We recall that we have fixed $H \in (1/3, 1/2)$ and $\gamma \in (1/3, H)$ for the whole study. 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 $\mathcal{C}((-\infty, 1]; \mathbb{R}^d)$. We also define the following two sets of functions, which will be used in order to define our perturbations on Wiener's space:

$$\begin{aligned} \mathcal{B}_-^c := \big\{ & g : (-\infty, 0] \to \mathbb{R}^d; \\ & g \text{ is bounded measurable with compact support} \big\}, \end{aligned} \tag{6.19}$$

$$\mathcal{B}_+ := \big\{ g : [0, 1] \to \mathbb{R}^d;\ g \text{ is bounded and measurable} \big\}. \tag{6.20}$$

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*

$$\mathcal{D}_X^-\varphi(t) := \alpha_H \int_{-\infty}^0 \big((t-r)^{H-\frac{1}{2}} - (-r)^{H-\frac{1}{2}}\big)\, d\varphi(r) \qquad \text{if } t \in (0,1], \tag{6.21}$$

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 function $g_W^- \in \mathcal{B}_-^c$, 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*

$$\mathcal{D}_X^+\varphi(t) = \alpha_H \int_0^t (t-r)^{H-\frac{1}{2}}\, d\varphi(r), \qquad t \in [0,1]. \tag{6.22}$$

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*

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

(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

$$\mathcal{D}_X^-\varphi(t) := \alpha_H\Big(H - \frac{1}{2}\Big)\int_{-\infty}^0 \big((t-r)^{H-\frac{3}{2}} - (-r)^{H-\frac{3}{2}}\big)\varphi(r)\, dr \quad \text{if } t \in (0,1] \tag{6.23}$$

and $\mathcal{D}_X^-\varphi(0) = 0$, for every test-function φ vanishing at 0. Fixing a parameter $\varepsilon' > 0$, we can now specify the space Ω_- alluded to in our claim above:

$$\Omega_- := \Big\{ w_- : (-\infty,0] \to \mathbb{R}^d;\ w_-(0) = 0,\ \lim_{t\to-\infty} \frac{w_-(t)}{|t|^{\frac{1}{2}+\varepsilon'}} = 0 \text{ and } w_- \text{ is } \Big(\frac{1}{2} - \varepsilon\Big)\text{-Hölder continuous on compact intervals}\Big\}. \tag{6.24}$$

Owing to some classical properties on the Wiener process, this subspace is of Wiener measure 1 for any fixed $\varepsilon, \varepsilon' > 0$, and from (6.23), 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 $k \geq 1$ and every $t \in (0, 1]$, it holds that, for some constant $c_k \geq 0$,

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

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

$$\mathcal{D}_X^+\varphi(t) := \alpha_H\varphi(t)t^{H-\frac{1}{2}} + \alpha_H\left(H-\frac{1}{2}\right)\int_0^t (t-r)^{H-\frac{3}{2}}\big(\varphi(t)-\varphi(r)\big)\,dr \tag{6.25}$$

if $t \in (0, 1]$ 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 [14], Theorem 15.33, which allows us to assert that the subspace

$$\Omega_+ := \big\{w_+ \in \widetilde{\Omega}_+ : \mathcal{D}_X^+ w_+ \text{ can be canonically lifted as a } \gamma\text{-rough path}\big\} \tag{6.26}$$

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 [9], Proposition 3.1, since $\int_0^\cdot g_W^+(s)\,ds$ obviously belongs to $\mathcal{C}^1([0, 1])$.

(iii) By [14], 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 [12], 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) (i.e., we can lift $\mathcal{D}_X^+ w_+$) and [9], Proposition 3.1 (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 (6.23)–(6.25) for $\mathcal{D}_X^-$ and $\mathcal{D}_X^+$, we can apply the 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 [18], page 741.

LEMMA 6.6. *Consider $w_-, \widetilde{w}_- \in \Omega_-$, $w_+ \in \Omega_+$ and $g_W^- \in \mathcal{B}_-^c$ [recall that those spaces are respectively defined by (6.24), (6.26) and (6.19)]. We assume that*

$$\widetilde{w}_- = w_- + \int_{-\infty}^{\cdot} g_W^-(s)\, ds. \tag{6.27}$$

Also consider a generic function $g_X^+ \in \mathcal{B}_+$. Then there exists a map $\mathcal{K} : \mathcal{B}_-^c \times \Omega_+ \times \mathcal{B}_+ \to \Omega_+$ such that

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

Specifically, the map $\mathcal{K}$ is given by the following formula:

$$\mathcal{K}(g_W^-, \omega_+, g_X^+) = w_+ + \int_0^{\cdot} \mathcal{H}(g_W^-, g_X^+)\, ds = T\big(w_+, \mathcal{H}(g_W^-, g_X^+)\big),$$

where we recall that T has been introduced in (6.12), and where the integral transformation $\mathcal{H}$ is defined by

$$\begin{aligned} \mathcal{H}(g_1, g_2)_t &:= C_1 \int_{-\infty}^0 \frac{t^{\frac{1}{2}-H}(-s)^{H-\frac{1}{2}}}{t-s} g_1(s)\, ds \\ &\quad + \alpha_H \frac{d}{dt}\bigg(\int_0^t (t-s)^{\frac{1}{2}-H} g_2(s)\, ds\bigg) \\ &= C_1 \mathcal{R}_0 g_1(t) + C_2 \int_0^t (t-s)^{-\frac{1}{2}-H} g_2(s)\, ds. \end{aligned} \tag{6.28}$$

In (6.28), observe that the notation $\mathcal{R}_T$ has been introduced in (4.17).

The above formula can be interpreted as follows: in the spirit of (4.10), 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 4.2, 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

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

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 $\pi = (a, \tilde{a}, w_-, \tilde{w}_-) \in (\mathbb{R}^d)^2 \times \Omega_-^2$. Also, consider $g_W^- \in \mathcal{B}_-^c$, where $\mathcal{B}_-^c$ is defined by (6.24), such that $\tilde{w}_- = w_- + \int_{-\infty}^{\cdot} g_W^-(s)\,ds$. 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*

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

(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 $w_+ \in \Omega_+$,*

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

(v) *For every fixed $w_+ \in \Omega_+$, consider the function $g_W(w_+, \cdot) : (-\infty, 1] \to \mathbb{R}^d$ defined as $g_W(w_+, t) = g_W^-(t)$ if $t \leq 0$ and $g_W(w_+, t) = g_W^+(w_+, t)$ if $t \in (0,1]$, and denote by $g_X(w_+, \cdot)$ the image of $g_W(w_+, \cdot)$ through the transformation (4.11). Then there exists $C_K > 0$ depending only on K such that for $\mathbb{P}_W$-every $w_+ \in \Omega_+$,*

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

PROOF. Set $h_{w_-} := \mathcal{D}_X^- w_-$ [where we recall that $\mathcal{D}_X^-$ is defined by (6.21)] and write, for every $w_+ \in \Omega_+$, $\mathbf{z}_{w_+} := \mathfrak{L}(\mathcal{D}_X^+ w_+)$ [see Lemma 6.5(ii)].

