Random attractors for dissipative systems with rough noises

Luu Hoang Duc *

Abstract

We provide an analytic approach to study the asymptotic dynamics of rough differential equations, with the driving noises of Hölder continuity. Such systems can be solved with Lyons' theory of rough paths, in particular the rough integrals are understood in the Gubinelli sense for controlled rough paths. Using the framework of random dynamical systems and random attractors, we prove the existence and upper semi-continuity of the global pullback attractor for dissipative systems perturbed by bounded noises. Moreover, if the unperturbed system is strictly dissipative then the random attractor is a singleton for sufficiently small noise intensity.

Keywords: stochastic differential equations (SDE), rough path theory, rough integrals, random dynamical systems, random attractors, stochastic perturbation, stochastic stability.

1 Introduction

This paper studies the asymptotic behavior of the stochastic differential equation

$$dy_t = f(y_t)dt + g(y_t)dX_t \tag{1.1}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$, $g : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^d)$ are of enough regularity, and $X_t \in \mathbb{R}^m$ is a stochastic process with stationary increments, such that almost sure all realizations are ν - Hölder continuous for some $\nu \in (\frac{1}{3}, 1)$ and $d, m \in \mathbb{N}$ (e.g. fractional Brownian motions [32] with Hurst indices $H \in (\frac{1}{3}, 1)$). It is well known that such equation can be solved by using Lyons' theory of rough paths (see [30], [31] and also [14]), namely one attempts to solve the controlled differential equation

$$dy_t = f(y_t)dt + g(y_t)dx_t, \tag{1.2}$$

for the driving path x to be a realization of X in the space $C^\nu(\mathbb{R}, \mathbb{R}^m)$ of continuous paths with finite ν - Hölder norm on any finite time interval. The solution of (1.2) is often understood in the sense of either Lyons-Davie [30], [31], or of Friz-Victoir [14], [35], which needs not to specify rough integrals. On the other hand, equation (1.2) can also be understood in the integral form

$$y_t = y_0 + \int_0^t f(y_s)ds + \int_0^t g(y_s)dx_s, \quad \forall t \geq 0, \tag{1.3}$$

where the second integral is a rough integral for controlled rough paths in the sense of Gubinelli [17]. As such, system (1.3) is recently proved in [10] to admit a unique path-wise solution given the initial condition. An alternative approach is to define rough integrals using fractional calculus, as studied for example in [25], [16], [24].

Our aim is to investigate the role of the driving noise in the longterm behavior of system (1.1). This question is studied in a probabilistic approach in the series [18], [19], [20], [21], in which they prove that, under the dissipativity and some additional regularity conditions there exists a unique

*Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany, & Institute of Mathematics, Viet Nam Academy of Science and Technology duc.luu@mis.mpg.de, lhduc@math.ac.vn

adapted stationary solution for (1.1) in the sense that the generated *stochastic dynamical system* over a stationary noise process has a unique invariant probability measure [19]. Moreover, the convergence is of probability type, i.e. other probability measures converge to the unique invariant measure in the total variation norm.

In this paper, we propose an analytic approach to study the stochastic perturbation problem. Namely we impose assumptions for the drift coefficient so that there exists a global attractor for the deterministic system

$$\dot{\mu} = f(\mu) \tag{1.4}$$

which is asymptotically stable, and then raise the questions on the asymptotic dynamics of the perturbed system (1.1), in particular the existence of stationary states and their asymptotic (stochastic) stability with respect to almost sure convergence.

Note that the classical methods [36], [26], [33] on stochastic stability depends crucially on the nature of Itô calculus, since one can take advantage of the martingale property of the noise to apply Ito's formula for a Lyapunov function and then take the expectation to eliminate the noise part. As a result, the expectation of the solution norm can be proved to decay exponentially to zero, which is enough to prove that the solution norm itself converges exponentially and almost surely to zero due to Borel-Catelli lemma. The situation is however different here with a general noise Z, which is neither a Markov process nor a semimartingale (e.g. fractional Brownian motion B^H [34]), hence the noise part does not vanish by taking the expectation. This challenge suggests that a new approach to study stochastic stability is necessary.

Results in [15] and recently in [2], [11], [12], [7] suggest that the above questions could be studied in the framework of random dynamical systems [1], hence asymptotic structures like random attractors are well-understood. In this scenarios, system (1.1) has no deterministic equilibrium but is expected to possess a random attractor, although little is known on the inside structure of the attractor and much less on whether or not the attractor is a (random) singleton.

Assumptions and main results

Throughout the paper, we will assume that.

($\mathbf{H}_f$) f is locally Lipschitz continuous and dissipative, i.e. there exist constants $D_1 \geq 0, D_2 > 0$ such that

$$\langle y, f(y)\rangle \leq \|y\|(D_1 - D_2\|y\|), \quad \forall y \in \mathbb{R}^d; \tag{1.5}$$

in addition f is of linear growth in the perpendicular direction, i.e. there exists $C_f > 0$ such that

$$\Big\| f(y) - \frac{\langle f(y), y\rangle}{\|y\|^2} y \Big\| \leq C_f \Big(1 + \|y\|\Big), \quad \forall y \neq 0; \tag{1.6}$$

($\mathbf{H}_g$) g belongs to $C^3_b(\mathbb{R}^d, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^d))$ such that

$$C_g := \max\Big\{\|g\|_\infty, \|Dg\|_\infty, \|D^2_g\|_\infty, \|D^3_g\|_\infty\Big\} < \infty; \tag{1.7}$$

($\mathbf{H}_X$) for a given $\nu \in (\frac{1}{3}, \frac{1}{2})$, x belongs to the space $C^\nu(\mathbb{R}, \mathbb{R}^m)$ of all continuous paths which is of finite $\nu-$Hölder norm on any interval $[s,t]$. In particular, x is a realization of a stochastic process $X_t(\omega)$ with stationary increments, such that x can be lifted into a realized component $\mathbf{x} = (x, \mathbb{X})$ of a stochastic process $(x_\cdot(\omega), \mathbb{X}_{\cdot,\cdot}(\omega))$ with stationary increments, and the estimate

$$E\Big(\|x_{s,t}\|^p + \|\mathbb{X}_{s,t}\|^q\Big) \leq C_{T,\nu}|t-s|^{p\nu}, \forall s,t \in [0,T] \tag{1.8}$$

holds for any $[0,T]$, with $p\nu \geq 1, q = \frac{p}{2}$ and some constant $C_{T,\nu}$.

Let us comment on the assumptions. As presented in Remark 3.1, assumption $(\mathbf{H}_X)$ is satisfied if X is a fractional Brownian motion B^H [32] with Hurst exponent $H \in (\frac{1}{3}, 1)$, i.e. a family of centered Gaussian processes $B^H = \{B^H_t\}_{t\in\mathbb{R}}$ with continuous sample paths and

$$E\|B^H_t - B^H_s\| = |t-s|^{2H}, \forall t, s \in \mathbb{R}.$$

Meanwhile, the local Lipchitz continuity, condition (1.6) and the one-sided Lipschitz continuity

$$\exists C > 0 : \langle y, f(y)\rangle \leq C(1+\|y\|^2), \quad \forall y \in \mathbb{R}^d$$

are require for f in order to prove the existence and uniqueness of the solution of (1.2), as well as the continuity of the solution semiflow and the generation of a continuous random dynamical system, see e.g. [35, Theorem 4.3] and [2]. In our situation, condition (1.5) is stronger than the one-sided Lipschitz continuity, and in fact is equivalent to the classical dissipativity, as shown in the following lemma.

Lemma 1.1 *Condition* (1.5) *is equivalent to the following condition: there exist constants* $d_1 \geq 0, d_2 > 0$ *such that*

$$\langle y, f(y)\rangle \leq d_1 - d_2\|y\|^2, \quad \forall y \in \mathbb{R}^d; \tag{1.9}$$

Proof: Assume (1.5) is satisfied, then Cauchy inequality yields

$$\langle y, f(y)\rangle \leq \frac{D_1^2}{2D_2} - \frac{D_2}{2}\|y\|^2 - \frac{1}{2}\Big(\sqrt{D_2}\|y\| - \frac{D_1}{\sqrt{D_2}}\Big)^2 \leq \frac{D_1^2}{2D_2} - \frac{D_2}{2}\|y\|^2,$$

which proves (1.9) by choosing $d_1 := \frac{D_1^2}{2D_2}$ and $d_2 := \frac{D_2}{2}$. For the other direction, one can easily show that

$$\langle y, f(y)\rangle \leq \|y\|\Big(\sup_{\|y\|\leq 1}\|f(y)\| + d_1 + d_2 - d_2\|y\|\Big), \forall y \in \mathbb{R}^d.$$

Indeed, if $\|y\| \leq 1$, then

$$\langle y, f(y)\rangle \leq \|y\| \sup_{\|y\|\leq 1}\|f(y)\| + d_2\|y\|(1-\|y\|) \leq \|y\|\Big(\sup_{\|y\|\leq 1}\|f(y)\| + d_1 + d_2 - d_2\|y\|\Big).$$

On the other hand, if $\|y\| \geq 1$ then by (1.9)

$$\langle y, f(y)\rangle \leq d_1 - d_2\|y\|^2 \leq d_1\|y\| - d_2\|y\|^2 \leq \|y\|\Big(\sup_{\|y\|\leq 1}\|f(y)\| + d_1 + d_2 - d_2\|y\|\Big).$$

Hence (1.5) is followed by choosing $D_1 := \sup_{\|y\|\leq 1}\|f(y)\| + d_1 + d_2$ and $D_2 := d_2$.

□

Due to Lemma 1.1, the deterministic system (1.4) is dissipative and admits a global attractor. In addition, the addition technical condition (1.6) is equivalent to the following: for $y \in \mathbb{R}^d$ and $y \neq 0$, $f(y)$ is decomposed in the unique form

$$f(y) = \frac{\langle f(y), y\rangle}{\|y\|^2}y + \pi_y^\perp(f(y)), \quad \text{where} \quad \pi_y^\perp = 1 - \pi_y \quad \text{and} \quad \|\pi_y^\perp(f(y))\| \leq C_f(1+\|y\|). \tag{1.10}$$

Condition (1.6) is automatically satisfied if f is globally Lipschitz continuous, i.e.

$$\|f(y_1) - f(y_2)\| \leq L_f\|y_1 - y_2\|, \quad \forall y_1, y_2 \in \mathbb{R}^d, \tag{1.11}$$

or if f is simply of linear growth, i.e. $\|f(y)\| \leq L_f(1+\|y\|)$. Thus the assumption $(\mathbf{H}_f)$ is weaker than the one in [19]. Nontrivial examples are presented in the following examples.

Example 1.2 Consider the vector field $f(y) = \chi y - \|y\|^2 y$ for all $y \in \mathbb{R}^d$, where $\chi > 0$ is a constant. Then it follows from Cauchy inequality that

$$\langle y, f(y) \rangle = \|y\|(\chi - \|y\|^3) \leq \|y\|(\chi + 2 - 3\|y\|).$$

On the other hand, $\pi_y^\perp(f(y)) = 0$ whenever $y \neq 0$. Hence (1.5) and (1.6) are satisfied.

By similar computations, one can easily check that the Poincaré-Andronov-Hopf vector field [22, Example 7.26, p. 208]

$$f(y) = \begin{pmatrix} by_2 + y_1(a - y_1^2 - y_2^2) \\ -by_1 + y_2(a - y_1^2 - y_2^2) \end{pmatrix}, \ \forall y = (y_1, y_2)^{\mathrm{T}} \in \mathbb{R}^2,$$

for constants $a, b > 0$, also satisfies conditions (1.5) and (1.6). The function $a - y_1^2 - y_2^2$ can also be generalized to $F(a, \|y\|)$ for a function F that makes f dissipative in the strong sense (see [22, Example 11.13, p.345]).

In addition, (1.5) ensures that there exists a global attractor $\mathcal{A}$ for the deterministic system (1.4) satisfying: for any solution μ_t starting at point $\mu_0 \in \mathcal{A}$, we have

$$\|\mu_t\| \leq \underbrace{\max\{\|\mu\| : \mu \in \mathcal{A}\}}_{=:|\mathcal{A}|}, \quad \|\mu_{s,t}\| \leq \int_s^t \underbrace{\max\{\|f(p_1)\| : p_1 \in \mathcal{A}\}}_{=:\|f\|_{\infty,\mathcal{A}}} du = \|f\|_{\infty,\mathcal{A}}(t-s), \quad \forall 0 \leq s < t, \tag{1.12}$$

thus $\mu \in C^{1-\mathrm{var}}$.

One approach to show stochastic perturbation is to prove that a global random attractor does exist and is upper semi-continuous w.r.t the intensity of the stochastic noise (see e.g. [3], [4], [23], [38]). To do that, we need to impose an additional property of *uniform attraction* for the global attractor $\mathcal{A}$ as follows.

($\mathbf{H}_\mathcal{A}$) There exists a duration $r > 0$ and constants $D_3 > 0$ of the deterministic system (1.4) such that, for any starting point $y_0 \notin \mathcal{A}$, there exists a point $\mu_0 = \mu_0(y_0) \in \mathcal{A}$ satisfying

$$\|\mu_r(y_0) - \mu_r(\mu_0)\| \leq e^{-D_3}\|y_0 - \mu_0\|. \tag{1.13}$$

For instance, asumption ($\mathbf{H}_\mathcal{A}$) is satisfied when f is *strictly dissipative*, i.e. $D_1 = 0$ in (1.5), by choosing $D_3 = D_2$ and $r = 1$. Another example is any planar system satisfying ($\mathbf{H}_f$) which admits a periodic orbit that also acts as the boundary of the global attractor, see e.g. [22, Chapter 11]. Condition (1.13) is then equivalent to the exponential stability of the fixed point of the Poincaré map.

Our main results (Theorem 3.3, Theorem 3.4, Theorem 3.7) show that, under the assumptions $(\mathbf{H}_f), (\mathbf{H}_g), (\mathbf{H}_X)$, there exists a random pullback attractor $\mathcal{A}(\omega)$ such that $|\mathcal{A}(\cdot)| \in \mathcal{L}^\rho$ for any $\rho \geq 1$. In addition, if condition ($\mathbf{H}_f$) is replaced by conditions on the relative dissipativity (will be specified later in Theorem 3.4), the global Lipschitz continuity (1.11) and ($\mathbf{H}_\mathcal{A}$), then the random attractor is upper semi-continuous with respect to the noise intensity in the sense that $\mathcal{A}(\omega) \to \mathcal{A}$ (w.r.t. the Hausdorff semi-distance) as $C_g \to 0$, both in the almost sure and in $\mathcal{L}^\rho$ senses. Moreover, if f is strictly dissipative then $\mathcal{A}(\omega)$ is a singleton provided that C_g is sufficiently small.

Our idea of the proof uses a well-known Doss-Sussmann technique [37], which was developed in [28], [27], [35], [10] for stochastic systems, i.e. using the transformation $y_t = \phi_t(\mathbf{x}, z_t)$ generated from the pure rough differential equation $d\phi_t = g(\phi_t)dx_t$. The solution of the transformed system

$$\dot{z}_t = \Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}, z_t)\Big]^{-1} f(\phi_t(\mathbf{x}, z_t)) \tag{1.14}$$

can then be estimated on each interval of a greedy sequence of stopping times generated from the rough path $\mathbf{x}$ [5]. The case $\nu > \frac{1}{3}$ is therefore just for the aim of simple presentation. Our results and methods in this paper still hold for smaller ν, provided that almost all realizations of the stochastic noise are truly rough so that the Gubinelli derivative can be uniquely defined (see details in Subsection 2.2). Plus, we would need additional information in the signatures of rough paths to define rough integrals for controlled rough paths.

Finally, we emphasize here that there is of course a similar way to achieve the results for solutions of (1.2) understood in the sense of Lyons-Davie, by using [5]. However, our usage of rough integrals in the Gubinelli sense is not just a matter of taste, but because it provides short and self-contained proofs, and can be generalized for studying infinite dimensional systems with rough noises, as partly seen in [7] for stochastic systems with time delays.

2 Preliminaries

2.1 Rough paths

Let us introduce the concept of rough paths, following [30] and [13]. Given any compact time interval $I = [\min I, \max I] \subset \mathbb{R}$, we write $|I| := \max I - \min I$ and $I^2 := I \times I$. For any finite dimensional vector space W, denote by $C(I, W)$ the space of all continuous paths $y : I \to W$ equipped with the sup norm $\|\cdot\|_{\infty,I}$ given by $\|y\|_{\infty,I} = \sup_{t\in I} \|y_t\|$, where $\|\cdot\|$ is the norm in W. We write $y_{s,t} := y_t - y_s$. For $p \geq 1$, denote by $C^{p-\rm{var}}(I, W) \subset C(I, W)$ the space of all continuous paths $y : I \to W$ of finite p-variation $|\!|\!| y |\!|\!|_{p-\rm{var},I} := \Big(\sup_{\Pi(I)} \sum_{i=1}^n \|y_{t_i,t_{i+1}}\|^p\Big)^{1/p} < \infty$, where the supremum is taken over the whole class of finite partition of I. It is well known [14] that $|\!|\!| y |\!|\!|^p_{p-\rm{var},I}$ is a control, i.e. it satisfies

$$|\!|\!| y |\!|\!|^p_{p-\rm{var},[s,s]} = 0, \qquad |\!|\!| y |\!|\!|^p_{p-\rm{var},[s,u]} + |\!|\!| y |\!|\!|^p_{p-\rm{var},[u,t]} \leq |\!|\!| y |\!|\!|^p_{p-\rm{var},[s,t]}, \quad \forall s \leq u \leq t. \tag{2.1}$$

Then $C^{p-\rm{var}}(I, W)$ with the equipped $p-$var norm $\|y\|_{p\text{-var},I} := \|y_{\min I}\| + |\!|\!| y |\!|\!|_{p-\rm{var,I}}$ is a nonseparable Banach space [14, Theorem 5.25, p. 92]. Also for each $0 < \alpha < 1$, we denote by $C^\alpha(I, W)$ the space of Hölder continuous functions with exponent α on I equipped with the norm

$$\|y\|_{\alpha,I} := \|y_{\min I}\| + |\!|\!| y |\!|\!|_{\alpha,I}, \quad \text{where} \quad |\!|\!| y |\!|\!|_{\alpha,I} := \sup_{s,t\in I,\ s<t} \frac{\|y_{s,t}\|}{(t-s)^\alpha} < \infty. \tag{2.2}$$

For $\alpha \in (\frac{1}{3}, \frac{1}{2})$, a couple $\mathbf{x} = (x, \mathbb{X}) \in \mathbb{R}^m \oplus (\mathbb{R}^m \otimes \mathbb{R}^m)$, where $x \in C^\alpha(I, \mathbb{R}^m)$ and

$$\mathbb{X} \in C^{2\alpha}(I^2, \mathbb{R}^m \otimes \mathbb{R}^m) := \{\mathbb{X} \in C(I^2, \mathbb{R}^m \otimes \mathbb{R}^m) : \sup_{s,t\in I,\ s<t} \frac{\|\mathbb{X}_{s,t}\|}{|t-s|^{2\alpha}} < \infty\},$$

is called a *rough path* if it satisfies Chen's relation

$$\mathbb{X}_{s,t} - \mathbb{X}_{s,u} - \mathbb{X}_{u,t} = x_{s,u} \otimes x_{u,t}, \qquad \forall \min I \leq s \leq u \leq t \leq \max I. \tag{2.3}$$

$\mathbb{X}$ is called a *Lévy area* for x and is viewed as *postulating* the value of the quantity $\int_s^t x_{s,r} \otimes dx_r := \mathbb{X}_{s,t}$ where the right hand side is taken as a definition for the left hand side. Denote by $\mathscr{C}^\alpha(I, \mathbb{R}^m \oplus (\mathbb{R}^m \otimes \mathbb{R}^m)) \subset C^\alpha(I, \mathbb{R}^m) \oplus C^{2\alpha}(I^2, \mathbb{R}^m \otimes \mathbb{R}^m)$ the set of all rough paths $\mathbf{x}$ on I (or in short $\mathscr{C}^\alpha(I)$), then $\mathscr{C}^\alpha(I)$ is a closed set (but not a linear space), equipped with the rough path semi-norm

$$|\!|\!| \mathbf{x} |\!|\!|_{\alpha,I} := |\!|\!| x |\!|\!|_{\alpha,I} + |\!|\!| \mathbb{X} |\!|\!|^{\frac{1}{2}}_{2\alpha,I^2}, \quad \text{where} \quad |\!|\!| \mathbb{X} |\!|\!|_{2\alpha,I^2} := \sup_{s,t\in I; s<t} \frac{\|\mathbb{X}_{s,t}\|}{|t-s|^{2\alpha}} < \infty. \tag{2.4}$$

Throughout this paper, we will fix parameters $\frac{1}{3} < \alpha < \nu < \frac{1}{2}$ and $p = \frac{1}{\alpha}$ so that $C^\alpha(I, W) \subset C^{p-\rm{var}}(I, W)$. We also set $q = \frac{p}{2}$ and consider the $p-$var semi-norm

$$|\!|\!| \mathbf{x} |\!|\!|_{p-\rm{var},I} := \Big(|\!|\!| x |\!|\!|^p_{p-\rm{var},I} + |\!|\!| \mathbb{X} |\!|\!|^q_{q-\rm{var},I^2} \Big)^{\frac{1}{p}}, \quad |\!|\!| \mathbb{X} |\!|\!|_{q-\rm{var},I^2} := \left(\sup_{\Pi(I)} \sum_{i=1}^n \|\mathbb{X}_{t_i,t_{i+1}}\|^q\right)^{1/q}, \tag{2.5}$$

where the supremum is taken over the whole class of finite partitions $\Pi(I)$ of I. Sometimes, we write $\mathscr{C}^\alpha(I)$ for abbreviation to neglect the value space for simplicity of presentation.

2.2 Rough integrals

Following [17], a rough integral can be defined for a continuous path $y \in C^\alpha(I, W)$ which is *controlled by* $x \in C^\alpha(I, \mathbb{R}^m)$ in the sense that, there exists a couple (y', R^y) with $y' \in C^\alpha(I, \mathcal{L}(\mathbb{R}^m, W))$, $R^y \in C^{2\alpha}(I^2, W)$ such that

$$y_{s,t} = y'_s x_{s,t} + R^y_{s,t}, \qquad \forall \min I \le s \le t \le \max I. \tag{2.6}$$

y' is called the *Gubinelli derivative* of y, which is uniquely defined as long as x is *truly rough* [13, Definition 6.3 & Proposition 6.4], namely there exists a dense set of instants s of I such that x is "rough at time s", i.e.

$$\forall h^* \in (\mathbb{R}^m)^* \setminus \{0\}: \quad \limsup_{t \downarrow s} \frac{|\langle h^*, x_{s,t}\rangle|}{|t-s|^{2\alpha}} = \infty.$$

For instance, almost all trajectories of a fractional Brownian motion B^H with $H > \frac{1}{3}$ is truly rough [13, Section 6].

Denote by $\mathscr{D}^{2\alpha}_x(I)$ the space of all the couples (y, y') controlled by x, then $\mathscr{D}^{2\alpha}_x(I)$ is a Banach space equipped with the norm

$$\|(y,y')\|_{x,2\alpha,I} := \|y_{\min I}\| + \|y'_{\min I}\| + \left\vert\!\left\vert\!\left\vert (y,y') \right\vert\!\right\vert\!\right\vert_{x,2\alpha,I}, \quad \left\vert\!\left\vert\!\left\vert (y,y') \right\vert\!\right\vert\!\right\vert_{x,2\alpha,I} := \left\vert\!\left\vert\!\left\vert y' \right\vert\!\right\vert\!\right\vert_{\alpha,I} + \left\vert\!\left\vert\!\left\vert R^y \right\vert\!\right\vert\!\right\vert_{2\alpha,I^2},$$

Then for a fixed rough path $\mathbf{x} = (x, \mathbb{X})$ and any controlled rough path $(y, y') \in \mathscr{D}^{2\alpha}_x(I)$, the integral $\int_s^t y_u dx_u$ can be defined as the limit of the Darboux sum

$$\int_s^t y_u dx_u := \lim_{|\Pi| \to 0} \sum_{[u,v] \in \Pi} \Big(y_u \otimes x_{u,v} + y'_u \mathbb{X}_{u,v} \Big)$$

where the limit is taken on all finite partitions Π of I with $|\Pi| := \max_{[u,v]\in\Pi} |v-u|$. Moreover, there exists a constant $C_\alpha = C_{\alpha,|I|} > 1$, such that

$$\Big\| \int_s^t y_u dx_u - y_s \otimes x_{s,t} - y'_s \mathbb{X}_{s,t} \Big\| \le C_\alpha |t-s|^{3\alpha} \Big(\left\vert\!\left\vert\!\left\vert x \right\vert\!\right\vert\!\right\vert_{\alpha,[s,t]} \left\vert\!\left\vert\!\left\vert R^y \right\vert\!\right\vert\!\right\vert_{2\alpha,[s,t]^2} + \left\vert\!\left\vert\!\left\vert y' \right\vert\!\right\vert\!\right\vert_{\alpha,[s,t]} \left\vert\!\left\vert\!\left\vert \mathbb{X} \right\vert\!\right\vert\!\right\vert_{2\alpha,[s,t]^2} \Big). \tag{2.7}$$

In our paper, we often use the p-variation norm

$$\|(y,y')\|_{x,p,I} := \|y_{\min I}\| + \|y'_{\min I}\| + \left\vert\!\left\vert\!\left\vert (y,y') \right\vert\!\right\vert\!\right\vert_{x,p,I}, \quad \left\vert\!\left\vert\!\left\vert (y,y') \right\vert\!\right\vert\!\right\vert_{x,p,I} := \left\vert\!\left\vert\!\left\vert y' \right\vert\!\right\vert\!\right\vert_{p-\mathrm{var},I} + \left\vert\!\left\vert\!\left\vert R^y \right\vert\!\right\vert\!\right\vert_{q-\mathrm{var},I^2},$$

and a similar version to (2.7) under $p-$variation semi-norm as follows

$$\Big\| \int_s^t y_u dx_u - y_s \otimes x_{s,t} - y'_s \mathbb{X}_{s,t} \Big\| \le C_p \Big(\left\vert\!\left\vert\!\left\vert x \right\vert\!\right\vert\!\right\vert_{p-\mathrm{var},[s,t]} \left\vert\!\left\vert\!\left\vert R^y \right\vert\!\right\vert\!\right\vert_{q-\mathrm{var},[s,t]^2} + \left\vert\!\left\vert\!\left\vert y' \right\vert\!\right\vert\!\right\vert_{p-\mathrm{var},[s,t]} \left\vert\!\left\vert\!\left\vert \mathbb{X} \right\vert\!\right\vert\!\right\vert_{q-\mathrm{var},[s,t]^2} \Big), \tag{2.8}$$

with constant $C_p > 1$ independent of $\mathbf{x}$ and (y, y').

2.3 Rough differential equations

The existence and uniqueness theorem for system (1.2) is first proved in [35], where the solution is understood in the sense of Friz-Victoir [14]. By using rough integrals, we would like to interpret the rough differential equation (1.2) by writing it in the integral form

$$y_t = y_{\min I} + \int_{\min I}^t f(y_s)ds + \int_{\min I}^t g(y_s)dx_s, \quad \forall t \in I, \tag{2.9}$$

for any interval I and an initial value $y_{\min I} \in \mathbb{R}^d$, and we search for a solution $y \in \mathscr{D}_x^{2\alpha}(I,\mathbb{R}^d)$. This is possible because for $g : \mathbb{R}^d \to \mathcal{L}(\mathbb{R}^m,\mathbb{R}^d)$ satisfying $(\mathbf{H}_2)$, it is easy to prove (see e.g. [17]) that

$$y \in \mathscr{D}_x^{2\alpha}(I,\mathbb{R}^d) \Rightarrow g(y) \in \mathscr{D}_x^{2\alpha}(I,\mathcal{L}(\mathbb{R}^m,\mathbb{R}^d)), \text{ with } [g(y)]'_s = Dg(y_s)y'_s \in \mathcal{L}(\mathbb{R}^m,\mathcal{L}(\mathbb{R}^m,\mathbb{R}^d)),$$

thus the second integral in (2.9) is well defined.

The existence and uniqueness theorem and the norm estimates for solution of (2.9) are recently proved in [10] under the Lipschitz continuity assumption (1.11), by using the Doss-Sussmann technique [37] and the so-called *greedy sequence of stopping times* in [5]. Namely, for any fixed $\gamma \in (0,1)$ the sequence of greedy times $\{\tau_i(\gamma,\mathbf{x},I)\}_{i\in\mathbb{N}}$ is defined by

$$\tau_0 = \min I, \quad \tau_{i+1} := \inf\Big\{t > \tau_i : \|\mathbf{x}\|_{p-\mathrm{var},[\tau_i,t]} = \gamma\Big\} \wedge \max I. \tag{2.10}$$

Define $N(\gamma,\mathbf{x},I) := \sup\{i \in \mathbb{N} : \tau_i \le \max I\}$, then it is easy to show a rough estimate

$$N(\gamma,\mathbf{x},I) \le 1 + \gamma^{-p}\,\|\mathbf{x}\|^p_{p-\mathrm{var},I}\,. \tag{2.11}$$

Other studies on continuity and properties of stopping times can also be founded in [6, Section 2.2] or [11, Section 4].

Note that from [10, Theorem 3.4], the solution $\phi_\cdot(\mathbf{x},\phi_a)$ of the *pure* rough differential equation

$$d\phi_u = g(\phi_u)dx_u, \quad u \in [a,b], \phi_a \in \mathbb{R}^d \tag{2.12}$$

is C^1 w.r.t. ϕ_a, and $\frac{\partial\phi}{\partial\phi_a}(\cdot,\mathbf{x},\phi_a)$ is the solution of the linearized system

$$d\xi_u = Dg(\phi_u(\mathbf{x},\phi_s))\xi_u dx_u, \quad u \in [a,b], \xi_a = Id, \tag{2.13}$$

where $Id \in \mathbb{R}^{d\times d}$ denotes the identity matrix.

We introduce the semi-norm $\|\kappa,R^\kappa\|_{p-\mathrm{var},[s,t]} := \|\kappa\|_{p-\mathrm{var},[s,t]} + \|R^\kappa\|_{q-\mathrm{var},[s,t]^2}$. The following result shows solution norm estimates for equation (2.12).

Proposition 2.1 *Assume that $\phi_t, \bar\phi_t$ are the solutions of* (2.12)*. Then for any interval $[a,b]$ such that $16C_pC_g\,\|\mathbf{x}\|_{p-\mathrm{var},[a,b]} \le 1$, the following estimates hold*

$$\begin{aligned}
\Big\|\phi, R^\phi\Big\|_{p-\mathrm{var},[a,b]} &\le 8C_pC_g\,\|\mathbf{x}\|_{p-\mathrm{var},[a,b]}\,; &(2.14)\\
\Big\|\bar\phi-\phi, R^{\bar\phi-\phi}\Big\|_{p-\mathrm{var},[a,b]} &\le 16C_pC_g\,\|\mathbf{x}\|_{p-\mathrm{var},[a,b]}\,\|\bar\phi_a-\phi_a\|; &(2.15)\\
\|\bar\phi-\phi\|_{\infty,[a,b]} &\le 2\|\bar\phi_a-\phi_a\|. &(2.16)
\end{aligned}$$

Proof: Because

$$\begin{aligned}
g(\phi_t) - g(\phi_s) &= \int_0^1 Dg(\phi_s+\eta\phi_{s,t})\phi_{s,t}d\eta\\
&= D_g(\phi_s)\phi'_s\otimes x_{s,t} + \int_0^1 Dg(\phi_s+\eta\phi_{s,t})R^\phi_{s,t}d\eta + \int_0^1 [Dg(\phi_s+\eta\phi_{s,t})-Dg(\phi_s)]\phi'_{s,t}\otimes x_{s,t}d\eta,
\end{aligned}$$

it follows that $[g(\phi)]'_s = Dg(\phi_s)g(\phi_s)$, where we use (1.7) to estimate

$$\begin{aligned}\|R^{g(\phi)}_{s,t}\| &\leq \int_0^1 \|Dg(\phi_s+\eta\phi_{s,t})\|\|R^\phi_{s,t}\|d\eta + \int_0^1 \|Dg(\phi_s+\eta\phi_{s,t}) - Dg(\phi_s)\|\|g(\phi_s)\|\|x_{s,t}\|d\eta \\ &\leq C_g\|R^\phi_{s,t}\| + \frac{1}{2}C_g^2\|\phi_{s,t}\|\|x_{s,t}\|.\end{aligned}$$

This together with Hölder inequality yields

$$\begin{aligned}\left\|[g(\phi)]'\right\|_{p-\mathrm{var},[a,b]} &\leq 2C_g^2 \left\|\phi\right\|_{p-\mathrm{var},[a,b]}, \quad \|[g(\phi)]'\|_{\infty,[a,b]} \leq C_g^2, \\ \left\|R^{g(\phi)}\right\|_{q-\mathrm{var},[a,b]^2} &\leq C_g \left\|R^\phi\right\|_{q-\mathrm{var},[a,b]^2} + \frac{1}{2}C_g^2 \left\|x\right\|_{p-\mathrm{var},[a,b]} \left\|\phi\right\|_{p-\mathrm{var},[a,b]}.\end{aligned}$$

Assumption $16C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[a,b]} \leq 1$ follows that $4C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[a,b]} \leq C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[a,b]} < 1$. By applying (2.5) and (2.8), we obtain for any $a \leq s < t \leq b$

$$\begin{aligned}\|\phi_{s,t}\| &\leq \Big\|\int_s^t g(\phi_u)dx_u\Big\| \\ &\leq C_g \left\|x\right\|_{p-\mathrm{var},[s,t]} + C_g^2 \left\|\mathbb{X}\right\|_{q-\mathrm{var},[s,t]^2} \\ &\quad + C_p\Big\{ \left\|x\right\|_{p-\mathrm{var},[s,t]} \left\|R^{g(\phi)}\right\|_{q-\mathrm{var},[s,t]^2} + \left\|\mathbb{X}\right\|_{q-\mathrm{var},[s,t]^2} \left\|[g(\phi)]'\right\|_{p-\mathrm{var},[a,b]} \Big\} \\ &\leq 2\Big\{C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \vee 4C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} \Big\}\Big(1 + C_p \left\|\phi, R^\phi\right\|_{p-\mathrm{var},[s,t]}\Big) \\ &\leq 2C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \Big(1 + \left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]}\Big),\end{aligned}$$

which, by the definition of p-variation seminorm and (2.1), derives

$$\begin{aligned}\left\|\phi\right\|_{p-\mathrm{var},[a,b]} &\leq 2C_pC_g\Big\{\sup_{\Pi[a,b]} \sum_{[s,t]\in\Pi([a,b])} \left\|\mathbf{x}\right\|^p_{p-\mathrm{var},[s,t]}\Big\}^{\frac{1}{p}}\Big(1 + \left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]}\Big) \\ &\leq 2C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[a,b]} \Big(1 + \left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]}\Big).\end{aligned}$$

The same estimate for R^ϕ is actually included in the above estimate, hence

$$\begin{aligned}\left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]} &\leq 4C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[a,b]} \Big(1 + \left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]}\Big) \\ &\leq 4C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[a,b]} + \frac{1}{2}\left\|\phi, R^\phi\right\|_{p-\mathrm{var},[a,b]},\end{aligned}$$

which deduces (2.14).