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

$$\big[g_X(w_+)\big](t) := -\sigma\big(y_t(1)\big)^{-1} \int_0^1 d\eta\, \varphi_K\big(j_t(\eta)\big), \tag{6.29}$$

with $(y, j) := \Psi_{V_1}((a, \widetilde{a}), \varphi_K, (h_{w_-}, \mathbf{z}_{w_+}))$. Then, with the notation of Lemma 6.6, we define Λ on $[0, 1]$ as

$$\begin{aligned}\Lambda(w_+)_t &:= \mathcal{K}\big(g_W^-, w_+, g_X(w_+)\big) = w_+(t) + \int_0^t \mathcal{H}\big(g_W^-, g_X(w_+)\big)_s\, ds \\ &= T\big(w_+, \mathcal{H}\big(g_W^-, g_X(w_+)\big)\big)_t.\end{aligned} \tag{6.30}$$

(ii) By the very definition of $\mathcal{K}$, $\mathcal{H}$ and invoking Lemma 6.6, we have, for every $w_+ \in \Omega_+$,

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

Besides, it is readily checked that

$$\begin{aligned}&\Phi\big(\widetilde{a}, \mathfrak{L}\big(\mathcal{D}_X\big(\widetilde{w}_- \sqcup \Lambda(w_+)\big)\big)\big) \\ &\quad = \Phi\bigg(\widetilde{a}, \mathfrak{L}\bigg(\mathcal{D}_X(w_- \sqcup w_+) + \int_0^{\cdot} \big[g_X(w_+)\big](s)\, ds\bigg)\bigg) = y^{(1)},\end{aligned}$$

where $y^{(1)}$ is defined by (6.14). In addition, by the admissibility condition (4.19), we know that $|||h_{w_-}|||_{1;\gamma} \le K$ and $\max(|a|, |\widetilde{a}|, |\widetilde{a} - a|) \le 2K$. 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

$$\begin{aligned}\mathbb{P}_W\left(w_+ \in \Omega_+ : \Phi\left(a; \mathfrak{L}\big(\mathcal{D}_X(w_- \sqcup w_+)\big)_1\right)\right) &= \Phi\big(\tilde{a}; \mathfrak{L}\big(\mathcal{D}_X\big(\widetilde{w}_- \sqcup \Lambda(w_+)\big)\big)_1\big) \\ &\ge \mathbb{P}_W\big(w_+ \in \Omega_+ : \|\mathbf{z}_{w_+}\|_{\gamma;[0,1]} \le M_K\big).\end{aligned}$$

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

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

$$\bar{g}_X(w_+, t) := \sigma\big(\bar{y}_t(1)\big)^{-1} \int_0^1 d\eta \varphi_K\big(\bar{j}_t(\eta)\big),$$

with $(\bar{y}, \bar{j}) := \bar{\Psi}_{V_1}((a, \widetilde{a}), \varphi_K, (h_{\widetilde{w}_-}, \mathbf{z}_{w_+}))$, where the flow $\bar{\Psi}_{V_1}$ has been introduced in relation (6.11). Then consider the application $\bar{\Lambda} : \Omega_+ \to \Omega_+$ given by

$$\bar{\Lambda}(w_+)_t := \mathcal{K}\big(-g_W^-, w_+, \bar{g}_X(w_+)\big) = T_t\big(w_+, \mathcal{H}\big(-g_W^-, \bar{g}_X(w_+, \cdot)\big)\big). \tag{6.31}$$

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 $\bar{w}_+ := \Lambda(w_+) = T(w_+, g_W^+(w_+, \cdot))$, where, according to (i), $g_W^+(w_+, \cdot) := \mathcal{H}(g_W^-, g_X(w_+, \cdot))$, with $g_X(w_+, \cdot)$ given by (6.29). Then consider the functions $G_W^- = \int_{-\infty}^{\cdot} g_W^-(s)\, ds$ and $G_W^+(w_+, \cdot) := \int_0^{\cdot} g_W^+(w_+, s)\, ds$ (defined respectively on $(-\infty, 0]$ and $[0, 1]$), so that, by construction,

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

Using the identification result [9], Corollary 3.2, we get

$$\begin{aligned}
&\bar{\Psi}_{V_1}\big((a,\widetilde{a}),\varphi_K,(h_{\widetilde{w}_-},\mathbf{z}_{\bar{w}_+})\big)\\
&\quad=\bar{\Psi}_{V_1}\big((a,\widetilde{a}),\varphi_K,\big(h_{\widetilde{w}_-}+\mathcal{D}_X^+\big(G_W^+(w_+,\cdot)\big),\mathbf{z}_{w_+}\big)\big)\\
&\quad=\bar{\Psi}_{V_1}\big((a,\widetilde{a}),\varphi_K,\big(h_{w_-}+\mathcal{D}_X^-G_W^-+\mathcal{D}_X^+\big(G_W^+(w_+,\cdot)\big)\big),\mathbf{z}_{w_+}\big).
\end{aligned}\tag{6.32}$$

At this point, observe that by the inversion formula (4.11), one has for any $t\in(0,1]$,

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

Therefore, $h_{w_-}+\mathcal{D}_1G_W^-+\mathcal{D}_2G_W^+=T(h_{w_-},g_X(w_+,\cdot))$, which, going back to (6.32), gives us

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

We can now apply identity (6.16) to assert that

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

which readily entails that $\bar{g}_X(\bar{w}_+,\cdot)=-g_X(w_+,\cdot)$. The conclusion is now immediate: according to (6.31) we have

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

The fact that $\Lambda\circ\bar{\Lambda}=\mathrm{Id}$ follows from symmetric arguments [by using (6.17) instead of (6.16)].

(iv) Let us recall that according to (6.28) we have

$$\begin{aligned}
g_W^+(w_+,t)&=\mathcal{H}\big(g_W^-,g_X(w_+,\cdot)\big)_t\\
&=C_1(\mathcal{R}_0g_W^-)(t)+C_2\int_0^t(t-s)^{-\frac{1}{2}-H}\big[g_X(w_+)\big](s)\,ds.
\end{aligned}$$

The desired uniform bound on the L^2-norm of $g_W^+(w_+,\cdot)$ then follows from two elementary facts: (a) It holds that $\int_0^1|(\mathcal{R}_0g_W^-)(s)|^2\,ds\le1$ thanks to the admissibility condition (4.18). (b) The function $g_X(w_+)$ defined by (6.29) is bounded. This trivially stands from our assumption (2.7) on σ^{-1} and from the definition of the

cutoff function φ_K. The same arguments can be used for the bound on the L^2-norm of $\bar{g}_W^+(w_+,\cdot)$.

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

REMARK 6.8. We have written our results for a time origin $\tau = 0$ for notational sake. However, notice that the generalizations of Lemma 6.6 and Theorem 6.7 to a shifted time origin τ are straightforward. Let us just describe the transformation $\mathcal{H}$ in this context: consider $w_-, \widetilde{w}_- \in \Theta_\tau\Omega_-$, $w_+ \in \Theta_\tau\Omega_+$ and $g_W^- \in \Theta_\tau\mathcal{B}_-^c$, where Θ_τ denotes the shift of a path by τ. We still assume that relation (6.27) between $\widetilde{w}_-$ and w_- holds true. Then for $g_X^+ \in \Theta_\tau\mathcal{B}_+$ we have

$$\mathcal{D}_X(w_- \sqcup w_+)_{|[\tau,\tau+1]} + \int_\tau^\cdot g_X^+(s)\,ds = \mathcal{D}_X\big(\widetilde{w}_- \sqcup \mathcal{K}_\tau\big(g_W^-, \omega_+, g_X^+\big)\big)_{|[\tau,\tau+1]},$$

where $\mathcal{K}_\tau$ is defined as follows:

$$\mathcal{K}_\tau\big(g_W^-, \omega_+, g_X^+\big) := w_+ + \int_\tau^\cdot \mathcal{H}_\tau\big(g_W^-, g_X^+\big)_s\,ds,$$

with

$$\mathcal{H}_\tau(g_1, g_2)_t := C_1\mathcal{R}_0 g_1^\tau(t) + C_2\int_\tau^t (t-s)^{-\frac{1}{2}-H} g_2(s)\,ds,$$

and where we recall that the notation $\mathcal{R}_T$ has been introduced in (4.17). These transforms are then used in the successive binding trials alluded to in Section 4.2.

6.3. *Achievement of Step* 1. As a conclusion of this section, we obtain the following result.

PROPOSITION 6.9. *Assume* (H1) *and* (H3). *Then, for all* $\alpha > 0$ *and* $K > 0$, *there exist constants* $\delta =: \boldsymbol{\delta}(\alpha, K) > 0$ *and* $C_K > 0$ *such that for each* $k \geq 1$, $(W, \widetilde{W})$ *can be built on* $[\tau_{k-1}, \tau_{k-1}+1]$ *in such a way that the following properties hold*:

(i) *One has* $\mathbb{P}(Y_{\tau_{k-1}+1} = \widetilde{Y}_{\tau_{k-1}+1} | \mathcal{E}_{k-1} \cap A_k(K,\alpha,\gamma)) \geq \delta$.

(ii) *If* $\omega \in A_k(K,\alpha,\gamma)$ *and Step* 1 *is successful, then*

$$\sup_{t\in[0,1]} \big|g_X(t+\tau_{k-1})\big| \leq C_K.$$

(iii) $\int_{\tau_{k-1}}^{\tau_{k-1}+1} |g_W(s)|^2\,ds \leq C_K$ *a.s.*

PROOF. At the price of a change of variable, we can assume that $\tau_{k-1} = 0$. The construction of the coupling follows the lines of [18] and [11]. Let us however recall the principles below.

Step 1: Definition of $\mathcal{L}(W, \widetilde{W})$: With the notation 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's theorem, $\Lambda^*\mathbb{P}_W(dw) = D_\Lambda(w)\mathbb{P}_W(dw)$ where, with the notation of Theorem 6.7,

$$D_\Lambda(w) = \exp\biggl(\int_0^1 g_W^+(w,s)\,dw(s) - \frac{1}{2}\int_0^1 |g_W^+(w,s)|^2\,ds\biggr). \tag{6.33}$$

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 $(\mu_1 \wedge \mu_2)(dw) = D_1(w) \wedge D_2(w)\mu(dw)$. According to Theorem 6.7(ii), the way of gluing Y and $\widetilde{Y}$ at time $\tau_{k-1}+1$ implies the construction of a coupling $(W, \widetilde{W})$ such that $\widetilde{W} = \Lambda(W)$ on $[\tau_{k-1}, \tau_{k-1}+1]$ with lower-bounded probability. However, let us recall that in this non-Markovian setting, we also need to control the distance between W and $\widetilde{W}$ on the event where the coalescent coupling fails. This particular feature leads to a construction of the coupling which slightly differs from the classical maximal coupling, where the components are independent conditionally to the failure (see, e.g., [25]). Namely, with the help of the invertibility of Λ proved in Theorem 6.7, one defines a nonnegative measure $\mathbf{P}_1$ on Ω_+^2 by

$$\mathbf{P}_1 = \frac{1}{2}\bigl(\Lambda_1^*\mathbb{P}_W \wedge \Lambda_2^*\mathbb{P}_W\bigr),$$

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

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

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). Indeed, if one goes back to the definition (6.30) of Λ, it is readily checked that the function $g_X(w_+)$ therein depends also on w_-. In addition, for any bounded measurable function F defined on Ω_+^2, we have

$$\begin{aligned}
\mathbb{E}_{\Lambda_1^*\mathbb{P}}[F] &= \int_{\Omega_+} F\bigl(w, \Lambda(w)\bigr)\mathbb{P}(dw) \\
&= \int_{\Omega_+} F(w_1, w_2)\mathbf{1}_{(w_2=\Lambda(w_1))} D_\Lambda(w_2)\mathbb{P}(dw_2) \\
&= \int_{\Omega_+} F(w_1, w_2)\mathbf{1}_{(w_1=\Lambda^{-1}(w_2))} D_\Lambda(w_2)\mathbb{P}(dw_2).
\end{aligned}$$

Therefore we get

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

where D_Λ is defined by (6.33). This implies that $\mathbf{P}_1$ satisfies

$$\mathbf{P}_1(dw_1, dw_2) = \frac{1}{2}\mathbf{1}_{\{(\Lambda^{-1}(w),w)\}}(w_1, w_2)\big(D_\Lambda(w) \wedge 1\big)\mathbb{P}_W(dw). \tag{6.34}$$

Write $S(w_1, w_2) = (w_2, w_1)$ and denote by $\widetilde{\mathbf{P}}_1$ the "symmetrized" nonnegative 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:

$$\mathcal{L}\big((W_t^{\tau_{k-1}}, \widetilde{W}_t^{\tau_{k-1}})_{t\in[0,1]}\big) = \widetilde{\mathbf{P}}_1 + \Delta^*\big(\mathbb{P}_W - \Pi_1^*\widetilde{\mathbf{P}}_1\big) = \mathbf{P}_1 + \mathbf{P}_2, \tag{6.35}$$

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 (6.34), we check that for nonnegative functions f,

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

hence $\mathbf{P}_2$ is the sum of two nonnegative 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 $\mathbb{P}_W$. In conclusion, the coupling (6.35) between W and $\widetilde{W}$ has been achieved in such a way that:

(1) Each marginal is the distribution of a Brownian motion.

(2) Only three possibilities occur: $w_2 = \Lambda(w_1)$ (under $\mathbf{P}_1$) or $w_2 = \Lambda^{-1}(w_1)$ (under $S^*\mathbf{P}_1$) or $w_1 = w_2$ [under $\Delta^*(\mathbb{P}_W - \Pi_1^*\widetilde{\mathbf{P}}_1)$] so that $g_W = 0$ in that last case. In particular, whenever the coupling fails, the distance between the two Brownian motions is still controlled.

Step 2: Proof of statements (ii) and (iii): Statement (ii) is a direct consequence of the last statement of Theorem 6.7. For (iii), the result is obvious if $\omega \notin A_k(K, \alpha, \gamma)$. Otherwise, this is a consequence of the statement (2) above and from Theorem 6.7(iv).

Step 3: Proof of statement (i): Recall that $\tilde{P}_1$ denotes the distribution of $(W^{\tau_{k-1}}, \widetilde{W}^{\tau_{k-1}})$ on $[\tau_{k-1}, \tau_{k-1} + 1]$ and that $\widetilde{W}_t^{\tau_{k-1}} = \Lambda(W_t^{\tau_{k-1}})$ under the subprobability $\mathbf{P}_1$. Set $\mathcal{R} := \{w, y^{\pi,w}(1) = \tilde{y}^{\pi,\Lambda(w)}(1)\}$ 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 $(w, \Lambda(w))$. With the notation of Theorem 6.7, $\mathcal{R}$ can be more precisely written as follows:

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

where $\pi := (a, \tilde{a}, w_-, \tilde{w}_-) \in (\mathbb{R}^d)^2 \times \Omega_-^2$ stands for the past of the system up to time $\tau_{k-1} = 0$. As a consequence,

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

Now, by Theorem 6.7, we know that there exists a positive δ_K (depending only on K) such that $\mathbb{P}_W(\mathcal{R}) \ge \delta_K$. Our claim (i) is thus reduced to lower bound $\|\mathbf{1}_{\mathcal{R}\times\Lambda(\mathcal{R})}\mathbf{P}_1\|_{\mathrm{TV}}$ in terms of $\mathbb{P}_W(\mathcal{R})$.