Next, for any two solutions $\phi_t(\mathbf{x},\phi_a)$ and $\bar\phi_t(\mathbf{x},\bar\phi_a)$ of (2.12), consider their difference $\bar\phi_t - \phi_t$, which satisfies the rough differential equation $d(\bar\phi_t - \phi_t) = [g(\bar\phi_t) - g(\phi_t)]dx_t$. Because

$$\begin{aligned}& g(\bar\phi_t) - g(\phi_t) - g(\bar\phi_s) + g(\phi_s) \\ =\ & \Big[Dg(\bar\phi_s)g(\bar\phi_s) - Dg(\phi_s)g(\phi_s)\Big] \otimes x_{s,t} \\ & + \int_0^1 \Big\{Dg(\bar\phi_s+\eta\bar\phi_{s,t})R^{\bar\phi-\phi}_{s,t} + \Big[Dg(\bar\phi_s+\eta\bar\phi_{s,t}) - Dg(\phi_s+\eta\phi_{s,t})\Big]R^\phi_{s,t}\Big\}d\eta \\ & + \int_0^1 \Big[Dg(\bar\phi_s+\eta\bar\phi_{s,t}) - Dg(\bar\phi_s)\Big]\Big[g(\bar\phi_s) - g(\phi_s)\Big] \otimes x_{s,t}d\eta \\ & + \Big(\int_0^1\int_0^1 D^2g(\bar\phi_s+\mu\eta\bar\phi_{s,t})\eta(\bar\phi_{s,t} - \phi_{s,t})d\mu d\eta\Big)g(\phi_s) \otimes x_{s,t}\end{aligned}$$

$$+\Big(\int_0^1\int_0^1\Big[D^2g(\bar\phi_s+\mu\eta\bar\phi_{s,t})-D^2g(\phi_s+\mu\eta\phi_{s,t})\eta\phi_{s,t}\Big]d\mu d\eta\Big)g(\phi_s)\otimes x_{s,t}.$$

it follows that $[g(\bar\phi)-g(\phi)]'_s = Dg(\bar\phi_s)g(\bar\phi_s)-Dg(\phi_s)g(\phi_s)$ which has the form $Q(\bar\phi_s)-Q(\phi_s)$. Notice that $\|Q(\bar\phi_s)-Q(\phi_s)\|\leq 2C_g^2\|\bar\phi_s-\phi_s\|$ and

$$\begin{aligned}\left\|\!\left|Q(\bar\phi)-Q(\phi)\right|\!\right\|_{p-{\rm var},[s,t]} &\leq C_Q\Big(\left\|\!\left|\bar\phi-\phi\right|\!\right\|_{p-{\rm var},[s,t]}+\|\bar\phi-\phi\|_{\infty,[s,t]}\left\|\!\left|\phi\right|\!\right\|_{p-{\rm var},[s,t]}\Big)\\ &\leq 2C_g^2\Big(\left\|\!\left|\bar\phi-\phi\right|\!\right\|_{p-{\rm var},[s,t]}+\|\bar\phi-\phi\|_{\infty,[s,t]}\left\|\!\left|\phi\right|\!\right\|_{p-{\rm var},[s,t]}\Big).\end{aligned}$$

On the other hand

$$\begin{aligned}&\left\|\!\left|R^{g(\bar\phi)-g(\phi)}\right|\!\right\|_{q-{\rm var},[s,t]^2}\\ \leq\ & C_g\left\|\!\left|R^{\bar\phi-\phi}\right|\!\right\|_{q-{\rm var},[s,t]^2}+C_g\|\bar\phi-\phi\|_{\infty,[s,t]}\left\|\!\left|R^{\phi}\right|\!\right\|_{q-{\rm var},[s,t]^2}\\ &+\frac12 C_g^2\left\|\!\left|x\right|\!\right\|_{p-{\rm var},[s,t]}\Big[\left\|\!\left|\bar\phi-\phi\right|\!\right\|_{p-{\rm var},[s,t]}+\|\bar\phi-\phi\|_\infty\Big(\left\|\!\left|\bar\phi\right|\!\right\|_{p-{\rm var},[s,t]}+\left\|\!\left|\phi\right|\!\right\|_{p-{\rm var},[s,t]}\Big)\Big].\end{aligned}$$

This leads to the estimate

$$\begin{aligned}&\|\bar\phi_{s,t}-\phi_{s,t}\|\leq\Big\|\int_s^t[g(\bar\phi_u)-g(\phi_u)]dx_u\Big\|\\ \leq\ & C_g\|\bar\phi_s-\phi_s\|\left\|\!\left|x\right|\!\right\|_{p-{\rm var},[s,t]}+2C_g^2\|\bar\phi_s-\phi_s\|\left\|\!\left|\mathbb{X}\right|\!\right\|_{q-{\rm var},[s,t]^2}\\ &+C_p\Big\{\left\|\!\left|x\right|\!\right\|_{p-{\rm var},[s,t]}\left\|\!\left|R^{g(\bar\phi)-g(\phi)}\right|\!\right\|_{q-{\rm var},[s,t]^2}+\left\|\!\left|\mathbb{X}\right|\!\right\|_{q-{\rm var},[s,t]^2}\left\|\!\left|[g(\bar\phi)-g(\phi)]'\right|\!\right\|_{p-{\rm var},[s,t]}\Big\},\end{aligned}\tag{2.17}$$

which yields

$$\begin{aligned}\left\|\!\left|\bar\phi-\phi\right|\!\right\|_{p-{\rm var},[a,b]} &\leq 2C_p\Big\{C_g\left\|\!\left|\mathbf{x}\right|\!\right\|_{p-{\rm var},[a,b]}\vee 4C_g^2\left\|\!\left|\mathbf{x}\right|\!\right\|^2_{p-{\rm var},[a,b]}\Big\}\times\\ &\times\Big(1+\left\|\!\left|\bar\phi,R^{\bar\phi}\right|\!\right\|_{p-{\rm var},[a,b]}+\left\|\!\left|\phi,R^{\phi}\right|\!\right\|_{p-{\rm var},[a,b]}\Big)\times\\ &\times\Big(\|\bar\phi_a-\phi_a\|+\left\|\!\left|\bar\phi-\phi,R^{\bar\phi-\phi}\right|\!\right\|_{p-{\rm var},[a,b]}\Big).\end{aligned}\tag{2.18}$$

The similar estimate for $\left\|\!\left|R^{\bar\phi-\phi}\right|\!\right\|_{q-{\rm var},[a,b]}$ is already included in the estimate (2.17), hence

$$\begin{aligned}\left\|\!\left|\bar\phi-\phi,R^{\bar\phi-\phi}\right|\!\right\|_{p-{\rm var},[a,b]} &\leq 4C_p\Big\{C_g\left\|\!\left|\mathbf{x}\right|\!\right\|_{p-{\rm var},[a,b]}\vee C_g^2\left\|\!\left|\mathbf{x}\right|\!\right\|^2_{p-{\rm var},[a,b]}\Big\}\times\\ &\times\Big(1+\left\|\!\left|\bar\phi,R^{\bar\phi}\right|\!\right\|_{p-{\rm var},[a,b]}+\left\|\!\left|\phi,R^{\phi}\right|\!\right\|_{p-{\rm var},[a,b]}\Big)\times\\ &\times\Big(\|\bar\phi_a-\phi_a\|+\left\|\!\left|\bar\phi-\phi,R^{\bar\phi-\phi}\right|\!\right\|_{p-{\rm var},[a,b]}\Big),\end{aligned}$$

which, together with (2.14), leads to (2.15) and (2.16). □

Since $\phi_\cdot(\mathbf{x},\phi_a)$ is C^1 w.r.t. ϕ_a [10, Theorem 3.4], by dividing both sides of (2.15) by $\|\bar\phi_a-\phi_a\|$ and then letting $\bar\phi_a-\phi_a$ to zero, we obtain

$$\Big\|\frac{\partial\phi}{\partial\phi_a}(t,\mathbf{x},\phi_a)-Id\Big\|\leq\left\|\!\left|\frac{\partial\phi}{\partial\phi_a}(\cdot,\mathbf{x},\phi_a),R^{\frac{\partial\phi}{\partial\phi_a}(\cdot,\mathbf{x},\phi_a)}\right|\!\right\|_{p-{\rm var},[a,b]}\leq 16C_pC_g\left\|\!\left|\mathbf{x}\right|\!\right\|_{p-{\rm var},[a,b]}.\tag{2.19}$$

Note that (2.14), (2.15) still hold for the backward equation

$$h_b=h_t+\int_t^b g(h_u)dx_u,\quad \forall t\in[a,b],\tag{2.20}$$

thus (2.19) still holds if $\frac{\partial \phi}{\partial \phi_s}(t,\mathbf{x},\phi_s)$ is replaced by $\Big[\frac{\partial \phi}{\partial \phi_s}(t,\mathbf{x},\phi_s)\Big]^{-1}$, which is also the linearization of the solution of (2.20) (see e.g. [10, Corollary 3.5, Theorem 3.7]).

As shown in [35, Theorem 4.3], the local Lipschitz continuity, the one-sided Lipschitz continuity and (1.6) for f are enough to prove the existence and uniqueness of solution of rough equation (1.2). Here we need to go one more step to prove the solution estimate of (1.2), under condition $(\mathbf{H}_f)$.

Theorem 2.2 *Under the assumptions ($\boldsymbol{H}_f$), ($\boldsymbol{H}_g$), ($\boldsymbol{H}_X$), there exists a solution of* (1.2) *on any interval* $[0,T]$. *Moreover, for any* $\lambda > 0$ *small enough, there exist constants* $\delta_\lambda, C_\lambda > 0$ *such that the following estimates hold*

$$\|y_t\| \leq \|y_0\| e^{-\delta_\lambda t} + C_\lambda N\Big(\frac{\lambda}{16 C_p C_g}, \mathbf{x}, [0,t]\Big), \quad \forall t\in[0,T]. \tag{2.21}$$

Proof: The idea is to prove the existence and uniqueness of the solution on each small interval between two consecutive stopping times, and then concatenate to obtain the conclusion on any interval. The Doss-Sussmann technique used in [35] and [10, Theorem 3.7] ensures that, by a transformation $y_t = \phi_t(\mathbf{x}, z_t)$ there is an one-one correspondence between a solution y_t of (1.2) on a certain interval $[0,\tau]$ and a solution z_t of the associate ordinary differential equation

$$\dot{z}_t = \Big[\frac{\partial \phi}{\partial z}(t,\mathbf{x},z_t)\Big]^{-1} f(\phi_t(\mathbf{x},z_t)), \quad t\in[0,\tau],\ z_0 = y_0. \tag{2.22}$$

Since f is locally Lipschitz continuous, there exists a unique solution for (2.22) on some local interval τ_{local}. To estimate the solution norm growth, assign $\gamma_t := y_t - z_t$ and $\psi_t := \Big[\frac{\partial \phi}{\partial z}(t,\mathbf{x},z_t)\Big]^{-1} - Id$ for $t\in[0,\tau\wedge\tau_{local}]$, where $\tau>0$ is chosen such that $16 C_p C_g \left\| \mathbf{x} \right\|_{p-\mathrm{var},[0,\tau]} \leq \lambda$ for some $\lambda \in (0,1)$ small enough (which will be specified later). With such τ, it then follows from Proposition 2.1 and (2.19) that

$$\|\gamma_t\| = \|\phi_t(\mathbf{x},z_t) - z_t\| \leq \frac{\lambda}{2} \quad \text{and} \quad \|\psi_t\| \leq \lambda, \quad \forall t\in[0,\tau\wedge\tau_{local}]. \tag{2.23}$$

To estimate $\|z_t\|$, we rewrite (2.22) as

$$\dot{z}_t = (Id + \psi_t) f(z_t + \gamma_t). \tag{2.24}$$

First, we are going to prove that there exists constants $\bar{C}_\lambda, \delta_\lambda > 0$ such that

$$\frac{d}{2dt}\|z_t\|^2 \leq \bar{C}_\lambda - \delta_\lambda \|z_t\|^2. \tag{2.25}$$

Indeed, consider two cases.

Case 1: $z_t + \gamma_t \neq 0$. From assumption $(\mathbf{H}_f)$ and condition (1.10), we can check that

$$\begin{aligned}
\frac{d}{2dt}\|z_t\|^2 &= \Big\langle z_t, (Id+\psi_t)\Big[\frac{\langle z_t+\gamma_t, f(z_t+\gamma_t)\rangle}{\|z_t+\gamma_t\|^2}(z_t+\gamma_t) + \pi^\perp_{z_t+\gamma_t}(f(z_t+\gamma_t))\Big]\Big\rangle \\
&= \underbrace{\Big\langle z_t, (Id+\psi_t)\frac{(z_t+\gamma_t)}{\|z_t+\gamma_t\|}\Big\rangle}_{=:M_1} \underbrace{\Big\langle \frac{z_t+\gamma_t}{\|z_t+\gamma_t\|}, f(z_t+\gamma_t)\Big\rangle}_{=:M_2} + \underbrace{\Big\langle z_t, (Id+\psi_t)\pi^\perp_{z_t+\gamma_t}(f(z_t+\gamma_t))\Big\rangle}_{=:M_3}.
\end{aligned} \tag{2.26}$$

Observe that from (1.5) and (2.23),

$$M_1 \;\leq\; (1+\|\psi_t\|)\|z_t\| \leq (1+\lambda)\|z_t\|; \tag{2.27}$$

$$\begin{aligned} M_1 &\geq \left\langle z_t, \frac{z_t+\gamma_t}{\|z_t+\gamma_t\|}\right\rangle - \|\psi_t\|\|z_t\| \geq \|z_t+\gamma_t\| - \|\gamma_t\| - \|\psi_t\|\|z_t\| \geq (1-\lambda)\|z_t\| - \lambda; \quad (2.28)\\ M_2 &\leq D_1 - D_2\|z_t+\gamma_t\| \leq D_1 + D_2\lambda - D_2\|z_t\|. \quad (2.29)\end{aligned}$$

As a result, (2.27) deduces

$$M_1M_2 \leq (1+\lambda)\|z_t\|M_2 \qquad \text{if} \qquad M_2 \geq 0, \tag{2.30}$$

while (2.28) follows

$$M_1M_2 \leq \Big[(1-\lambda)\|z_t\| - \lambda\Big]M_2 \qquad \text{if} \qquad M_2 < 0. \tag{2.31}$$

If $M_2 \geq 0$ then (2.30) and (2.29) lead to

$$M_1M_2 \leq (1+\lambda)\|z_t\|\Big[D_1 + D_2\lambda - D_2\|z_t\|\Big]. \tag{2.32}$$

If $M_2 < 0$ and $(1-\lambda)\|z_t\| - \lambda \geq 0$, then (2.31) and (2.29) yield

$$M_1M_2 \leq \Big[(1-\lambda)\|z_t\| - \lambda\Big]\Big[D_1 + D_2\lambda - D_2\|z_t\|\Big]. \tag{2.33}$$