The aforementioned lower bound is obtained as follows: by (6.34) and Lemma C.1 of [26] (applied to $p = 2$, $\mu_1 = \Lambda^* \mathbb{P}_W$, $\mu_2 = \mathbb{P}_W$ and $X = \mathcal{R}$) we have

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

Following the lines of the proof of [11], Lemma 3.1, and using that $w \mapsto \int_0^1 |g_W^+(w,s)|^2\,ds$ is bounded by a constant depending only on K [see Theorem 6.7(iv)], one deduces that

$$\|\mathbf{1}_{\mathcal{R}\times\Lambda(\mathcal{R})}\mathbf{P}_1\|_{\mathrm{TV}} \geq C\big[\mathbb{P}_W(\mathcal{R})\big]^2 \geq C\delta_K > 0,$$

which is the desired lower bound. □

7. About Step 2. As explained in Section 4.2, and following the ideas of [11, 18], Step 2 consists in a series of trials to keep Y and $\widetilde{Y}$ as equal on successive intervals $I_{k,\ell} := [s_{k,\ell}, s_{k,\ell+1}]$ of length $c_2 2^\ell$. To be more specific, for every fixed $k \geq 1$, we define $(s_{k,\ell})_{\ell\geq 0}$ by

$$s_{k,0} = s_{k,1} = \tau_{k-1} + 1 \quad \text{and} \quad \text{for every } \ell \geq 1 \qquad s_{k,\ell+1} = s_{k,\ell} + c_2 2^\ell. \tag{7.1}$$

From a pathwise point of view, an obvious way to achieve our goal here, that is, to keep the paths Y and $\widetilde{Y}$ glued together is to set $g_X(t) = 0$ after time $\tau_{k-1} + 1$, which by the one-to-one connection of Lemma 4.2, amounts to setting

$$g_W(t) = g_S(t) := \big(\mathcal{R}_0 g_W^{\tau_{k-1}+1}\big)(t). \tag{7.2}$$

The aim then is to extend the previous Brownian coupling $(W, \widetilde{W})$ in such a way that, with some controlled probability, Condition (7.2) is indeed satisfied on the successive intervals $I_{k,\ell}$. Using the notation $\mathcal{B}_{k,\ell}$ introduced in (4.23) and with the proof of Theorem 2.5 in mind [see (4.34) and (4.39)], we are more precisely interested in the control of the related quantity $\mathbb{P}(\mathcal{B}_{k,\ell}|\mathcal{B}_{k,\ell-1})$, that is, the probability of respecting (7.2) on the interval $I_{k,\ell} = [s_{k,\ell}, s_{k,\ell+1}]$ provided it holds up to time $s_{k,\ell}$. This specific issue has been studied in [11], Section 3.2, for $H > \frac{1}{2}$. It turns out that the result therein obtained, as well as its proof, can be transposed into our setting without any change, which leads us directly to the following assertion.

PROPOSITION 7.1. *Let g_S be defined by (7.2) for $t \geq \tau_{k-1} + 1$. Then for every $\alpha \in (0, H)$ and every $K > 0$, there exist constants $\mathbf{C}^{\mathbf{2}}_{\alpha,K} \geq 1$, $\mathbf{C}^{\mathbf{2},\mathbf{1}}_{\alpha,K} \geq 1$ and $\rho^1_{\alpha,K}, \rho^2_{\alpha,K} \in (0,1)$, which do not depend on k and such that the following properties hold*:

(a) *On the event $A_k(K, \alpha, \gamma)$, one has*

$$\int_0^{+\infty} (1+t)^{2\alpha} |g_S(\tau_{k-1} + 1 + t)|^2\,dt \leq \mathbf{C}^{\mathbf{2}}_{\alpha,K}.$$

(b) *One can extend the coupling* $(W, \widetilde{W})$ *along Condition* (7.2) *in such a way that, calibrating Step* 2 *by the formula* $c_2 := (\mathbf{C}^{\mathbf{2}}_{\alpha,K})^{\frac{1}{2\alpha}}$, *one has*

$$\rho^1_{\alpha,K} \le \mathbb{P}(\mathcal{B}_{k,1}|\mathcal{B}_{k,0}) \le \rho^2_{\alpha,K},$$

and for all $k \ge 0$, $\ell \ge 2$,

$$1 - 2^{-\alpha\ell} \le \mathbb{P}(\mathcal{B}_{k,\ell}|\mathcal{B}_{k,\ell-1}) \le 1 - 2^{-\alpha\ell-1}. \tag{7.3}$$

(c) *On the event* $F_{k,\ell}$ ($\ell \ge 1$) *defined by* (4.22) *and under the same calibration* $c_2 := (\mathbf{C}^{\mathbf{2}}_{\alpha,K})^{\frac{1}{2\alpha}}$, *one has*

$$\int_{s_{k,1}}^{s_{k,2}} |g_W(t)|^2\, dt \le \mathbf{C}^{\mathbf{2,1}}_{\alpha,K},$$

and if $\ell \ge 2$,

$$\int_{s_{k,\ell}}^{s_{k,\ell+1}} |g_W(t)|^2\, dt \le \big(2(\ell+3)\big)^2,$$

$$\int_{s_{k,p-1}}^{s_{k,p}} |g_W(t)|^2\, dt \le 2^{-2\alpha p}, \qquad p \in \{2, \dots, \ell\}.$$

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. In the subsequent statements, we recall that for all $\alpha \in (0, H)$ and $K > 0$, the notation $\mathbf{C}^{\mathbf{2}}_{\alpha,K}$ refers to the constant in $(1, +\infty)$ provided by Proposition 7.1.

PROPOSITION 8.1. *Let* $H \in (1/3, 1/2)$ *and assume* (H1), (H2), (H3) *hold true. For all* $\varepsilon \in (0, 1)$, $\alpha \in (0, H)$, $\beta > (1 - 2\alpha)^{-1}$ *and* $\varsigma > 1$, *there exist strictly positive constants*

$$K =: \mathbf{K}(\varepsilon, \alpha), \qquad c_3 =: \mathbf{c_3}(\varepsilon, \alpha, \beta, \varsigma) \tag{8.1}$$

such that calibrating Step 2 *and Step* 3 *along the formulas*

$$c_2 := \big(\mathbf{C}^{\mathbf{2}}_{\alpha,K}\big)^{\frac{1}{2\alpha}}, \qquad \Delta_3(k, \ell) := c_3 \varsigma^k 2^{\beta\ell} \tag{8.2}$$

yields that $c_3 \ge 2c_2$ *and for every* $k \ge 1$

$$\mathbb{P}\big(A_{k+1}(K, \alpha, \gamma)|\mathcal{E}_k\big) \ge 1 - \varepsilon. \tag{8.3}$$

REMARK 8.2. Let us insist on the fact that, in accordance with our notation in (8.1), the function $\mathbf{K}$ so defined depends on the two parameters ε and α only, whereas $\mathbf{c_3}$ depends both on (ε, α) and on (β, ς). This dependence issue is of paramount importance in the proof of Theorem 2.5, as we have seen it in Section 4.3.

The proof of Proposition 8.1 will actually be obtained as a consequence of the three following lemmas. We assume here that $H \in (1/3, 1/2)$ is fixed, and that Hypotheses (H1), (H2), (H3) hold true.

LEMMA 8.3. *For all $\alpha \in (0, H)$, $\beta > (1-2\alpha)^{-1}$, $K > 0$ and $\varsigma > 1$, there exists a constant $\mathbf{c_{3,1}}(\alpha, \beta, K, \varsigma) > 0$ such that for every $c_3 \geq \mathbf{c_{3,1}}(\alpha, \beta, K, \varsigma)$, calibrating Step 2 and Step 3 along the formulas in (8.2) yields that for every $k \geq 1$,*

$$\mathbb{P}\Big(\sup_{T\geq 0}\int_0^{+\infty}(1+t)^{2\alpha}\big((\mathcal{R}_T|g_W^{\tau_k}|)(t)\big)^2\,dt \leq 1\Big|\mathcal{E}_k\Big) = 1.$$

LEMMA 8.4. *For all $\alpha \in (0, H)$ and $\varsigma > 1$, there exist constants $\mathbf{C}_\alpha^{\mathbf{3,2}} > 0$ and $\mathbf{c_{3,2}}(\alpha, \varsigma) > 0$ such that for all $K > 0$ and $c_3 \geq \mathbf{c_{3,2}}(\alpha, \varsigma)$, calibrating the scheme along the formulas in (8.2) yields that*

$$\sup_{k\geq 0}\max\big(\mathbb{E}[|||D^{\tau_k}(W)|||_{1;\gamma}|\mathcal{E}_k], \mathbb{E}[|||D^{\tau_k}(\widetilde{W})|||_{1;\gamma}|\mathcal{E}_k]\big) \leq \mathbf{C}_\alpha^{\mathbf{3,2}}.$$