If $M_2 < 0$ and $(1-\lambda)\|z_t\| - \lambda < 0$, then $\|z_t\| \leq \frac{\lambda}{1-\lambda}$ and $\|z_t+\gamma_t\| \leq \|z_t\| + \|\gamma_t\| \leq \frac{\lambda}{1-\lambda} + \lambda$. In this case (2.31) and (2.29) deduce

$$\begin{aligned} M_1M_2 &\leq (1-\lambda)\|z_t\|M_2 + \lambda|M_2| \leq (1-\lambda)\|z_t\|\Big[D_1 + D_2\lambda - D_2\|z_t\|\Big] + \lambda\|f(z_t+\gamma_t)\|\\ &\leq (1-\lambda)\|z_t\|\Big[D_1 + D_2\lambda - D_2\|z_t\|\Big] + \lambda\max\Big\{\|f(\xi)\| : \|\xi\| \leq \frac{\lambda}{1-\lambda} + \lambda\Big\}. \quad (2.34)\end{aligned}$$

Combining all these three cases (2.32), (2.33), (2.34) and applying Cauchy inequality, we can show that there exists a generic constant $\bar{C}_\lambda > 0$ such that

$$M_1M_2 \leq \bar{C}_\lambda - \frac{D_2}{2}(1-\lambda)\|z_t\|^2.$$

On the other hand,

$$\begin{aligned} M_3 &= \Big\langle z_t+\gamma_t, \pi^\perp_{z_t+h_t}(f(z_t+\gamma_t))\Big\rangle - \Big\langle \gamma_t, \pi^\perp_{z_t+h_t}(f(z_t+\gamma_t))\Big\rangle + \Big\langle z_t, \psi_t\pi^\perp_{z_t+h_t}(f(z_t+\gamma_t))\Big\rangle\\ &= -\Big\langle \gamma_t, \pi^\perp_{z_t+h_t}(f(z_t+\gamma_t))\Big\rangle + \Big\langle z_t, \psi_t\pi^\perp_{z_t+h_t}(f(z_t+\gamma_t))\Big\rangle\\ &\leq (\|\gamma_t\| + \|\psi_t\|\|z_t\|)C_f(1+\|z_t+\gamma_t\|) \leq \bar{C}_\lambda + 2C_f\lambda\|z_t\|^2,\end{aligned}$$

for some generic $\bar{C}_\lambda$. As a result, there exists a generic constant $\bar{C}_\lambda$ such that

$$\frac{d}{2dt}\|z_t\|^2 \quad \leq \quad \bar{C}_\lambda + \Big[2C_f\lambda - \frac{D_2}{2}(1-\lambda)\Big]\|z_t\|^2. \tag{2.35}$$

Case 2: $z_t + \gamma_t = 0$. Then the same arguments show that

$$\begin{aligned} \frac{d}{2dt}\|z_t\|^2 &= \langle z_t+\gamma_t, f(z_t+\gamma_t)\rangle - \langle \gamma_t, f(z_t+\gamma_t)\rangle + \langle z_t, \psi_t f(z_t+\gamma_t)\rangle\\ &= \Big\langle z_t+\gamma_t, f(z_t+\gamma_t)\Big\rangle - \langle \gamma_t, f(0)\rangle + \langle z_t, \psi_t f(0)\rangle\\ &\leq D_1\|z_t+\gamma_t\| - D_2\|z_t+\gamma_t\|^2 + (\|\gamma_t\| + \|\psi_t\|\|z_t\|)\|f(0)\|\\ &\leq \bar{C}_\lambda + \Big[2C_f\lambda - \frac{D_2}{2}(1-\lambda)\Big]\|z_t\|^2,\end{aligned}$$

where one can apply Cauchy inequality to obtain the last inequality for some generic constant $\bar{C}_\lambda$. Hence (2.35) holds for all $z_t \in \mathbb{R}^d$ where $t \in [0, \tau \wedge \tau_{local}]$, with a generic constant $\bar{C}_\lambda$ and a sufficiently small $\lambda < 1$. This proves (2.25) by choosing

$$\delta_\lambda := \frac{D_2}{2}(1-\lambda) - 2C_f\lambda > 0 \quad \text{for} \quad 0 < \lambda < \frac{D_2}{D_2 + 4C_f} < 1.$$

Next, (2.25) implies that $\|z_t\|$ is bounded by $\sqrt{\frac{\bar{C}_\lambda}{\delta_\lambda}} + \|z_0\| = \sqrt{\frac{\bar{C}_\lambda}{\delta_\lambda}} + \|y_0\|$ as long as $t \in [0, \tau \wedge \tau_{local}]$, thereby proving the existence and uniqueness of the solution z_t of equation (2.22) on $[0, \tau \wedge \tau_{local}]$, and so is the solution y_t of (1.2) on $[0, \tau \wedge \tau_{local}]$. In addition, whenever $\tau > \tau_{local}$ then (2.23) is satisfied and the above arguments can be applied to prove the existence and uniqueness of the solution by concatenation, until the interval $[0, \tau]$ is fully covered.

Finally, with such $\lambda > 0$, construct a greedy sequence of stopping times $\{\tau_i(\frac{\lambda}{16C_pC_g}, \mathbf{x}, [0,t])\}$. On each interval $[\tau_i, \tau_{i+1}]$ it is similar to prove the existence and uniqueness of the solution of the two differential equations (1.2) and (2.22) with the shifted time

$$\begin{aligned} dy_{t+\tau_i} &= f(y_{t+\tau_i})dt + g(y_{t+\tau_i})dx_{t+\tau_i}, \quad \forall t \in [0, \tau_{i+1} - \tau_i]; \\ \dot{z}_{t+\tau_i} &= \Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}_{\cdot+\tau_i}, z_{t+\tau_i})\Big]^{-1} f(\phi_t(\mathbf{x}_{\cdot+\tau_i}, z_{t+\tau_i})), \quad \forall t \in [0, \tau_{i+1} - \tau_i]. \end{aligned}$$

As a result, the existence and uniqueness of the solution of the two systems (1.2) and (2.22) on $[0, T]$ is proved by concatenation. To estimate the solution norm, observe from (2.35) that

$$\begin{aligned} \|z_t\| &\leq \sqrt{\frac{\bar{C}_\lambda}{\delta_\lambda}} + \|z_{\tau_i}\| \exp\Big\{-\delta_\lambda(t-\tau_i)\Big\}, \quad \forall t \in [\tau_i, \tau_{i+1}], i \in \mathbb{N}. \quad \text{In particular,} \\ \|y_{\tau_{i+1}}\| &\leq \frac{\lambda}{2} + \sqrt{\frac{\bar{C}_\lambda}{\delta_\lambda}} + \|y_{\tau_i}\| \exp\Big\{-\delta_\lambda(\tau_{i+1}-\tau_i)\Big\}, \quad \forall i \in \mathbb{N}. \end{aligned}$$

Assign $C_\lambda := \frac{\lambda}{2} + \sqrt{\frac{\bar{C}_\lambda}{\delta_\lambda}}$. By induction, one can easily show that

$$\|y_{\tau_i}\| \leq \|y_0\| \exp\Big\{-\delta_\lambda \tau_i\Big\} + iC_\lambda, \quad \forall i \in \mathbb{N}.$$

By the definition of stopping times (2.10), $\tau_{N(\frac{\lambda}{16C_pC_g}, \mathbf{x}, [0,t])} = t$, which deduces (2.21). □

3 Random attractors

3.1 Generation of random dynamical systems

In this subsection we would like to present the generation of a random dynamical system from rough differential equation (1.2), which is based mainly on the work in [2] with only a small modification for Hölder spaces. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a so-called measurable *metric dynamical system* $\theta : \mathbb{R} \times \Omega \to \Omega$ such that $\theta_t : \Omega \to \Omega$ is $\mathbb{P}-$ preserving, i.e $\mathbb{P}(B) = \mathbb{P}(\theta_t^{-1}(B))$ for all $B \in \mathcal{F}, t \in \mathbb{R}$, and $\theta_{t+s} = \theta_t \circ \theta_s$ for all $t, s \in \mathbb{R}$. A continuous *random dynamical system* $\varphi : \mathbb{R} \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$, $(t, \omega, y_0) \mapsto \varphi(t, \omega)y_0$ is then defined as a measurable mapping which is also continuous in t and y_0 such that the cocycle property

$$\varphi(t+s, \omega)y_0 = \varphi(t, \theta_s\omega) \circ \varphi(s, \omega)y_0, \quad \forall t, s \in \mathbb{R}, \omega \in \Omega, y_0 \in \mathbb{R}^d \tag{3.1}$$

is satisfied [1].

In our setting, denote by $T_1^2(\mathbb{R}^m) = 1 \oplus \mathbb{R}^m \oplus (\mathbb{R}^m \otimes \mathbb{R}^m)$ the set with the tensor product

$$(1, g^1, g^2) \otimes (1, h^1, h^2) = (1, g^1 + h^1, g^1 \otimes h^1 + g^2 + h^2),$$

for all $\mathbf{g} = (1, g^1, g^2), \mathbf{h} = (1, h^1, h^2) \in T_1^2(\mathbb{R}^m)$. Then it can be shown that $(T_1^2(\mathbb{R}^m), \otimes)$ is a topological group with unit element $\mathbf{1} = (1, 0, 0)$ and $\mathbf{g}^{-1} = (1, -g^1, g^1 \otimes g^1 - g^2)$.

Given $\alpha \in (\frac{1}{3}, \nu)$, denote by $\mathscr{C}^{0,\alpha}(I, T_1^2(\mathbb{R}^m))$ the closure of $\mathscr{C}^\infty(I, T_1^2(\mathbb{R}^m))$ in the Hölder space $\mathscr{C}^\alpha(I, T_1^2(\mathbb{R}^m))$, and by $\mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m))$ the space of all paths $\mathbf{g} : \mathbb{R} \to T_1^2(\mathbb{R}^m))$ such that $\mathbf{g}|_I \in \mathscr{C}^{0,\alpha}(I, T_1^2(\mathbb{R}^m))$ for each compact interval $I \subset \mathbb{R}$ containing 0. Then $\mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m))$ is equipped with the compact open topology given by the $\alpha-$ Hölder norm (2.2), i.e the topology generated by the metric

$$d_\alpha(\mathbf{g}, \mathbf{h}) := \sum_{k \geq 1} \frac{1}{2^k} (\|\mathbf{g} - \mathbf{h}\|_{\alpha, [-k,k]} \wedge 1).$$

As a result, it is separable and thus a Polish space.

Let us consider a stochastic process $\bar{\mathbf{X}}$ defined on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ with realizations in $(\mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m)), \mathcal{F})$. Assume further that $\bar{\mathbf{X}}$ has stationary increments. Assign $\Omega := \mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m))$ and equip it with the Borel σ - algebra $\mathcal{F}$ and let $\mathbb{P}$ be the law of $\bar{\mathbf{X}}$. Denote by θ the *Wiener-type shift*

$$(\theta_t \omega)_\cdot = \omega_t^{-1} \otimes \omega_{t+\cdot}, \forall t \in \mathbb{R}, \omega \in \mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m)), \tag{3.2}$$

and define the so-called *diagonal process* $\mathbf{X} : \mathbb{R} \times \Omega \to T_1^2(\mathbb{R}^m), \mathbf{X}_t(\omega) = \omega_t$ for all $t \in \mathbb{R}, \omega \in \Omega$. Due to the stationarity of $\bar{\mathbf{X}}$, it can be proved that θ is invariant under $\mathbb{P}$, then forming a continuous (and thus measurable) dynamical system on $(\Omega, \mathcal{F}, \mathbb{P})$ [2, Theorem 5]. Moreover, $\mathbf{X}$ forms an α - *rough path cocycle*, namely, $\mathbf{X}_\cdot(\omega) \in \mathscr{C}_0^{0,\alpha}(\mathbb{R}, T_1^2(\mathbb{R}^m))$ for every $\omega \in \Omega$, which satisfies the *cocyle relation*:

$$\mathbf{X}_{t+s}(\omega) = \mathbf{X}_s(\omega) \otimes \mathbf{X}_t(\theta_s \omega), \forall \omega \in \Omega, t, s \in \mathbb{R},$$

in the sense that $\mathbf{X}_{s,s+t} = \mathbf{X}_t(\theta_s \omega)$ with the increment notation $\mathbf{X}_{s,s+t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_{s+t}$. It is important to note that the two-parameter flow property

$$\mathbf{X}_{s,u} \otimes \mathbf{X}_{u,t} = \mathbf{X}_{s,t}, \forall s, t \in \mathbb{R}$$

is equivalent to the fact that $\mathbf{X}_t(\omega) = (1, \mathbf{x}_t(\omega)) = (1, x_t(\omega), \mathbb{X}_{0,t}(\omega))$, where $x_\cdot(\omega) : \mathbb{R} \to \mathbb{R}^m$ and $\mathbb{X}_{\cdot,\cdot}(\omega) : I^2 \to \mathbb{R}^m \otimes \mathbb{R}^m$ are random funtions satisfying Chen's relation relation (2.3).

To fulfill the Hölder continuity of almost all realizations, assume condition (1.8) that the estimate $\mathbb{E}\Big(\|x_{s,t}\|^p + \|\mathbb{X}_{s,t}\|^q\Big) \leq C_{T,\nu}|t-s|^{p\nu}$ holds for all $s, t \in [0,T]$ and any interval $[0,T]$, with $p\nu \geq 1, q = \frac{p}{2}$ and some constant $C_{T,\nu}$. Then due to the Kolmogorov criterion for rough paths [14, Appendix A.3], for any $\alpha \in (\frac{1}{3}, \nu)$ and $p = \frac{1}{\alpha}$, there exists a version of $\omega-$wise $(x, \mathbb{X})$ and random variables $K_\beta \in L^p, \mathbb{K}_\beta \in L^q$, such that, $\omega-$wise speaking and an abuse of notation, $\|x_{s,t}\| \leq K_\alpha |t-s|^\alpha$, $\|\mathbb{X}_{s,t}\| \leq \mathbb{K}_\alpha |t-s|^{2\alpha}$, for all $s, t \in [0,T]$, so that $\mathbf{x} = (x, \mathbb{X}) \in \mathscr{C}^\alpha(I)$. Moreover, we could modify α such that

$$x \in C^{0,\alpha}(I) := \{x \in C^\alpha(I) : \lim_{\Delta \to 0} \sup_{0<t-s<\Delta} \frac{\|x_{s,t}\|}{|t-s|^\alpha} = 0\},$$

$$\mathbb{X} \in C^{0,2\alpha}(I^2) := \{\mathbb{X} \in C^{2\alpha}(I^2) : \lim_{\Delta \to 0} \sup_{0<t-s<\Delta} \frac{\|\mathbb{X}_{s,t}\|}{|t-s|^{2\alpha}} = 0\},$$

thus $\mathscr{C}^{0,\alpha}(I) \subset C^{0,\alpha}(I) \oplus C^{0,2\alpha}(I^2)$ is separable due to the separability of $C^{0,\alpha}(I)$ and $C^{0,2\alpha}(I^2)$. In particular, the Wiener shift (3.2) implies that

$$\|\mathbf{x}(\theta_h \omega)\|_{p-\mathrm{var},[s,t]} = \|\mathbf{x}(\omega)\|_{p-\mathrm{var},[s+h,t+h]}, \quad N_{[s,t]}(\mathbf{x}(\theta_h \omega)) = N_{[s+h,t+h]}(\mathbf{x}(\omega)). \tag{3.3}$$

Remark 3.1 Due to [2, Corollary 9], the above construction is possible for X_t to be a continuous, centered Gaussian process with stationary increments and independent components, satisfying: there exists for any $T>0$ a constant C_T such that for all $p\geq \frac{1}{\nu}$, $\mathbb{E}\|X_t-X_s\|^p \leq C_T|t-s|^{p\nu}$ for all $s,t\in[0,T]$. Then $\mathbf{X}$ can be chosen to be the natural lift of X in the sense of Friz-Victoir [14, Chapter 15] with sample paths in the space $\mathscr{C}^{0,\alpha}_0(\mathbb{R},T^2_1(\mathbb{R}^m))$, for a certain $\alpha\in(0,\nu)$.