LEMMA 8.5. *There exists a constant $p \in (0,1)$ and for all $\alpha \in (0, H)$ and $\varsigma > 1$, there exist constants $\mathbf{C}_\alpha^{\mathbf{3,3}} > 0$ and $\mathbf{c_{3,3}}(\alpha, \varsigma) > 0$ such that for all $K > 0$ and $c_3 \geq \mathbf{c_{3,3}}(\alpha, \varsigma)$, calibrating the scheme along the formulas in (8.2) yields that*

$$\sup_{k\geq 0}\max\big(\mathbb{E}[|Y_{\tau_k}|^p|\mathcal{E}_k], \mathbb{E}[|\widetilde{Y}_{\tau_k}|^p|\mathcal{E}_k]\big) \leq \mathbf{C}_\alpha^{\mathbf{3,3}}.$$

REMARK 8.6. Just as above, let us stress the fact that, as indicated by our notation, the constants $\mathbf{C}_\alpha^{\mathbf{3,2}}$ and $\mathbf{C}_\alpha^{\mathbf{3,3}}$ in Lemmas 8.4 and 8.5 only depend on α, and not on ς. This will be an essential point in the subsequent proof of Proposition 8.1.

Before we turn to the proof of these three lemmas, let us see how their combination can lead to the desired proposition.

PROOF OF PROPOSITION 8.1. Fix $\varepsilon \in (0,1)$, $\alpha \in (0, H)$, $\beta > (1-2\alpha)^{-1}$, $\varsigma > 1$, and let $p \in (0,1)$, $\mathbf{C}_\alpha^{\mathbf{3,2}}$, $\mathbf{C}_\alpha^{\mathbf{3,3}}$, $\mathbf{c_{3,2}}(\alpha, \varsigma)$ and $\mathbf{c_{3,3}}(\alpha, \varsigma)$ be defined as in Lemmas 8.4 and 8.5. Then for all $K > 0$ and $c_3 \geq \max(\mathbf{c_{3,2}}(\alpha, \varsigma), \mathbf{c_{3,3}}(\alpha, \varsigma))$, calibrating the scheme as in (8.2) yields that

$$\begin{aligned}
&\mathbb{P}\big(|Y_{\tau_k}| + |\widetilde{Y}_{\tau_k}| + |||D^{(\tau_k)}(W)|||_{1;\gamma} + |||D^{(\tau_k)}(\widetilde{W})|||_{1;\gamma} \leq K|\mathcal{E}_k\big)\\
&\quad\geq 1 - \mathbb{P}\big(|Y_{\tau_k}|^p + |\widetilde{Y}_{\tau_k}|^p + |||D^{(\tau_k)}(W)|||_{1;\gamma}^p + |||D^{(\tau_k)}(\widetilde{W})|||_{1;\gamma}^p > K^{p\alpha}|\mathcal{E}_k\big)\\
&\quad\geq 1 - \frac{1}{K^p}\big\{\mathbb{E}[|Y_{\tau_k}|^p|\mathcal{E}_k] + \mathbb{E}[|\widetilde{Y}_{\tau_k}|^p|\mathcal{E}_k]\\
&\qquad + \mathbb{E}[|||D^{(\tau_k)}(W)|||_{1;\gamma}^p|\mathcal{E}_k] + \mathbb{E}[|||D^{(\tau_k)}(\widetilde{W})|||_{1;\gamma}^p|\mathcal{E}_k]\big\}\\
&\quad\geq 1 - \frac{2}{K^p}\big\{\mathbf{C}_\alpha^{\mathbf{3,3}} + (\mathbf{C}_\alpha^{\mathbf{3,2}})^p\big\}.
\end{aligned}$$

Therefore, setting from now on

$$K = \mathbf{K}(\varepsilon, \alpha) := \big(2\varepsilon^{-1}\{\mathbf{C}_\alpha^{\mathbf{3,3}} + (\mathbf{C}_\alpha^{\mathbf{3,2}})^p\}\big)^{1/p},$$

we get that for every $c_3 \geq \max(\mathbf{c_{3,2}}(\alpha, \varsigma), \mathbf{c_{3,3}}(\alpha, \varsigma))$ and for the calibration in (8.2),

$$\mathbb{P}\big(|Y_{\tau_k}| + |\widetilde{Y}_{\tau_k}| + \|\!|D^{(\tau_k)}(W)|\!\|_{1;\gamma} + \|\!|D^{(\tau_k)}(\widetilde{W})|\!\|_{1;\gamma} \leq K|\mathcal{E}_k\big) \geq 1 - \varepsilon. \tag{8.4}$$

Then, appealing also to the notation of Lemma 8.3, we define

$$\mathbf{c_3}(\varepsilon, \alpha, \beta, \varsigma) := \max\big(\mathbf{c_{3,1}}(\alpha, \beta, K, \varsigma), \mathbf{c_{3,2}}(\alpha, \varsigma), \mathbf{c_{3,3}}(\alpha, \varsigma), 2(\mathbf{C}_{\alpha,K}^{\mathbf{2}})^{\frac{1}{2\alpha}}\big).$$

In this way, setting $c_3 := \mathbf{c_3}(\varepsilon, \alpha, \beta, \varsigma)$ and still calibrating the scheme as in (8.2), we deduce, on top of (8.4), that $c_3 \geq 2c_2$ and by Lemma 8.3

$$\mathbb{P}\bigg(\sup_{T\geq 0}\int_0^{+\infty}(1+t)^{2\alpha}\big((\mathcal{R}_T|g_W^{\tau_k}|)(t)\big)^2\,dt \leq 1\Big|\mathcal{E}_k\bigg) = 1. \tag{8.5}$$

The bound (8.3) immediately follows from (8.4) and (8.5). □

It remains us to prove Lemma 8.3, Lemma 8.4 and Lemma 8.5. It turns out that Lemma 8.3 can be shown along the very same arguments as in [11], Proposition 4.6, and, therefore, we will not return to this proof for the sake of conciseness. As for the strategy toward Lemma 8.4, respectively, Lemma 8.5, it is the topic of the subsequent Section 8.1, respectively, Section 8.2.

8.1. *Proof of Lemma 8.4.* Let us first remark that $[D^{(\tau)}(W)]'$ is well defined (see Lemma 6.5) and satisfies

$$\big[D^{(\tau)}(W)\big]'_t = \alpha_H\bigg(H - \frac{1}{2}\bigg)\int_{-\infty}^{\tau}(t+\tau-r)^{H-\frac{3}{2}}\,dW_r.$$

In this section, our computations will hinge on a related incremental process, defined as follows: for $u \leq v \leq \tau$,

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

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

$$\big[D^{(\tau_k)}(W)\big]'_t = \alpha_H\bigg(H - \frac{1}{2}\bigg)\bigg(\mathcal{D}_{-\infty,0}^{\tau_k}(t) + \sum_{m=1}^{k}\mathcal{D}_{\tau_{m-1},\tau_m}^{\tau_k}(t)\bigg). \tag{8.6}$$

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

LEMMA 8.7. *For every $0 \le u < v < \tau$ and every $t > 0$, one has almost surely*

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

and

$$|\mathcal{D}^{\tau}_{-\infty,0}(t)| \le c_H \int_{-\infty}^{0} (t+\tau-r)^{H-\frac{5}{2}} |W_r| \, dr$$

for some deterministic constant $c_H > 0$.