For example, consider X to be a m - dimensional fractional Brownian motion B^H with independent components [32] and Hurst exponent $H\in(\frac{1}{3},1)$, i.e. a family of $B^H=\{B^H_t\}_{t\in\mathbb{R}}$ with continuous sample paths and $\mathbb{E}[B^H_t B^H_s]=\frac{1}{2}\Big(t^{2H}+s^{2H}-|t-s|^{2H}\Big)I^{m\times m}$, for all $t,s\in\mathbb{R}_+$. Given a fixed interval $[0,T]$, the covariance of increments of fractional Brownian motions $R:[0,T]^4\to\mathbb{R}^{m\times m}$, defined by $R\begin{pmatrix} s & t\\ s' & t'\end{pmatrix}:=E(B^H_{s,t}B^H_{s',t'})$ is of finite $\varrho-$ variation norm for $\varrho=\frac{1}{2H}$, i.e.

$$\|R\|_{I\times I',\varrho}:=\Big\{\sup_{\Pi(I),\Pi'(I')}\sum_{\substack{[s,t]\in\Pi(I),\\ [s',t']\in\Pi(I')}}\Big|R\begin{pmatrix} s & t\\ s' & t'\end{pmatrix}\Big|^\varrho\Big\}^{\frac{1}{\varrho}}<\infty, \tag{3.4}$$

and there exists a constant $M_{\varrho,T}$ such that $\|R\|_{[s,t]^2,\varrho}\leq M_{\varrho,T}|t-s|^{\frac{1}{\varrho}}, \quad \forall t,s\in[0,T]$.

Then one can prove that the integral

$$\mathbb{X}^{i,j}_{s,t}=\lim_{|\Pi|\to 0}\int_\Pi X^i_{s,r}dX^j_r=\lim_{|\Pi|\to 0}\sum_{[u,v]\in\Pi}X^i_{s,u}X^j_{u,v}\quad \text{in } \mathcal{L}^2\text{- sense},\quad \forall s,t\in[0,T],$$

is well-defined regardless of the chosen partition Π of $[s,t]$; in addition $\mathbb{X}^{i,i}_{s,t}=\frac{1}{2}(X^i_{s,t})^2$ and $\mathbb{X}^{i,j}_{s,t}+\mathbb{X}^{j,i}_{s,t}=X^i_{s,t}X^j_{s,t}$. Furthermore, for $\frac{1}{p}<\nu<\frac{1}{2\varrho}=H$, there exist constants $C(p,\varrho,m,T), C(p,\varrho,m,T,\nu)$ such that

$$\begin{aligned}\mathbb{E}\Big[\|X_{s,t}\|^p+\|\mathbb{X}_{s,t}\|^q\Big] &\leq C(p,\varrho,m,T)|t-s|^{pH},\quad \forall s,t\in[0,T]\quad\text{and}\\ \mathbb{E}\Big[\,|\!|\!|X|\!|\!|^p_{\nu,[0,T]}+|\!|\!|\mathbb{X}|\!|\!|^q_{2\nu,[0,T]}\Big] &\leq C(p,\varrho,m,T,\nu)M^q_{\varrho,T}.\end{aligned}$$

Therefore for $\frac{1}{3}<\alpha<\nu<H$, almost sure all realizations $\mathbf{x}=(X,\mathbb{X})$ belong to the set $\mathscr{C}^{0,\alpha}([0,T])$ and satisfy Chen's relation (2.3) and satisfy condition (1.8).

We reformulate the conclusion in [2, Theorem 21] in our scenarios as follows.

Proposition 3.2 *Given the measurable metric dynamical system $(\Omega,\mathscr{F},\mathbb{P},\theta)$ and the p- rough cocycle $\mathbf{X}:\mathbb{R}\times\Omega\to T^2_1(\mathbb{R}^m)$ as above, the system (1.1) generates a continuous random dynamical system φ over $(\Omega,\mathscr{F},\mathbb{P},\theta)$, such that for any $[0,T]$ and all $\omega\in\Omega$, $\varphi(t,\omega)y_0$ is the unique solution (in the Gubinelli sense) of (1.2), which is understood in the pathwise integral form (2.9) on [0,T], where $\mathbf{x}=(x,\mathbb{X})$ is the projection of $\mathbf{X}.(\omega)$ on $\mathbb{R}^m\oplus(\mathbb{R}^m\otimes\mathbb{R}^m)$.*

Proof: We present here only a sketch of the proof. Fix a realization $\omega\in\Omega$ of the diagonal process $\mathbf{X}$, then $\omega_t=\mathbf{X}_t(\omega)=(1,x_t(\omega),\mathbb{X}_{0,t}(\omega))$. Since $\mathbf{X}$ is a rough cocycle, the shift property (3.2) yields

$$\Big(1,x_{u,v}(\omega),\mathbb{X}_{u,v}(\omega)\Big)=\omega_u^{-1}\otimes\omega_v=(\theta_u\omega)_{v-u},\quad \forall 0\leq s\leq t. \tag{3.5}$$

We therefore can rewrite the definition of the rough integral as

$$\int_a^b y_u d\omega_u:=\lim_{|\Pi|\to 0}\sum_\Pi\underbrace{\Big(y_u\otimes x_{u,v}(\omega)+y'_u\mathbb{X}_{u,v}(\omega)\Big)}_{=:(y_u,y'_u)\bar{\otimes}\Big(1,x_{u,v}(\omega),\mathbb{X}_{u,v}(\omega)\Big)}=\lim_{|\Pi|\to 0}\sum_\Pi(y_u,y'_u)\bar{\otimes}(\theta_u\omega)_{v-u}, \tag{3.6}$$

where the operator $\bar{\otimes}$ is well defined. Because $\theta_{u+r}\omega = \theta_u \circ \theta_r\omega$, it is easy to check that the rough integral in (3.6) satisfies the additivity and the shift properties, i.e.

$$\int_a^c y_u d\omega_u = \int_a^b y_u d\omega_u + \int_b^c y_u d\omega_u, \quad \forall a \leq b \leq c; \tag{3.7}$$

$$\int_{a+r}^{b+r} y_u d\omega_u = \int_a^b y_{u+r} d(\theta_r\omega)_u, \quad \forall a \leq b, r \in \mathbb{R}. \tag{3.8}$$

These two properties (3.7), (3.8) and Theorem 2.2 then suffice to prove the cocycle property (3.1) of the generated random dynamical system from stochastic rough differential equation (1.1). □

3.2 Existence of random attractors

Given a random dynamical system φ on the phase space $\mathbb{R}^d$, we follow [8] (see also [9], [1, Chapter 9] and the references therein) to present the notion of random pullback attractors. Recall that a set $\hat{M} := \{M(\omega)\}_{\omega\in\Omega}$ is a *random set*, if $\omega \mapsto d(y|M(\omega)) := \inf\{d(y,z)|z \in M(\omega)\}$ is $\mathscr{F}$-measurable for each $y \in \mathbb{R}^d$. An *universe* $\mathscr{D}$ is a family of random sets which is closed w.r.t. inclusions (i.e. if $\hat{D}_1 \in \mathscr{D}$ and $\hat{D}_2 \subset \hat{D}_1$ then $\hat{D}_2 \in \mathscr{D}$). In our setting, we define the universe $\mathscr{D}$ to be a family of *tempered* random sets $D(\omega)$, which means the following: A random variable $\rho(\omega) > 0$ is called *tempered* if it satisfies $\lim\limits_{t\to\pm\infty} \frac{1}{t}\log^+ \rho(\theta_t\omega) = 0$ a.s. (see e.g. [1, pp. 164, 386]) which is equivalent to the sub-exponential growth $\lim\limits_{t\to\pm\infty} e^{-c|t|}\rho(\theta_t\omega) = 0$ a.s. for all $c > 0$ [28, p. 220]). A random set $D(\omega)$ is called *tempered* if it is contained in a ball $B(0,\rho(\omega))$ a.s., where the radius $\rho(\omega)$ is a tempered random variable.

A random subset A is called invariant, if $\varphi(t,\omega)A(\omega) = A(\theta_t\omega)$ for all $t \in \mathbb{R}$, $\omega \in \Omega$. An invariant random compact set $\mathcal{A} \in \mathscr{D}$ is called a *pullback attractor* in $\mathscr{D}$, if $\mathcal{A}$ attracts any closed random set $\hat{D} \in \mathscr{D}$ in the pullback sense, i.e.

$$\lim_{t\to\infty} d_H(\varphi(t,\theta_{-t}\omega)\hat{D}(\theta_{-t}\omega)|\mathcal{A}(\omega)) = 0, \tag{3.9}$$

where $d_H(\cdot|\cdot)$ is the Hausdorff semi-distance, i.e. $d_H(D|A) := \sup_{d\in D}\inf_{a\in A}\|d-a\|$. $\mathcal{A}$ is called a *forward attractor* in $\mathscr{D}$, if $\mathcal{A}$ is invariant and attracts any closed random set $\hat{D} \in \mathscr{D}$ in the forward sense, i.e. $\lim\limits_{t\to\infty} d_H(\varphi(t,\omega)\hat{D}(x)|\mathcal{A}(\theta_t\omega)) = 0$.

The existence of a pullback attractor follows from the existence of a pullback absorbing set (see [9, Theorem 3]), namely a random set $\mathcal{B} \in \mathscr{D}$ is called *pullback absorbing* in the universe $\mathscr{D}$ if $\mathcal{B}$ absorbs all closed random sets in $\mathscr{D}$, i.e. for any closed random set $\hat{D} \in \mathscr{D}$, there exists a time $t_0 = t_0(\omega,\hat{D})$ such that

$$\varphi(t,\theta_{-t}\omega)\hat{D}(\theta_{-t}\omega) \subset \mathcal{B}(\omega), \text{ for all } t \geq t_0. \tag{3.10}$$

Then given the universe $\mathscr{D}$ and a compact pullback absorbing set $\mathcal{B} \in \mathscr{D}$, there exists a unique pullback attractor $\mathcal{A}(\omega)$ in $\mathscr{D}$, given by

$$\mathcal{A}(\omega) = \bigcap_{t\geq 0}\overline{\bigcup_{s\geq t}\varphi(s,\theta_{-s}\omega)\mathcal{B}(\theta_{-s}\omega)}. \tag{3.11}$$

Our first main result is formulated as follows.

Theorem 3.3 *Under the assumptions* ($\boldsymbol{H}_f$)*,* ($\boldsymbol{H}_g$)*,* ($\boldsymbol{H}_X$) *and an additional condition that* $f \in C^1$*, there exists a pullback attractor* $\mathcal{A}(\omega)$ *for the generated random dynamical system of the stochastic system* (1.1) *such that* $|\mathcal{A}(\cdot)| \in \mathcal{L}^\rho$ *for any* $\rho \geq 1$.

Proof: First (2.21) and Jensen's inequality deduce that, for any $\rho \geq 1$ there exists an $\eta \in (0,1)$ and an integrable random variable $\xi_1(\omega) = \xi_1(C_g \left\|\!\left|\mathbf{x}(\omega)\right|\!\right\|_{p-\mathrm{var},[0,1]})$ such that

$$\|y_1\|^\rho \leq \eta \|y_0\|^\rho + \xi_1(\omega). \tag{3.12}$$

From (3.12) it is easy to prove by induction that

$$\|y_n(\mathbf{x},y_0)\|^\rho \leq \eta^n \|y_0\|^\rho + \sum_{i=0}^{n-1} \eta^i \xi_1(\theta_{n-i}\omega), \quad \forall n \geq 1;$$

thus replacing ω by $\theta_{-n}\omega$ yields

$$\|y_n(\theta_{-n}\omega, y_0(\theta_{-n}\omega))\|^\rho \leq \eta^n \|y_0(\theta_{-n}\omega)\|^\rho + \sum_{i=0}^{\infty} \eta^i \xi_1(\theta_{-i}\omega).$$

In other words, starting from a tempered random set $D(\omega) \in \mathscr{D}$ which is contained in a ball $B(0,r(\omega))$ with a tempered random radius $r(\omega)$, then any point $y_0 = y_0(\theta_{-t}\omega) \in D(\theta_{-t}\omega)$ satisfies

$$\begin{aligned} \Big\|y_n(\theta_{-n}\omega, y_0(\theta_{-n}\omega))\Big\|^\rho &\leq \eta^n r(\theta_{-n}\omega)^\rho + \underbrace{\sum_{i=0}^{\infty} \eta^i \xi_1(\theta_{-i}\omega)}_{=:R(\omega)} \\ \Rightarrow \quad \Big\|\varphi(n,\theta_{-n}\omega) D(\theta_{-n}\omega))\Big\|^\rho &\leq \eta^n r(\theta_{-n}\omega)^\rho + R(\omega). \end{aligned} \tag{3.13}$$

Due to the integrability of ξ_1, R is also integrable (and thus tempered) with $\mathbb{E}R(\cdot) = \frac{1}{1-\eta}\mathbb{E}\xi_1(\cdot)$. On the other hand, the cocycle property (3.1) yields

$$\varphi(t+n,\theta_{-t-n}\omega)D(\theta_{-t-n}\omega) = \varphi(n,\theta_{-n}\omega) \circ \varphi(t,\theta_{-t-n}\omega)D(\theta_{-t-n}\omega), \quad \forall t \in [0,1]. \tag{3.14}$$

It follows from (2.21) and the shift property (3.3) that

$$\begin{aligned} \|\varphi(t,\theta_{-t-n}\omega)y(\theta_{-t-n}\omega)\| &\leq \|y(\theta_{-t-n}\omega)\| + C_\lambda N\Big(\frac{\lambda}{16C_pC_g}, \mathbf{x}(\theta_{-t-n}\omega), [0,t]\Big) \\ &\leq r(\theta_{-t-n}\omega) + C_\lambda N\Big(\frac{\lambda}{16C_pC_g}, \mathbf{x}(\theta_{-n}\omega), [-1,0]\Big) \end{aligned} \tag{3.15}$$

for all $t \in [0,1]$ and $n \in \mathbb{N}$. Since the right hand side of (3.15) is a tempered random variable, we conclude from (3.13), (3.14) and (3.15) that there exists a pullback absorbing set $\mathcal{B}(\omega) = B(0,\hat{b}(\omega))$, with a tempered random variable $\hat{b}(\omega) = \Big[1 + R(\omega)\Big]^{\frac{1}{\rho}}$, containing our pullback attractor $\mathcal{A}(\omega)$. In particular, $|\mathcal{A}(\cdot)| \in \mathcal{L}^\rho$. □

Our second main result shows the existence and the upper-semi continuity of the pullback attractor in comparison to the deterministic attractor, under the additional Lipschitz continuity assumption (1.11) for the drift f and the uniform attraction assumption ($\mathbf{H}_\mathcal{A}$) for $\mathcal{A}$. Note that the Doss-Sussmann technique and the method in Theorem 2.2 do not work in this case because C_λ still contains D_1 and can not be arbitrarily small for sufficiently small C_g, thus we will provide a direct proof.

Theorem 3.4 *Assume that f is globally Lipschitz continuous with* (1.11) *and dissipative in the relative sense, i.e. there exists $D_1, D_2 > 0$ such that*

$$\langle y_1 - y_2, f(y_1) - f(y_2)\rangle \leq D_1 - D_2\|y_1 - y_2\|^2, \quad \forall y_1, y_2 \in \mathbb{R}^d. \tag{3.16}$$

Then under the assumptions ($\boldsymbol{H}_g$), ($\boldsymbol{H}_\mathcal{X}$), ($\boldsymbol{H}_\mathcal{A}$), the random attractor is upper semi-continuous, i.e.

$$\lim_{C_g \to 0} d_H\Big(\mathcal{A}(\omega)|\mathcal{A}\Big)^\rho = 0 \quad a.s. \quad and \quad \lim_{C_g \to 0} \mathbb{E}\; d_H\Big(\mathcal{A}(\cdot)|\mathcal{A}\Big)^\rho = 0, \quad \forall \rho \geq 1. \tag{3.17}$$

Proof: Fix any solution $y_t(\mathbf{x}, y_0)$ and associate it with the solution $\mu_t(\mu_0)$ of the deterministic system $\dot{\mu} = \bar{f}(\mu)$ which starts at μ_0. Consider the difference $y_t^* := y_t - \mu_t$ for $t \geq 0$, then y^* satisfies the equation

$$dy_t^* = [f(y_t^* + \mu_t) - f(\mu_t)]dt + g(y_t^* + \mu_t)dx_t = \bar{f}(y_t^*)dt + \bar{g}(y_t^*)dx_t.$$

First, we prove that there exists a constant $\eta \in (0,1)$ and an integrable random variable $\xi_1(\omega)$ such that

$$\|y_r^*\|^\rho \leq \eta \|y_0^*\|^\rho + \xi_1(\omega). \tag{3.18}$$

holds for a certain instant $r > 0$. Consider the difference $\mu_t^* = \mu_t(y_0) - \mu_t(\mu_0)$ of the two solutions of the deterministic system (1.4) starting at different points y_0 and μ_0, then μ_t^* is the solution of the nonautonomous deterministic system $\frac{d}{dt}\mu_t^* = f(\mu_t^* + \mu_t) - f(\mu_t)$ which starts at $\mu_0^* = y_0^*$. The relative dissipativity assumption (3.16) yields $\|\mu_t^*\| \leq \frac{D_1}{D_2} + \|\mu_0^*\|$ and

$$\|\mu_{s,t}^*\| \leq \int_s^t L_f \|\mu_u^*\| du \leq L_f(1 + \frac{D_1}{D_2})(1 + \|y_0^*\|)(t-s), \quad \forall 0 \leq s < t \leq r, \tag{3.19}$$

thus $\mu^* \in C^{1-\mathrm{var}}$. Moreover, because $(\mathbf{H}_\mathcal{A})$ is fulfilled with certain numbers r, D_3, we can choose μ_0 depending on y_0 such that (1.13) is satisfied, i.e.

$$\|\mu_r^*\| \leq \|\mu_0^*\| e^{-D_2 r}. \tag{3.20}$$

Assign $h_t := y_t^* - \mu_t^*$, then h satisfies

$$h_{s,t} = \int_s^t \Big[f(h_u + \mu_u + \mu_u^*) - f(\mu_u + \mu_u^*)\Big] du + \int_s^t g(h_u + \mu_u + \mu_u^*) dx_u, \tag{3.21}$$

thus h is also controlled by x with $h_s' = y_s' = g(h_s + \mu_s + \mu_s^*)$ and $R_{s,t}^h = R_{s,t}^y - \mu_{s,t} - \mu_{s,t}^*$. We need an auxiliary result below.

Proposition 3.5 *There exists a generic constant D such that the following estimate holds*

$$\|h_r\| \leq \underbrace{e^{4L_f r}\Big(1 + 4C_p r \|f\|_{\infty,\mathcal{A}} + 4C_p D\Big) 8 C_p C_g \,\|\!|\mathbf{x}|\!\|_{p-\mathrm{var},[0,r]}\, N\Big(\frac{1}{8C_pC_g}, \mathbf{x}, [0,r]\Big)}_{=:\xi_0(\mathbf{x})} (1 + \|y_0^*\|^\beta). \tag{3.22}$$

Assume that (3.22) in Proposition 3.5 holds, we then apply Jensen's inequality and Young inequality, for $\epsilon > 0$ small enough, and use (3.20) to conclude that

$$\begin{aligned}
\|y_r^*\|^\rho &\leq (\|h_r\| + \|\mu_r^*\|)^\rho \leq (1+\epsilon)^{\rho-1}\|\mu_r^*\|^\rho + (\frac{1+\epsilon}{\epsilon})^{\rho-1}\|h_r\|^\rho \\
&\leq (1+\epsilon)^{\rho-1}\|y_0^*\|^\rho (e^{-D_2})^\rho + (\frac{1+\epsilon}{\epsilon})^{\rho-1}\Big[(1+\epsilon)^{\rho-1}\xi_0(\mathbf{x})^\rho \|y_0^*\|^{\rho\beta} + (\frac{1+\epsilon}{\epsilon})^{\rho-1}\xi_0(\mathbf{x})^\rho\Big] \\
&\leq (1+\epsilon)^{\rho-1}\|y_0^*\|^\rho e^{-\rho D_2} + (\frac{1+\epsilon}{\epsilon})^{2\rho-2}\xi_0(\mathbf{x})^\rho \\
&\quad + (\frac{1+\epsilon}{\epsilon})^{\rho-1}(1+\epsilon)^{\rho-1}\Big[\beta\Big(\epsilon^{\rho\beta}\|y_0^*\|^{\rho\beta}\Big)^{\frac{1}{\beta}} + (1-\beta)\Big(\frac{1}{\epsilon^{\rho\beta}}\xi_0(\mathbf{x})^\rho\Big)^{\frac{1}{1-\beta}}\Big] \\
&\leq (1+\epsilon)^{2(\rho-1)}\Big(e^{-\rho D_2} + \epsilon\beta\Big)\|y_0^*\|^\rho + \xi_1(\mathbf{x}),
\end{aligned}$$

where

$$\xi_1(\mathbf{x}) = (\frac{1+\epsilon}{\epsilon})^{\rho-1}(1+\epsilon)^{\rho-1}(1-\beta)\Big(\frac{1}{\epsilon^{\rho\beta}}\xi_0(\mathbf{x})^\rho\Big)^{\frac{1}{1-\beta}} + (\frac{1+\epsilon}{\epsilon})^{2\rho-2}\xi_0(\mathbf{x})^\rho. \tag{3.23}$$

Note that $\xi_1(\mathbf{x}(\omega)) = \xi_1(\omega)$ is integrable in ω due to the assumption. By choosing $\epsilon \in (0,1)$ small enough such that

$$\eta := (1+\epsilon)^{2(\rho-1)}\Big[e^{-\rho D_2} + \epsilon\beta\Big] < 1,$$

we obtain (3.18).

Next, the same arguments as in the proof of Theorem 3.3 for discrete times nr lead to

$$\begin{aligned}\Big\|y_n(\theta_{-n}\omega, y_0(\theta_{-n}\omega)) - \mu_n(\mu_0)\Big\|^\rho &\leq \eta^n\Big(\|y_0(\theta_{-n}\omega)\| + \|\mu_0\|\Big)^\rho + \underbrace{\sum_{i=0}^\infty \eta^i \xi_1(\theta_{-i}\omega)}_{=:R(\omega)} \\ \Rightarrow d_H\Big(\varphi_n(\theta_{-n}\omega, D(\theta_{-n}\omega))|\mathcal{A}\Big)^\rho &\leq \eta^n\Big(|D(\theta_{-n}\omega)| + |\mathcal{A}|\Big)^\rho + R(\omega). \qquad (3.24)\end{aligned}$$

The final argument in the proof of Theorem 3.3 is then applied to prove that there exists a pullback absorbing set $\mathcal{B}(\omega) = B(0,\hat{b}(\omega))$, with a tempered random variable $\hat{b}(\omega) = |\mathcal{A}| + \Big[1 + R(\omega)\Big]^{\frac{1}{\rho}}$, containing our pullback attractor $\mathcal{A}(\omega)$. In particular, choose $D(\omega) = \mathcal{A}(\omega)$ (which is naturally tempered) and let n tends to infinity in the inequality (3.24), then the first term in the right hand side of (3.24) tends to zero due to the temperedness of $\mathcal{A}(\omega)$ and we obtain

$$d_H\Big(\mathcal{A}(\omega)|\mathcal{A}\Big)^\rho \leq R(\omega). \qquad (3.25)$$

Now because

$$8C_pC_g \left\|\!\left|\mathbf{x}\right|\!\right\|_{p-\mathrm{var},[0,r]} N\Big(\frac{1}{8C_pC_g}, \mathbf{x}, [0,r]\Big) \leq 8C_pC_g \left\|\!\left|\mathbf{x}\right|\!\right\|_{p-\mathrm{var},[0,r]} \Big[1 + \Big(8C_pC_g \left\|\!\left|\mathbf{x}\right|\!\right\|_{p-\mathrm{var},[0,r]}\Big)^p\Big] \to 0$$

as $C_g \to 0$, both in the almost sure and in the $\mathcal{L}^\rho$ senses, the definitions of $\xi_0(\omega)$ in (3.22) and of $\xi_1(\omega)$ in (3.23) show that $R(\omega) \to 0$ as $C_g \to 0$, both in the almost sure and in the $\mathcal{L}^\rho$ senses. This proves (3.17).

□

Proof: [**Proof of Proposition 3.5**] First, it follows from (3.21) that

$$\begin{aligned}\|h_{s,t}\| &\leq \int_s^t L_f\|h_u\|du + C_g\|x_{s,t}\| + C_g^2\|\mathbb{X}_{s,t}\| \\ &\quad + C_p\Big\{ \left\|\!\left|\mathbb{X}\right|\!\right\|_{q-\mathrm{var},[s,t]^2} \left\|\!\left|[g(y)]'\right|\!\right\|_{p-\mathrm{var},[s,t]} + \left\|\!\left|x\right|\!\right\|_{p-\mathrm{var},[s,t]} \left\|\!\left|R^{g(y)}\right|\!\right\|_{q-\mathrm{var},[s,t]^2}\Big\}. \qquad (3.26)\end{aligned}$$

Observe that for $\beta = \frac{2}{p} \in (\frac{2}{3}, 1)$,

$$\begin{aligned}&\|g(\mu_u + h_u + \mu^*_u) - g(\mu_v + h_v + \mu^*_v)\| \vee \|Dg(\mu_u + h_u + \mu^*_u) - Dg(\mu_v + h_v + \mu^*_v)\| \\ \leq\ & C_g\|h_{u,v}\| + C_g\|\mu_{u,v}\| + 2C_g\|\mu^*_{u,v}\|^\beta \\ \leq\ & C_g\|h_{u,v}\| + C_g\|\mu_{u,v}\| + C_gD(1 + \|y^*_0\|^\beta)(t-s)^\beta, \quad \forall 0 \leq u < v \leq r,\end{aligned}$$

for a generic constant D. This follows that

$$\begin{aligned}&\left\|\!\left|g(\mu + h + \mu^*)\right|\!\right\|_{p-\mathrm{var},[s,t]} \vee \left\|\!\left|Dg(\mu + h + \mu^*)\right|\!\right\|_{p-\mathrm{var},[s,t]} \\ \leq\ & C_g \left\|\!\left|h\right|\!\right\|_{p-\mathrm{var},[s,t]} + C_g \left\|\!\left|\mu\right|\!\right\|_{1-\mathrm{var},[s,t]} + C_gD(1 + \|y^*_0\|^\beta)(t-s)^\beta, \quad \forall 0 \leq s < t \leq r. \qquad (3.27)\end{aligned}$$

Inequality (3.27) together with $[g(y)]'_s = Dg(y_s)g(y_s)$ leads to

$$\left\|\!\left|[g(y)]'\right|\!\right\|_{p-\mathrm{var},[s,t]} \leq 2C_g^2\Big(\left\|\!\left|h\right|\!\right\|_{p-\mathrm{var},[s,t]} + \left\|\!\left|\mu\right|\!\right\|_{1-\mathrm{var},[s,t]} + D(1 + \|y^*_0\|^\beta)\Big), \quad \forall 0 \leq s < t \leq r. \qquad (3.28)$$

Furthermore,

$$\begin{aligned}
\|R^{g(y)}_{s,t}\| &= \|g(y_s + h'_s x_{s,t} + R^h_{s,t} + \mu_{s,t} + \mu^*_{s,t}) - g(y_s) - Dg(y_s)g(y_s)x_{s,t}\| \\
&\leq \|g\big(y_s + g(y_s)x_{s,t} + R^h_{s,t} + \mu_{s,t} + \mu^*_{s,t}\big) - g\big(y_s + g(y_s)x_{s,t}\big)\| \\
&\quad + \|g\big(y_s + g(y_s)x_{s,t}\big) - g(y_s) - Dg(y_s)g(y_s)x_{s,t}\| \\
&\leq C_g\|R^h_{s,t}\| + C_g\|\mu_{s,t}\| + 2C_g\|\mu^*_{s,t}\|^\beta \\
&\quad + \int_0^1 \|D_g\big(y_s + \chi g(y_s)x_{s,t}\big) - Dg(y_s)\|\|g(y_s)\|\|x_{s,t}\|d\chi \\
&\leq C_g\|R^h_{s,t}\| + C_g\|\mu_{s,t}\| + C_g D(1+\|y^*_0\|^\beta)(t-s)^\beta + \frac{1}{2}C_g^3\|x_{s,t}\|^2, \quad \forall 0 \leq s < t \leq r,
\end{aligned}$$

which, due to $\beta q = \beta \frac{p}{2} = 1$, yields

$$\left\|R^{g(y)}\right\|_{q-\mathrm{var},[s,t]^2} \leq C_g \left\|R^h\right\|_{q-\mathrm{var},[s,t]^2} + C_g \left\|\mu\right\|_{1-\mathrm{var},[s,t]} + C_g D(1+\|y^*_0\|^\beta) + \frac{1}{2}C_g^3 \left\|x\right\|^2_{p-\mathrm{var},[s,t]} \quad (3.29)$$

for all $0 \leq s < t \leq r$. Replacing (3.28) and (3.29) into (3.26) we obtain

$$\begin{aligned}
\|h\|_{p-\mathrm{var},[s,t]} &\leq \int_s^t L_f\|h_u\|du + C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} + C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} + \frac{1}{2}C_p C_g^3 \left\|\mathbf{x}\right\|^3_{p-\mathrm{var},[s,t]} \\
&\quad + 2C_p\Big(C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} \vee C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]}\Big)\Big(\left\|\mu\right\|_{1-\mathrm{var},[s,t]} + D(1+\|y^*_0\|^\beta)\Big) \\
&\quad + 2C_p\Big(C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} \vee C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]}\Big)\left\|h, R^h\right\|_{p-\mathrm{var},[s,t]}, \quad \forall 0 \leq s < t \leq r.
\end{aligned}$$

The estimate for R^h is already included in the right hand side of the above inequality (excluded the term $C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]}$). Since $\|h_u\| \leq \|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,u]}$, we finally get

$$\begin{aligned}
\left\|h, R^h\right\|_{p-\mathrm{var},[s,t]} &= \left\|h\right\|_{p-\mathrm{var},[s,t]} + \left\|R^h\right\|_{q-\mathrm{var},[s,t]^2} \\
&\leq \int_s^t 2L_f\Big(\|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,u]}\Big)du \\
&\quad + \Big(C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} + 2C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} + C_p C_g^3 \left\|\mathbf{x}\right\|^3_{p-\mathrm{var},[s,t]}\Big) \times \\
&\quad \times\Big(1 + 4C_p \left\|\mu\right\|_{1-\mathrm{var},[s,t]} + 4C_p D(1+\|y^*_0\|^\beta)\Big) \\
&\quad + 4C_p\Big(C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[s,t]} \vee C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]}\Big)\left\|h, R^h\right\|_{p-\mathrm{var},[s,t]} \\
&\leq \int_s^t 2L_f\Big(\|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,u]}\Big)du + \frac{1}{2}\left\|h, R^h\right\|_{p-\mathrm{var},[s,t]} \\
&\quad + \underbrace{4C_p C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]}\Big(1 + 4C_p \left\|\mu\right\|_{1-\mathrm{var},[s,t]} + 4C_p D(1+\|y^*_0\|^\beta)\Big)}_{=:L_1}, \quad (3.30)
\end{aligned}$$

whenever $4C_p C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \leq \frac{1}{2}$. Estimate (3.30) yields

$$\|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,t]} \leq \|h_s\| + 2L_1 + \int_s^t 4L_f\Big(\|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,u]}\Big)du$$

whenever $4C_p C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \leq \frac{1}{2}$, thus by the continuous Gronwall lemma,

$$\|h_s\| + \left\|h, R^h\right\|_{p-\mathrm{var},[s,t]} \leq (\|h_s\| + 2L_1)e^{4L_f(t-s)}$$

whenever $4C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \leq \frac{1}{2}$. Now by constructing the greedy sequence of stopping times of the form $\{\tau_i(\frac{1}{8C_pC_g}, \mathbf{x}, [s,t])\}_{i\in\mathbb{N}}$ as in (2.10), we can prove by induction that

$$\|h\|_{\infty,[\tau_k,\tau_{k-1}]} \leq e^{4L_f(\tau_k - s)}\Big(\|h_s\| + 2L_1 k\Big), \quad \forall k = 1,\ldots, N\Big(\frac{1}{8C_pC_g}, \mathbf{x}, [s,t]\Big).$$

This enables us to show that

$$\begin{aligned}\|h\|_{\infty,[s,t]} \quad &\leq \quad e^{4L_f(t-s)}\Big\{\|h_s\| + N\Big(\frac{1}{8C_pC_g}, \mathbf{x}, [s,t]\Big) \times \\ &\times 8C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[s,t]} \Big(1 + 4C_p \left\|\mu\right\|_{1-\mathrm{var},[s,t]} + 4C_pD(1 + \|y_0^*\|^\beta)\Big)\Big\}, \qquad (3.31)\end{aligned}$$

for all $0 \leq s < t \leq r$. Since $h_0 = y_0^* - \mu_0^* = 0$ and $\left\|\mu\right\|_{1-\mathrm{var},[0,r]} \leq r\|f\|_{\infty,\mathcal{A}}$ due to (1.12), (3.22) is proved. □

If f is strictly dissipative, i.e. $D_1 = 0$ in condition (3.16), then (1.5) is automatically satisfied and the attractor $\mathcal{A}$ is a singleton. However, it is not a trivial task to prove that $\mathcal{A}(\omega)$ is a singleton random attractor. In fact, we can only prove below that statement for sufficiently small C_g.

From now on, we follow the terminologies in the proof of Theorem 2.2 with

$$\gamma_t = y_t - z_t; \quad \bar{\gamma}_t = \bar{y}_t - \bar{z}_t; \quad \psi_t = \Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}, z_t)\Big]^{-1} - Id; \quad \bar{\psi}_t = \Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}, \bar{z}_t)\Big]^{-1} - Id.$$

Also from the proof of Theorem 2.2, given a time $\tau > 0$ such that $16C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[0,\tau]} \leq \lambda$ where $\lambda < \frac{D_2}{D_2 + 4C_f}$, it follows that $N\Big(\frac{1}{16C_pC_g}, \mathbf{x}, [0,t]\Big) = 1$ for all $t \in [0,\tau]$. We first need an auxiliary result.