PROOF. Both bounds follow from an integration-by-parts argument similar to the one we performed in (6.23). The second bound also involves the fact that $\lim_{u\to-\infty} u^{H-\frac{3}{2}} |W_u| = 0$ almost surely. □

We now state some controls related to the decomposition (8.6).

LEMMA 8.8. *Let $\alpha > 0$ and assume that for some (fixed) calibration of the scheme, there exists $\eta \in (0,1)$ such that for all $k \ge 1$, $\ell \ge 0$ and $K > 0$,*

$$\mathbb{P}(\mathcal{E}_k | \mathcal{E}_{k-1}) \ge \eta, \qquad \mathbb{P}(F_{k,\ell} | \mathcal{E}_{k-1}) \le 2^{-\alpha\ell} \quad \text{and} \quad \Delta\tau_k \ge a_k \qquad a.s., \tag{8.7}$$

where $F_{k,\ell}$ and $\mathcal{E}_k$ are respectively defined in (4.22) and (4.24), and $(a_k)_{k\ge1}$ is a deterministic sequence such that $a_k \ge 1$ for every $k \ge 1$. Then there exists a constant $C^2_{\eta,\alpha} > 0$ and for every $p > 0$ there exists a constant $C^1_{\eta,\alpha,p} > 0$ such that for every $k \ge 1$,

$$\begin{aligned} &\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma} |\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)|\Big) | \mathcal{E}_k\Big] \\ &\qquad \le \frac{C^1_{\eta,\alpha,p}}{a_k^{1/2-H} \eta^{(k-m)/p}}, \qquad m \in \{1,\dots,k-1\}, \end{aligned} \tag{8.8}$$

$$\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma} |\mathcal{D}^{\tau_k}_{-\infty,0}(t)|\Big) | \mathcal{E}_k\Big] \le \frac{C^1_{\eta,\alpha,p}}{a_k^{1/2-H} \eta^{k/p}}, \tag{8.9}$$

and

$$\begin{aligned} &\max\Big(\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma} |\mathcal{D}^{\tau_k}_{\tau_{k-1},\tau_k}(t)|\Big) | \mathcal{E}_k\Big], \mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma} |\mathcal{D}^{0}_{-\infty,0}(t)|\Big)\Big]\Big) \\ &\qquad \le C^2_{\eta,\alpha}. \end{aligned} \tag{8.10}$$

PROOF. For the sake of conciseness, the details behind these technical estimates are given in the Supplementary Material [9], Section 5. □

REMARK 8.9. Observe that the second assumption of (8.7) holds true by Proposition 7.1(b) as soon as $c_2 =:= (\mathbf{C}^{\mathbf{2}}_{\alpha,K})^{\frac{1}{2\alpha}}$. As well, with the help of the upper-bound in (7.3), we have for instance:

$$\mathbb{P}(\mathcal{E}_k|\mathcal{E}_{k-1}) \geq 1 - \mathbb{P}(\mathcal{B}_{k,2}|\mathcal{B}_{k,1}) \geq \eta = 2^{-2\alpha-1}.$$

Remark that the previous inequality corresponds to an upper bound for the probability of success of the attempt. In fact, this (possibly surprising) technical condition gives us a way to control the effect of conditioning by the event $\mathcal{E}_k$ throughout the computations (see [9], Section 5, for more details).

With Lemma 8.8 in hand, we can now turn to the proof of our result.

PROOF OF LEMMA 8.4. We only prove the result for W, the proof for $\widetilde{W}$ being completely similar. Fix $\alpha > 0$ and $\varsigma > 1$. By Remark 8.9, the conditions of Lemma 8.8 are satisfied by the sequence $a_k := c_3\varsigma^k$, provided $c_3 \geq 1$. Now, by decomposition (8.6), it holds that for every $k \geq 0$,

$$\begin{aligned}\mathbb{E}[|||D^{\tau_k}(W)|||_{1;\gamma}|\mathcal{E}_k] \leq c_H \Bigg\{ &\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{-\infty,0}(t)|\Big)|\mathcal{E}_k\Big] \\ &+ \sum_{m=1}^{k} \mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)|\Big)|\mathcal{E}_k\Big]\Bigg\}. \end{aligned} \tag{8.11}$$

For $k = 0$, we can use (8.10) to assert that

$$\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{0}_{-\infty,0}(t)|\Big)\Big] \leq C^2_{\eta,\alpha}. \tag{8.12}$$

For $k \geq 1$, combining the three bounds (8.8)–(8.9)–(8.10) gives that for every $p > 0$,

$$\begin{aligned}&\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{-\infty,0}(t)|\Big)|\mathcal{E}_k\Big] + \sum_{m=1}^{k} \mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)|\Big)|\mathcal{E}_k\Big] \\ &\quad\leq \frac{C^1_{\eta,\alpha,p}}{(c_3\varsigma^k)^{1/2-H}} \sum_{m=0}^{k-1} \frac{1}{\eta^{(k-m)/p}} + C^2_{\eta,\alpha} \leq \frac{C^{1,1}_{\eta,\alpha,p}}{c_3^{1/2-H}} \Big(\frac{\varsigma^{H-1/2}}{\eta^{1/p}}\Big)^k + C^2_{\eta,\alpha}.\end{aligned}$$

At this point, and since $\varsigma > 1$, we can pick $p = p(\varsigma,\eta) > 0$ such that

$$\frac{\varsigma^{H-1/2}}{\eta^{1/p}} \leq 1,$$

which entails that

$$\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{-\infty,0}(t)|\Big)|\mathcal{E}_k\Big] + \sum_{m=1}^{k}\mathbb{E}\Big[\Big(\sup_{t\in(0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_k}_{\tau_{m-1},\tau_m}(t)|\Big)|\mathcal{E}_k\Big] \leq \frac{C^{1,2}_{\eta,\alpha,\varsigma}}{c_3^{1/2-H}} + C^2_{\eta,\alpha}. \tag{8.13}$$

By injecting (8.12)–(8.13) into (8.11) and setting

$$\mathbf{c_{3,2}}(\alpha,\varsigma) := \max\big(1, (C^{1,2}_{\eta,\alpha,\varsigma})^{1/(1/2-H)}\big),$$

we can conclude that for every $c_3 \geq \mathbf{c_{3,2}}(\alpha,\varsigma)$ and every $k \geq 0$,

$$\mathbb{E}[\|\!|D^{\tau_k}(W)|\!\|_{1;\gamma}|\mathcal{E}_k] \leq c_H\{1 + C^2_{\eta,\alpha}\},$$

which corresponds to the desired estimate. □

8.2. *Proof of Lemma 8.5.* The argument is based on a combination of the Lyapunov control established in Theorem 3.5 and the properties of the noise shown in the previous section. Since the arguments are identical for Y and $\widetilde{Y}$, we only prove the statement for Y.

First, set $\rho := e^{-C_2/2}$ where C_2 is defined in Hypothesis (H2). By Theorem 3.5, we know that there exists a constant c (depending only on γ) such that for every $t \in \mathbb{R}_+$,

$$|Y_{t+1}|^2 \leq \rho|Y_t|^2 + c\{1 + \|\mathbf{X}^{(t)}\|^{\lambda}_{\gamma}\}, \tag{8.14}$$

where we have set $\lambda := \frac{8}{3\gamma-1}$, $X^{(t)}_s := X_{t+s} - X_t$ $(s \geq 0)$ and $\|\mathbf{X}^{(t)}\|_{\gamma} := \|\mathbf{X}^{(t)}\|_{\gamma,[0,1]}$. Accordingly, for every $p \in (0,2]$, we get that

$$|Y_{t+1}|^p \leq \rho^{p/2}|Y_t|^p + c^{p/2}\{1 + \|\mathbf{X}^{(t)}\|^{\lambda p/2}_{\gamma}\}. \tag{8.15}$$

A straightforward induction procedure then yields

$$|Y_{\tau_k}|^p \leq \big(\rho^{p/2}\big)^{\Delta\tau_k}|Y_{\tau_{k-1}}|^p + c^{p/2}\sum_{\ell=0}^{\Delta\tau_k-1}\big(\rho^{p/2}\big)^{\Delta\tau_k-\ell}\big(1 + \|\mathbf{X}^{(\tau_{k-1}+\ell)}\|^{\lambda p/2}_{\gamma}\big).$$

At this point, note that by Remark 8.9, we can rely on the existence of a parameter $\eta > 0$ (depending only on α) such that for every $k \geq 0$ and $K > 0$, $\mathbb{P}(\mathcal{E}_k|\mathcal{E}_{k-1}) \geq \eta$, and in particular $\mathbb{E}[|Y_{\tau_{k-1}}|^p|\mathcal{E}_k] \leq \eta^{-1}\mathbb{E}[|Y_{\tau_{k-1}}|^p|\mathcal{E}_{k-1}]$. Therefore, for all $p \in (0,2]$ and $c_3 \geq \frac{\log(\eta/2)}{\log\rho^{p/2}}$, we have, due to $\Delta\tau_k \geq c_3$,

$$\big(\rho^{p/2}\big)^{\Delta\tau_k} \leq \big(\rho^{p/2}\big)^{c_3} \leq \frac{\eta}{2},$$

and so

$$\begin{aligned}\mathbb{E}\big[|Y_{\tau_k}|^p|\mathcal{E}_k\big] &\le \frac{\eta}{2}\mathbb{E}\big[|Y_{\tau_{k-1}}|^p|\mathcal{E}_k\big] \\ &\quad + c^{p/2}\sum_{\ell=0}^{+\infty}\rho^\ell + c^{p/2}\mathbb{E}\bigg[\sum_{\ell=0}^{\Delta\tau_k}\big(\rho^{p/2}\big)^{\Delta\tau_k-\ell}\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{\lambda p/2}\Big|\mathcal{E}_k\bigg] \\ &\le \frac{1}{2}\mathbb{E}\big[|Y_{\tau_{k-1}}|^p|\mathcal{E}_{k-1}\big] \\ &\quad + \frac{c^{p/2}}{1-\rho} + c^{p/2}\mathbb{E}\bigg[\sum_{\ell=0}^{\Delta\tau_k}\big(\rho^{p/2}\big)^{\Delta\tau_k-\ell}\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{\lambda p/2}\Big|\mathcal{E}_k\bigg],\end{aligned}$$

which, by induction, entails that

$$\sup_{k\ge 0}\mathbb{E}\big[|Y_{\tau_k}|^p|\mathcal{E}_k\big] \le \mathbb{E}\big[|Y_{\tau_0}|^p|\mathcal{E}_0\big] + 2C_{\rho,p} = \mathbb{E}\big[|Y_0|^p\big] + 2C_{\rho,p},$$

where

$$C_{\rho,p} := \frac{c^{p/2}}{1-\rho} + c^{p/2}\mathbb{E}\bigg[\sum_{\ell=0}^{\Delta\tau_k}\big(\rho^{p/2}\big)^{\Delta\tau_k-\ell}\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{\lambda p/2}|\mathcal{E}_k\bigg]. \tag{8.16}$$

Let us recall here that we have assumed the existence of a parameter $r > 0$ such that $\mathbb{E}[|Y_0|^r] < +\infty$ (for $\widetilde{Y}_0$, one even knows that $\mathbb{E}[|\widetilde{Y}_0|^p] < +\infty$ for every $p > 0$, since the invariant measure has moments of any order). The conclusion now comes from the result of Proposition 8.10 below.

PROPOSITION 8.10. *For all $\alpha \in (0, H)$ and $\varsigma > 1$, there exists a constant $\bar{\mathbf{c}}(\alpha, \varsigma) > 0$ such that for all $K > 0$ and $c_3 \ge \bar{\mathbf{c}}(\alpha, \varsigma)$, calibrating the scheme along the formulas in (8.2) entails that for all $q \in (0, \frac{1}{4}]$ and $\rho \in (0, 1)$,*

$$\sup_{k\ge 1}\mathbb{E}\bigg[\sum_{\ell=0}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{q}\Big|\mathcal{E}_k\bigg] < +\infty.$$

PROOF. Let us again recall the existence of a parameter $\eta > 0$ (depending only on α) such that for every $k \ge 0$ and $K > 0$, $\mathbb{P}(\mathcal{E}_k|\mathcal{E}_{k-1}) \ge \eta$. Then, by using the result (and the notation) of [9], Lemma 5.2, one obtains that

$$\begin{aligned}&\mathbb{E}\bigg[\sum_{\ell=0}^{\Delta\tau_k}\rho^{\Delta\tau_k-\ell}\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{q}\Big|\mathcal{E}_k\bigg] \\ &\quad\le c_{\eta,\alpha}\cdot\sup_{m\ge 1}\sum_{\ell=0}^{\Delta(k,m)}\rho^{\Delta(k,m)-\ell}\mathbb{E}\big[\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{2q}|\mathcal{E}_{k-1}\big]^{\frac{1}{2}} \\ &\quad\le \frac{c_{\eta,\alpha}}{1-\rho}\cdot\sup_{\ell,k}\mathbb{E}\big[\big\|\mathbf{X}^{(\tau_{k-1}+\ell)}\big\|_\gamma^{1/2}|\mathcal{E}_{k-1}\big]^{2q}.\end{aligned}$$

Second, by [9], Corollary 3.3, we know that

$$\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma \le c\big(1+\|\mathbf{Z}^{(\tau_{k-1}+\ell)}\|_\gamma^2 + |||D^{(\tau_{k-1}+\ell)}|||_{1;\gamma}^2\big) \tag{8.17}$$

for some constant c that depends only on γ, and so

$$\begin{aligned}&\sup_{\ell,k}\mathbb{E}\big[\|\mathbf{X}^{(\tau_{k-1}+\ell)}\|_\gamma^{1/2}|\mathcal{E}_{k-1}\big]\\ &\quad\le c^{1/2}\Big(1+\sup_{\ell,k}\mathbb{E}\big[\|\mathbf{Z}^{(\tau_{k-1}+\ell)}\|_\gamma|\mathcal{E}_{k-1}\big]+\sup_{\ell,k}\mathbb{E}\big[|||D^{(\tau_{k-1}+\ell)}|||_{1;\gamma}|\mathcal{E}_{k-1}\big]\Big).\end{aligned}$$

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

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

where the latter expectation is known to be finite (see [14], Theorem 15.33).

Then, similarly to (8.6), one has the decomposition

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

On the one hand, the fact that, for any c_3 large enough (depending on α and ς), the quantity

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

is uniformly bounded in ℓ and k can be shown with similar arguments as those in the proof of Lemma 8.4. To be more specific, the idea is to start from a similar decomposition as the one in (8.6) and to control each term with the help of Lemma 8.7. Since the right-hand term in the latter lemma decreases with τ, the dependency in ℓ can be managed as follows: for all $\ell\ge 0$ and $u<v$,

$$\begin{aligned}&|\mathcal{D}^{\tau_{k-1}+\ell}_{u,v}(t)|\\ &\quad\le c_H\Big((t+\tau_{k-1}-u)^{H-\frac{3}{2}}|W_v-W_u|+\int_u^v(t+\tau_{k-1}-r)^{H-\frac{5}{2}}|W_v-W_r|\,dr\Big),\end{aligned}$$

and from here we can mimic the arguments of the proof of Lemma 8.8.