Proposition 3.6 *Assume ($\boldsymbol{H}_g$), ($\boldsymbol{H}_X$) and λ, τ as introduced above. Then there exist an increasing continuous function $K : [0,1] \to \mathbb{R}_+$ with $K(0) = 0$, such that the following estimates hold*

$$\|\bar{\gamma}_t - \gamma_t\| \leq \lambda \|\bar{z}_t - z_t\|; \quad \|\bar{\psi}_t - \psi_t\| \leq K(\lambda)\|\bar{z}_t - z_t\|, \quad \forall t \in [0,\tau]. \qquad (3.32)$$

Proof: i, The proof for the first estimate is simple, since one can write $\bar{\gamma}_t - \gamma_t$ in the form

$$\bar{\gamma}_t - \gamma_t = \int_0^t [g(\bar{\phi}_u(\mathbf{x}, \bar{z}_t)) - g(\phi_u(\mathbf{x}, z_t))]dx_u.$$

Then the estimates (2.17) and (2.18) together with (2.14), (2.15) enable us to obtain

$$\begin{aligned}\|\bar{\gamma}_t - \gamma_t\| \quad &\leq \quad \|g(\bar{z}_t) - g(z_t)\| \left\|x\right\|_{p-\mathrm{var},[0,t]} + \|Dg(\bar{z}_t)g(\bar{z}_t) - Dg(z_t)g(z_t)\| \left\|\mathbb{X}\right\|_{q-\mathrm{var},[0,t]^2} \\ &\quad + C_p\Big\{ \left\|x\right\|_{p-\mathrm{var},[0,t]} \left\|R^{g(\bar{\phi})-g(\phi)}\right\|_{q-\mathrm{var},[0,t]^2} + \left\|\mathbb{X}\right\|_{q-\mathrm{var},[0,t]^2} \left\|[g(\bar{\phi}) - g(\phi)]'\right\|_{p-\mathrm{var},[0,t]} \Big\} \\ &\leq \quad 2C_p\Big\{C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[0,t]} + C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[0,t]} \Big\} \times \\ &\quad \times\Big(1 + \left\|\bar{\phi}, R^{\bar{\phi}}\right\|_{p-\mathrm{var},[0,t]} + \left\|\phi, R^{\phi}\right\|_{p-\mathrm{var},[0,t]}\Big)\Big(\|\bar{z}_t - z_t\| + \left\|\bar{\phi} - \phi, R^{\bar{\phi}-\phi}\right\|_{p-\mathrm{var},[0,t]}\Big) \\ &\leq \quad 2C_p\Big\{C_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[0,t]} + C_g^2 \left\|\mathbf{x}\right\|^2_{p-\mathrm{var},[0,t]} \Big\}\Big(1 + 16C_pC_g \left\|\mathbf{x}\right\|_{p-\mathrm{var},[0,t]}\Big)^2 \|\bar{z}_t - z_t\| \\ &\leq \quad \|\bar{z}_t - z_t\| \frac{\lambda}{4}(1+\lambda)^2 \leq \lambda\|\bar{z}_t - z_t\|, \qquad \forall t \in [0,\tau]. \qquad (3.33)\end{aligned}$$

ii, The proof for the second estimate in (3.32) is more technical and lengthy. First observe from (2.19) that

$$\|\bar{\psi}_t - \psi_t\| \quad = \quad \Big\|\Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}, \bar{z}_t)\Big]^{-1} - \Big[\frac{\partial \phi}{\partial z}(t, \mathbf{x}, z_t)\Big]^{-1}\Big\|$$

$$
\begin{aligned}
&\le \Big\|\Big[\frac{\partial\phi}{\partial z}(t,\mathbf{x},\bar z_t)\Big]^{-1}\Big\|\Big\|\Big[\frac{\partial\phi}{\partial z}(t,\mathbf{x},z_t)\Big]^{-1}\Big\|\Big\|\frac{\partial\phi}{\partial z}(t,\mathbf{x},\bar z_t)-\frac{\partial\phi}{\partial z}(t,\mathbf{x},z_t)\Big\| \\
&\le (1+\lambda)^2\Big\|\frac{\partial\phi}{\partial z}(t,\mathbf{x},\bar z_t)-\frac{\partial\phi}{\partial z}(t,\mathbf{x},z_t)\Big\|,\quad \forall t\in[0,\tau],
\end{aligned}
$$

hence it is enough to estimate the last term in the right hand side of the above inequality. We will write in short $\|\bar\xi_t-\xi_t\|$ for some $t\in[0,\tau]$ fixed, so that $\bar z_t$ and z_t are also fixed as well. By definition $\bar\xi_t$ and ξ_t are respectively the values at time t of the solutions $\bar\xi_u(\mathbf{x},\bar z_t)$ and $\xi_u(\mathbf{x},z_t)$ of the linear matrix - valued rough differential equations

$$
\bar\xi_v = Id+\int_0^v Dg(\phi_u(\mathbf{x},\bar z_t))\bar\xi_u dx_u;\quad \xi_v = Id+\int_0^v Dg(\phi_u(\mathbf{x},z_t))\xi_u dx_u.
$$

As a result,

$$
\underbrace{\bar\xi_v-\xi_v}_{=:\zeta_v} = \int_0^v \underbrace{Dg(\phi_u(\mathbf{x},\bar z_t))}_{=:A_u}\underbrace{(\bar\xi_u-\xi_u)}_{=:\zeta_u}dx_u + \int_0^v \underbrace{\Big[Dg(\phi_u(\mathbf{x},\bar z_t))-Dg(\phi_u(\mathbf{x},z_t))\Big]}_{=:b_u}\xi_u dx_u. \tag{3.34}
$$

It follows from (3.34) that $\zeta_0=0$ and

$$
\begin{aligned}
\|\zeta_{u,v}\| &= \Big\|\int_u^v A_\chi\zeta_\chi dx_\chi\Big\| + \underbrace{\Big\|\int_u^v b_\chi\xi_\chi dx_\chi\Big\|}_{=:\Lambda_{u,v}} \\
&\le \|A_u\|\|\zeta_u\|\,\|\!|x|\!\|_{p-\mathrm{var},[u,v]} + \|A_u\otimes A_u\|\|\zeta_u\|\,\|\!|\mathbb{X}|\!\|_{q-\mathrm{var},[u,v]} + \Lambda_{u,v} \\
&\quad +C_p\Big\{\|\!|x|\!\|_{p-\mathrm{var},[u,v]}\Big\|\!\Big|R^{A\zeta}\Big|\!\Big\|_{q-\mathrm{var},[u,v]} + \|\!|\mathbb{X}|\!\|_{q-\mathrm{var},[u,v]}\|\!|(A'+A\otimes A)\zeta|\!\|_{p-\mathrm{var},[u,v]}\Big\}.
\end{aligned}
$$

Since

$$
A_v\zeta_v - A_u\zeta_u = (A'_u\zeta_u + A_u\zeta'_u)x_{u,v} + R^A_{u,v}\zeta_u + A_uR^\zeta_{u,v} + A_{u,v}\zeta_{u,v},
$$

ζ and $A\zeta$ are controlled by x with $\zeta'_u = A_u\zeta_u$, $(A\zeta)'_u = A'_u\zeta_u + A_u\otimes A_u\zeta_u$ and

$$
\begin{aligned}
\|\!|(A'+A\otimes A)\zeta|\!\|_{p-\mathrm{var},[u,v]} &\le \|\!|A'+A\otimes A|\!\|_{p-\mathrm{var},[u,v]}\|\zeta\|_{\infty,[u,v]} + \|A'+A\otimes A\|_{\infty,[u,v]}\|\!|\zeta|\!\|_{p-\mathrm{var},[u,v]} \\
&\le 2\Big(\|A'\|_{p-\mathrm{var},[u,v]} + 2\|A\|^2_{p-\mathrm{var},[u,v]}\Big)\|\zeta\|_{p-\mathrm{var},[u,v]}; \\
\|R^{A\zeta}_{u,v}\| &\le \|R^A_{u,v}\|\|\zeta_u\| + \|A_u\|\|R^\zeta_{u,v}\| + \|A_{u,v}\|\|\zeta_{u,v}\|; \\
\Big\|\!\Big|R^{A\zeta}\Big|\!\Big\|_{q-\mathrm{var},[u,v]} &\le \|\!|R^A|\!\|_{q-\mathrm{var},[u,v]}\|\zeta\|_{\infty,[u,v]} + \|A\|_{\infty,[u,v]}\Big\|\!\Big|R^\zeta\Big|\!\Big\|_{q-\mathrm{var},[u,v]} \\
&\quad + \|\!|A|\!\|_{p-\mathrm{var},[u,v]}\|\!|\zeta|\!\|_{p-\mathrm{var},[u,v]} \\
&\le \Big(\|A_u\| + \|\!|A,R^A|\!\|_{p-\mathrm{var},[u,v]}\Big)\Big(\|\zeta_u\| + \Big\|\!\Big|\zeta,R^\zeta\Big|\!\Big\|_{p-\mathrm{var},[u,v]}\Big).
\end{aligned}
$$

Because $A_u = Dg(\phi_u(\mathbf{x},\bar z_t))$, a direct computation shows that $A'_u = D^2g(\phi_u(\mathbf{x},\bar z_t))g(\phi_u(\mathbf{x},\bar z_t))$ with

$$
\begin{aligned}
\|A\|_{p-\mathrm{var},[u,v]} &\le C_g(1+\|\!|\phi_\cdot(\mathbf{x},\bar z_t)|\!\|_{p-\mathrm{var},[u,v]}) \\
\|A'\|_{p-\mathrm{var},[u,v]} &\le C_g^2(1+2\,\|\!|\phi_\cdot(\mathbf{x},\bar z_t)|\!\|_{p-\mathrm{var},[u,v]}) \\
\|\!|R^A|\!\|_{q-\mathrm{var},[u,v]} &\le C_g\,\|\!|\phi_\cdot(\mathbf{x},\bar z_t)|\!\|^2_{p-\mathrm{var},[u,v]} + C_g\Big\|\!\Big|R^{\phi_\cdot(\mathbf{x},\bar z_t)}\Big|\!\Big\|_{q-\mathrm{var},[u,v]}
\end{aligned}
$$

As a result, by combining all the above estimates and using (2.14), we can show that there exists a generic function $D(\lambda)\ge 1$ such that

$$
\|\!|\zeta|\!\|_{p-\mathrm{var},[u,v]} \le \|\!|\Lambda|\!\|_{p-\mathrm{var},[u,v]} + C_p\Big(C_g\,\|\!|x|\!\|_{p-\mathrm{var},[u,v]}\vee C_g^2\,\|\!|\mathbf{x}|\!\|^2_{q-\mathrm{var},[u,v]}\Big)\times
$$

$$\times D(\lambda)\Big(\|\zeta_u\| + \Big\|\zeta, R^\zeta\Big\|_{p-\mathrm{var},[u,v]}\Big).$$

The estimate for $\left\| R^\zeta \right\|_{q-\mathrm{var},[u,v]}$ is already included in the above estimate, hence

$$\begin{aligned}\Big\|\zeta, R^\zeta\Big\|_{p-\mathrm{var},[u,v]} &\leq 2\,\|\Lambda\|_{p-\mathrm{var},[u,v]} + 2C_p D(\lambda)\Big(C_g\,\|x\|_{p-\mathrm{var},[u,v]} \vee C_g^2\,\|\mathbf{x}\|^2_{q-\mathrm{var},[u,v]}\Big)\times \\ &\quad \times\Big(\|\zeta_u\| + \Big\|\zeta, R^\zeta\Big\|_{p-\mathrm{var},[u,v]}\Big), \quad \forall 0\leq u<v\leq t.\end{aligned}$$

This implies

$$\|\zeta_v\| \leq \|\zeta_u\| + \Big\|\zeta, R^\zeta\Big\|_{p-\mathrm{var},[u,v]} \leq 2\,\|\Lambda\|_{p-\mathrm{var},[0,t]} + 2\|\zeta_u\| \text{ whenever } \|\mathbf{x}\|_{p-\mathrm{var},[u,v]} \leq \underbrace{\Big[4C_pC_gD(\lambda)\Big]^{-1}}_{=:\lambda'}.$$

By constructing the stopping times $\Big\{\tau_i(\lambda',\mathbf{x},[0,t])\Big\}$ and using induction, we can show that

$$\begin{aligned}\|\zeta_t\| &\leq \exp\{N(\lambda',\mathbf{x},[0,t])\log 2\}\Big(\|\zeta_0\| + 2\,\|\Lambda\|_{p-\mathrm{var},[0,t]}\Big) \\ &\leq 2\,\|\Lambda\|_{p-\mathrm{var},[0,t]}\exp\{N(\lambda',\mathbf{x},[0,t])\log 2\}.\end{aligned} \tag{3.35}$$

It remains to estimate $\|\Lambda\|_{p-\mathrm{var},[0,t]}$. Observe that

$$\begin{aligned}\|\Lambda_{u,v}\| &\leq \|b_u\xi_u\|\,\|x\|_{p-\mathrm{var},[u,v]} + \|(b\xi)'_u\|\,\|\mathbb{X}\|_{q-\mathrm{var},[u,v]} \\ &\quad + C_p\Big(\|x\|_{p-\mathrm{var},[u,v]}\Big\|R^{b\xi}\Big\|_{q-\mathrm{var},[u,v]} + \|\mathbb{X}\|_{q-\mathrm{var},[u,v]}\,\|(b\xi)'\|_{p-\mathrm{var},[u,v]}\Big),\end{aligned}$$

which yields

$$\begin{aligned}\|\Lambda\|_{p-\mathrm{var},[0,t]} &\leq \|b\|_{\infty,[0,t]}\|\xi\|_{\infty,[0,t]}\,\|\mathbf{x}\|_{p-\mathrm{var},[0,t]} \\ &\quad + \Big(\|b'\|_{\infty,[0,t]}\|\xi\|_{\infty,[0,t]} + \|b\|_{\infty,[0,t]}\|\xi'\|_{\infty,[0,t]}\Big)\,\|\mathbf{x}\|^2_{q-\mathrm{var},[0,t]} \\ &\quad + C_p\,\|\mathbf{x}\|_{p-\mathrm{var},[0,t]}\Big(\|b_0\| + \Big\|b, R^b\Big\|_{p-\mathrm{var},[0,t]}\Big)\Big(\|\xi_0\| + \Big\|\xi, R^\xi\Big\|_{p-\mathrm{var},[0,t]}\Big) \\ &\quad + C_p\,\|\mathbf{x}\|^2_{q-\mathrm{var},[0,t]}\Big\{\|b'\|_{p-\mathrm{var},[0,t]}\,\|\xi\|_{\infty,[0,t]} + \|b'\|_{\infty,[0,t]}\,\|\xi\|_{p-\mathrm{var},[0,t]} \\ &\qquad + \|\xi'\|_{p-\mathrm{var},[0,t]}\,\|b\|_{\infty,[0,t]} + \|\xi'\|_{\infty,[0,t]}\,\|b\|_{p-\mathrm{var},[0,t]}\Big\}.\end{aligned}$$

It is easy to check from (2.19) that the estimates for $\left\|\xi, R^\xi\right\|_{p-\mathrm{var},[0,t]}, \|\xi\|_{\infty,[0,t]}, \|\xi'\|_{\infty,[0,t]}$ and $\|\xi'\|_{p-\mathrm{var},[0,t]}$ are functions of $16C_pC_g\,\|\mathbf{x}\|_{p-\mathrm{var},[0,t]}$ and can thus be bounded from above by functions of λ. On the other hand, similar to the computations in Proposition 2.1, we can show that

$$\begin{aligned}\|b\|_{\infty,[0,t]} &\leq C_g\|\bar{\phi}_\cdot(\mathbf{x},\bar{z}_t) - \phi_\cdot(\mathbf{x},z_t)\|_{\infty,[0,t]}; \\ \|b'\|_{\infty,[0,t]} &\leq \Big\|D^2g(\phi_\cdot(\mathbf{x},\bar{z}_t))g(\phi_\cdot(\mathbf{x},\bar{z}_t)) - D^2g(\phi_\cdot(\mathbf{x},z_t))g(\phi_\cdot(\mathbf{x},z_t))\Big\|_{\infty,[0,t]}; \\ \|b\|_{p-\mathrm{var},[0,t]} &\leq C_g\Big(\|\phi_\cdot(\mathbf{x},\bar{z}_t) - \phi_\cdot(\mathbf{x},z_t)\|_{p-\mathrm{var},[0,t]} \\ &\qquad + \|\phi_\cdot(\mathbf{x},\bar{z}_t) - \phi_\cdot(\mathbf{x},z_t)\|_{\infty,[0,t]}\,\|\phi_\cdot(\mathbf{x},z_t)\|_{p-\mathrm{var},[0,t]}\Big); \\ \|b'\|_{p-\mathrm{var},[0,t]} &\leq C_g\Big(\|\phi_\cdot(\mathbf{x},\bar{z}_t) - \phi_\cdot(\mathbf{x},z_t)\|_{p-\mathrm{var},[0,t]} \\ &\qquad + \|\phi_\cdot(\mathbf{x},\bar{z}_t) - \phi_\cdot(\mathbf{x},z_t)\|_{\infty,[0,t]}\,\|\phi_\cdot(\mathbf{x},z_t)\|_{p-\mathrm{var},[0,t]}\Big);\end{aligned}$$

$$
\begin{aligned}
\left\|R^b\right\|_{q-\mathrm{var},[0,t]} &\leq C_g \left\|R^{\phi.(\mathbf{x},\bar{z}_t)-\phi.(\mathbf{x},z_t)}\right\|_{q-\mathrm{var},[0,t]} \\
&+C_g\|\phi.(\mathbf{x},\bar{z}_t)-\phi.(\mathbf{x},z_t)\|_{\infty,[0,t]} \left\|R^{\phi.(\mathbf{x},z_t)}\right\|_{q-\mathrm{var},[0,t]} \\
&+\frac{1}{2}C_g^2 \|\mathbf{x}\|_{p-\mathrm{var},[0,t]} \Big[\|\phi.(\mathbf{x},\bar{z}_t)-\phi.(\mathbf{x},z_t)\|_{p-\mathrm{var},[0,t]} \\
&+\|\phi.(\mathbf{x},\bar{z}_t)-\phi.(\mathbf{x},z_t)\|_{\infty,[0,t]}\Big(\|\bar{\phi}\|_{p-\mathrm{var},[0,t]} + \|\phi\|_{p-\mathrm{var},[0,t]} \Big)\Big]. \qquad (3.36)
\end{aligned}
$$

Finally, the existence of function K satisfying (3.32) is a consequence of (3.35), (3.36) and the estimates (2.14), (2.16), (2.15) on $[0,\tau]$, which can be written as functions of λ.