As far as the process $\mathcal{D}^{\tau_{k-1}+\ell}_{\tau_{k-1},\tau_{k-1}+\ell}$ is concerned, we can use the independency and stationary properties of the Brownian motion, together with the bound of Lemma 8.7, to assert that for every $\ell\ge 1$,

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t)||\mathcal{E}_{k-1}\Big]\\ &\quad\le c_H\mathbb{E}\Big[\sup_{t\in[0,1]} t^{1-\gamma}\Big((t+\ell)^{H-\frac{3}{2}}|W_\ell|+\int_0^\ell(t+\ell-r)^{H-\frac{5}{2}}|W_r|\,dr\Big)\Big].\end{aligned}$$

For the first term, it is enough to observe that $t^{1-\gamma}(t+\ell)^{H-\frac{3}{2}}|W_\ell| \le \ell^{-\frac{1}{2}}|W_\ell|$. For the second term, we can write

$$\begin{aligned} t^{1-\gamma}&\int_0^\ell (t+\ell-r)^{H-\frac{5}{2}}|W_r - W_\ell|\,dr \\ &\le \int_0^{\ell-1} (\ell-r)^{H-2}(\ell-r)^{-\frac{1}{2}}|W_r - W_\ell|\,dr \\ &\quad + \int_{\ell-1}^{\ell} (\ell-r)^{H-1-\gamma}(\ell-r)^{-\frac{1}{2}}|W_r - W_\ell|\,dr, \end{aligned}$$

where we have used the fact that for every $r \in [\ell-1, \ell]$ and every $t \in (0,1]$,

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

The uniform boundedness of $\mathbb{E}[\sup_{t\in[0,1]} t^{1-\gamma}|\mathcal{D}^{\tau_{k-1}+\ell}_{-\infty,\tau_{k-1}}(t)||\mathcal{E}_{k-1}]$ follows immediately, and this achieves the proof of our assertion. □

Acknowledgments. 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).

SUPPLEMENTARY MATERIAL

Supplement to "Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise" (DOI: 10.1214/18-AOP1265SUPP; .pdf). Our supplement develops the proofs of several crucial but technical results in our paper. We first prove a Lyapunov-type property for rough differential equations with inward looking drifts. Then we handle rough differential equations involving singular drifts, a type of system which arises when one tries to condition in the highly non-Markovian fractional Brownian motion setting. Next, we show how to lift rough paths involving singularities. Finally, we evaluate some effects of our conditioning procedure on the underlying fractional Brownian motion X in equation (1.3).

REFERENCES

[1] ALÒS, E., MAZET, O. and NUALART, D. (2000). Stochastic calculus with respect to fractional Brownian motion with Hurst parameter lesser than $\frac{1}{2}$. *Stochastic Process. Appl.* **86** 121–139. MR1741199

[2] ARNOLD, L. (1998). *Random Dynamical Systems*. Springer, Berlin. MR1723992

[3] BAKRY, D., GENTIL, I. and LEDOUX, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Springer, Cham. MR3155209

[4] COHEN, S. and PANLOUP, F. (2011). Approximation of stationary solutions of Gaussian driven stochastic differential equations. *Stochastic Process. Appl.* **121** 2776–2801. MR2844540

[5] COHEN, S., PANLOUP, F. and TINDEL, S. (2014). Approximation of stationary solutions to SDEs driven by multiplicative fractional noise. *Stochastic Process. Appl.* **124** 1197–1225. MR3148010

[6] CRAUEL, H. (1993). Non-Markovian invariant measures are hyperbolic. *Stochastic Process. Appl.* **45** 13–28. MR1204858

[7] DAVIE, A. M. (2008). Differential equations driven by rough paths: An approach via discrete approximation. *Appl. Math. Res. Express. AMRX* **2008** Art. ID abm009. MR2387018

[8] DEYA, A., GUBINELLI, M., HOFMANOVA, M. and TINDEL, S. (2016). A priori estimates for rough PDE with application to rough conservation laws. Preprint. Available at arXiv:1604.00437.

[9] DEYA, A., PANLOUP, F. and TINDEL, S. (2019). Supplement to "Rate of convergence to equilibrium of fractional driven stochastic differential equations with rough multiplicative noise." DOI:10.1214/18-AOP1265SUPP.

[10] DOWN, D., MEYN, S. P. and TWEEDIE, R. L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. MR1379163

[11] FONTBONA, J. and PANLOUP, F. (2017). Rate of convergence to equilibrium of fractional driven stochastic differential equations with some multiplicative noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 503–538. MR3634264

[12] FRIZ, P. K., GESS, B., GULISASHVILI, A. and RIEDEL, S. (2016). The Jain–Monrad criterion for rough paths and applications to random Fourier series and non-Markovian Hörmander theory. *Ann. Probab.* **44** 684–738. MR3456349

[13] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths*: *With an Introduction to Regularity Structures*. Springer, Cham. MR3289027

[14] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669

[15] GARRIDO-ATIENZA, M. J., KLOEDEN, P. E. and NEUENKIRCH, A. (2009). Discretization of stationary solutions of stochastic systems driven by fractional Brownian motion. *Appl. Math. Optim.* **60** 151–172. MR2524684

[16] GUASONI, P. (2006). No arbitrage under transaction costs, with fractional Brownian motion and beyond. *Math. Finance* **16** 569–582. MR2239592

[17] GUBINELLI, M. (2004). Controlling rough paths. *J. Funct. Anal.* **216** 86–140. MR2091358

[18] HAIRER, M. (2005). Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.* **33** 703–758. MR2123208

[19] HAIRER, M. and MATTINGLY, J. C. (2008). Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. *Ann. Probab.* **36** 2050–2091. MR2478676

[20] HAIRER, M., MATTINGLY, J. C. and SCHEUTZOW, M. (2011). Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** 223–259. MR2773030

[21] HAIRER, M. and OHASHI, A. (2007). Ergodic theory for SDEs with extrinsic memory. *Ann. Probab.* **35** 1950–1977. MR2349580

[22] HAIRER, M. and PILLAI, N. S. (2013). Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. *Ann. Probab.* **41** 2544–2598. MR3112925

[23] JEON, J.-H., TEJEDOR, V., BUROV, S., BARKAI, E., SELHUBER-UNKEL, C., BERG-SØRENSEN, K., ODDERSHEDE, L. and METZLER, R. (2011). *In vivo* anomalous diffusion and weak ergodicity breaking of lipid granules. *Phys. Rev. Lett.* **106** 048103.

[24] KOU, S. C. (2008). Stochastic modeling in nanoscale biophysics: Subdiffusion within proteins. *Ann. Appl. Stat.* **2** 501–535. MR2524344

[25] LINDVALL, T. (1992). *Lectures on the Coupling Method*. Wiley, New York. MR1180522

[26] MATTINGLY, J. C. (2002). Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics. *Comm. Math. Phys.* **230** 421–462. MR1937652

[27] ODDE, D. J., TANAKA, E. M., HAWKINS, S. S. and BUETTNER, H. M. (1996). Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. *Biotechnol. Bioeng.* **50** 452–461.

[28] RÖCKNER, M. and WANG, F.-Y. (2001). Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. *J. Funct. Anal.* **185** 564–603. MR1856277

[29] VIENS, F. G. and ZHANG, T. (2008). Almost sure exponential behavior of a directed polymer in a fractional Brownian environment. *J. Funct. Anal.* **255** 2810–2860. MR2464192

A. DEYA
INSTITUT ELIE CARTAN
UNIVERSITÉ DE LORRAINE
B.P. 239
54506 VANDOEUVRE-LÈS-NANCY, CEDEX
FRANCE
E-MAIL: aurelien.deya@univ-lorraine.fr

F. PANLOUP
LAREMA
UNIVERSITÉ D'ANGERS
2, BD LAVOISIER
49045 ANGERS CEDEX 01
FRANCE
E-MAIL: fabien.panloup@univ-angers.fr

S. TINDEL
DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
150 N. UNIVERSITY STREET
WEST LAFAYETTE, INDIANA 47907
USA
E-MAIL: stindel@purdue.edu