□

Theorem 3.7 (Singleton attractor) *Assume ($\boldsymbol{H}_f$),($\boldsymbol{H}_g$),($\boldsymbol{H}_X$) and further that f is strictly dissipative (i.e. $D_1 = 0$ in condition* (3.16)*) and of C^1. Then for C_g small enough the random attractor $\mathcal{A}$ is a singleton, i.e. $\mathcal{A}(\omega) = \{a(\omega)\}$ a.s., thus it satisfies the upper semi-continuity* (3.17)*.*

Proof: We first prove that for any $\lambda < \frac{D_2}{2(D_2+C_f)}$, there exists a random variable $0 < \Delta(\lambda,\omega) \in \mathcal{L}^1$ such that for any two solutions $\bar{y}_t$ and y_t of system (1.2) starting respectively from $\bar{y}_0, y_0 \in \mathcal{A}(\omega)$, the following estimate holds

$$
\|\bar{y}_1 - y_1\| \leq \exp\{-D_2 + \Delta(\lambda,\mathbf{x}(\omega))\}\|\bar{y}_0 - y_0\|. \qquad (3.37)
$$

Indeed, given $\tau > 0$ as above, assign $\eta_t := \bar{z}_t - z_t$ and consider the equation

$$
\begin{aligned}
\dot{\eta}_t &= (Id + \bar{\psi}_t)f(\bar{z}_t + \bar{\gamma}_t) - (Id + \psi_t)f(z_t + \gamma_t) \\
&= [f(\bar{y}_t) - f(y_t)] + \bar{\psi}_t[f(\bar{y}_t) - f(y_t)] + (\bar{\psi}_t - \psi_t)f(y_t).
\end{aligned}
$$

It follows from Lagrange's mean value theorem, the strict dissipativity, estimates (2.21) and (3.32) that for all $t \in [0,\tau]$,

$$
\begin{aligned}
\frac{d}{dt}\|\eta_t\|^2 &\leq 2\Big\langle \bar{y}_t - y_t - (\bar{\gamma}_t - \gamma_t), f(\bar{y}_t) - f(y_t)\Big\rangle + 2\|\eta_t\|\|\bar{\psi}_t - \psi_t\|\|f(y_t)\| \\
&\quad +2\|\eta_t\|\|\bar{\psi}_t\|\|f(\bar{y}_t) - f(y_t)\| \\
&\leq -2D_2\|\bar{y}_t - y_t\|^2 + 2\|\eta_t\|\|\bar{\psi}_t - \psi_t\| \underbrace{\max\{\|f(p_1)\| : p_1 \in \mathcal{A}(\theta_t\omega)\}}_{=:\|f\|_{\infty,\mathcal{A}(\theta_t\omega)}} \\
&\quad +2\Big(\|\bar{\gamma}_t - \gamma_t\| + \|\bar{\psi}_t\|\|\eta_t\|\Big) \underbrace{\max\Big\{\|Df(p_1)\| : \|p_1\| \leq |\mathcal{A}(\theta_t\omega)|\Big\}}_{=:\|Df\|_{\infty,B(0,|\mathcal{A}(\theta_t\omega)|)}} \|\bar{y}_t - y_t\| \\
&\leq -2D_2\|\eta_t\|^2(1-\lambda) + 2\|f\|_{\infty,\mathcal{A}(\theta_t\omega)}K(\lambda)\|\eta_t\|^2 + 4\|Df\|_{\infty,B(0,|\mathcal{A}(\theta_t\omega)|)}(1+\lambda)\lambda\|\eta_t\|^2 \\
&\leq -2\Big\{D_2 - \underbrace{\Big(1 + \|Df\|_{\infty,B(0,|\mathcal{A}(\theta_t\omega)|)} + \|f\|_{\infty,\mathcal{A}(\theta_t\omega)}\Big)}_{=:\Xi(\theta_t\omega)} \underbrace{\Big[D_2\lambda + 2(1+\lambda)\lambda + K(\lambda)\Big]}_{=:G(\lambda)}\Big\}\|\eta_t\|^2.
\end{aligned}
$$

As a result $\|\eta_\tau\| \leq \|\eta_0\| \exp\Big\{-\int_0^\tau \Big[D_2 - G(\lambda)\Xi(\theta_t\omega)\Big]dt\Big\}$, which yields

$$
\begin{aligned}
\|\bar{y}_\tau - y_\tau\| &\leq (1+\lambda)\exp\Big\{-\int_0^\tau \Big[D_2 - G(\lambda)\Xi(\theta_t\omega)\Big]dt\Big\}\|\bar{y}_0 - y_0\| \\
&\leq \exp\Big\{\lambda - \int_0^\tau \Big[D_2 - G(\lambda)\Xi(\theta_t\omega)\Big]dt\Big\}\|\bar{y}_0 - y_0\|.
\end{aligned}
$$

By constructing the sequence of stopping times $\{\tau_i(\frac{\lambda}{16C_pC_g}, \mathbf{x}, [0,1])\}$ and using induction (taking into account the shift property (3.2)), we derive

$$\|\bar{y}_1 - y_1\| \leq \exp\Big\{ - D_2 + \underbrace{\Big[\lambda N\Big(\frac{\lambda}{16C_pC_g}, \mathbf{x}, [0,1]\Big) + G(\lambda)\int_0^1 \Xi(\theta_t\omega)dt\Big]}_{=:\Delta(\lambda,\mathbf{x}(\omega))}\Big\}\|\bar{y}_0 - y_0\|. \tag{3.38}$$

This proves (3.37). The integrability of $\Delta(\lambda,\cdot)$ follows from the integrability of $\|\!|\mathbf{x}(\cdot)|\!\|_{p-\mathrm{var},[0,1]}$ and of $\Xi(\cdot)$ (which is a consequence of $|\mathcal{A}(\cdot)| \in \mathcal{L}^\rho$ for any $\rho \geq 1$).

Next, take any two different points (if any) $a_1 \neq a_2 \in \mathcal{A}(\omega)$, we can also write $a_1(\omega), a_2(\omega) \in \mathcal{A}(\mathbf{x}(\omega))$ for a little abuse of notation to address the dependence on the path $\mathbf{x}$. For a given $n \in \mathbb{N}$, assign $\mathbf{x}^* := \mathbf{x}(\theta_{-n}\omega)$ and consider the equation

$$dy_t = f(y_t)dt + g(y_t)dx_t^*,$$

where $\mathbf{x}^* = (x^*, \mathbb{X}^*)$. Due to the invariance of $\mathcal{A}(\omega)$ under the flow, there exist $b_1, b_2 \in \mathcal{A}(\mathbf{x}^*)$ such that $a_i = y_n(\mathbf{x}^*, b_i)$. We write in short $y_t^1 = y_t(\mathbf{x}^*, b_1)$. Then by (3.37) and induction, one can use the shift property (3.3) to show that

$$\begin{aligned}\|a_2(\omega) - a_1(\omega)\| &\leq \exp\Big\{ - D_2 n + \sum_{k=0}^{n-1}\Delta(\lambda, \mathbf{x}^*(\theta_k\omega))\Big\}\|b_2 - b_1\| \\ &\leq 2\exp\Big\{ - n\Big[D_2 - \frac{1}{n}\sum_{k=1}^{n}\Delta(\lambda, \mathbf{x}(\theta_{-k}\omega))\Big]\Big\}\Big(|\mathcal{A}| + R(\theta_{-n}\omega)\Big) \end{aligned}\tag{3.39}$$

Applying Birkhorff ergodic theorem and using (3.38), one gets

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\Delta(\lambda, \mathbf{x}(\theta_{-k}\omega)) = \mathbb{E}\Delta(\lambda, \mathbf{x}(\cdot)) = \lambda\mathbb{E}N\Big(\frac{\lambda}{16C_pC_g}, \mathbf{x}(\cdot), [0,1]\Big) + G(\lambda)\mathbb{E}\int_0^1 \Xi(\theta_t\cdot)dt \quad \text{a.s.}$$

where the second term in the right hand side is small by choosing λ small enough. Meanwhile, the first term can be controlled as small as possible by choosing $\lambda := C_g$ for sufficiently small C_g so that N is fixed to $N\Big(\frac{1}{16C_p}, \mathbf{x}, [0,1]\Big)$. On the other hand $|\mathcal{A}| + R(\theta_{-n}\omega)$ is a tempered random variable. Hence for sufficiently small C_g, the right hand side of (3.39) tends to zero exponentially as n tends to infinity a.s., which proves that $\mathcal{A}(\omega)$ is a singleton a.s.

Finally, because $D_1 = 0$, the constants $\bar{C}_\lambda$ and C_λ in the proof of Theorem 2.2 vanish at $\lambda = 0$. Hence ξ_1 and R are functions of C_λ and can be as small as possible by choosing $\lambda = C_g$ for small enough C_g so that N in (2.21) is fixed to $N\Big(\frac{1}{16C_p}, \mathbf{x}, [0,1]\Big)$. This proves the upper semi-continuity (3.17). □

3.3 Discussion on estimation of $\mathbb{E}N(\lambda, \mathbf{x}(\cdot), [0,1])$

Estimate (3.39) in the proof of Theorem 3.7 leads to the question whether the conclusion still holds for any C_g. To get an answer, we need to check if the following limit is zero

$$\limsup_{\lambda\to 0}\lambda\mathbb{E}N(\lambda, \mathbf{x}(\cdot), [0,1]). \tag{3.40}$$

A direct computation shows that

$$\begin{aligned}\mathbb{E}N(\lambda, \mathbf{x}(\cdot), [0,1]) &= \sum_{n=1}^{\infty} n\mathbb{P}\{\omega : N(\lambda, \mathbf{x}(\cdot), [0,1]) = n\} \\ &= \sum_{n=1}^{\infty} n\Big(\mathbb{P}\{\omega : N(\lambda, \mathbf{x}(\cdot), [0,1]) > n-1\} - \mathbb{P}\{\omega : N(\lambda, \mathbf{x}(\cdot), [0,1]) > n\}\Big) \\ &= \sum_{n=0}^{\infty} \mathbb{P}\{\omega : N(\lambda, \mathbf{x}(\cdot), [0,1]) > n\}.\end{aligned}$$

Therefore at a first try, we would like to estimate the limit (3.40) for Gaussian noises. Unfortunately, we will show below that simply applying the estimate of $N(\lambda, \mathbf{x}, [0,1])$ in [5] would lead to a failure.

More specifically, following [13, Chapter 10 & Chapter 11], let $\mathcal{W} = C(I, \mathbb{R}^m)$ be the probability space equipped with a Gaussian measure $\mathbb{P}$ and let (X_t) be a continuous centered Gaussian process. The associated *Cameron-Martin space* $\mathcal{H} \subset \mathcal{W}$ consists of paths $t \mapsto h_\cdot = E(ZX_\cdot)$, where $Z \in \mathcal{W}^1$ is an element in the so-called *first Wiener chaos*. If $\bar{h}_\cdot = E(\bar{Z}X_\cdot)$ denotes another element in $\mathcal{H}$ then the inner product $\langle h, \bar{h}\rangle_{\mathcal{H}} := E(Z\bar{Z})$ makes $\mathcal{H}$ a Hilbert space and $Z \mapsto h$ is an isometry between $\mathcal{W}^1$ and $\mathcal{H}$. The triple $(\mathcal{W}, \mathcal{H}, \mathbb{P})$ is then called the *abstract Wiener space*. It follows from [13, Proposition 11.2] that given the covariance property (3.4), $\mathcal{H}$ is continuously embedded in the space of continuous paths of finite q-variation, i.e. $\mathcal{H} \hookrightarrow C^{q-\mathrm{var}}([0,1], \mathbb{R}^d)$, and there exists a constant $C_{\mathrm{emb}} > 0$ such that

$$\|h\|_{q-\mathrm{var},[s,t]} \leq \|h\|_{\mathcal{H}} \sqrt{\|R\|_{q-\mathrm{var},[s,t]^2}} \leq C_{\mathrm{emb}} \|h\|_{\mathcal{H}}, \quad \forall h \in \mathcal{H}, \forall 0 \leq s < t \leq 1. \tag{3.41}$$

It then makes sense (see e.g. [14] or [5]) to define the so-called *translated rough path* $T_h\mathbf{x}$ as

$$T_h\mathbf{x} := \Big(x + h, \mathbb{X} + \int h \otimes dx + \int x \otimes dh + \int h \otimes dh\Big).$$

According to [5], $T_h : C^{p-\mathrm{var}} \to C^{p-\mathrm{var}}$ satisfies the estimate

$$\|T_h\mathbf{x}\|_{p-\mathrm{var},[s,t]} \leq \bar{C}_p\Big(\|\mathbf{x}\|_{p-\mathrm{var},[s,t]} + \|h\|_{q-\mathrm{var},[s,t]}\Big), \quad \forall 0 \leq s < t \leq 1. \tag{3.42}$$

In addition, assume that X has a natural lift to a geometric p - variation rough path $\mathbf{X}$. It is proved in [5, Proposition 6.2, Theorem 6.3] that there exists a set $E \subset \mathcal{W}$ of $\mathbb{P}$-full measure, with the property

$$\forall \omega \in E, \forall h \in \mathcal{H}, \forall \lambda > 0: \text{ if } \|\mathbf{X}(\omega - h)\|_{p-\mathrm{var},[0,1]} \leq \lambda \text{ then } \|h\|^q_{q-\mathrm{var},[0,1]} \geq N(2\bar{C}_p\lambda, \mathbf{X}(\omega), [0,1]).$$

Moreover,

$$\mathbb{P}\{\omega : N(2\bar{C}_p\lambda, \mathbf{X}(\omega), [0,1]) > n\} \leq \exp\{2a_\lambda^2\} \exp\Big\{\frac{-\lambda^2 n^{\frac{4}{p}}}{2C^2_{\mathrm{emb}}}\Big\}, \tag{3.43}$$

where C_{emb} and $\bar{C}_p$ are given in (3.41), (3.42) respectively, Φ^{-1} is the inverse of the standard normal cumulative distribution function and $a_\lambda := \Phi^{-1}(\mathbb{P}(B_\lambda))$ where $B_\lambda := \{\omega \in \mathcal{W} : \|\mathbf{X}(\omega)\|_{p-\mathrm{var},[0,1]} \leq \lambda\}$. In fact, a closer look shows that (3.43) follows from Borell's theorem [29, Theorem 4.3] and

$$\mathbb{P}\{\omega : N(2\bar{C}_p\lambda, \mathbf{X}(\omega), [0,1]) > n\} \leq 1 - \Phi\Big(a_\lambda + \frac{\lambda n^{\frac{2}{p}}}{C_{\mathrm{emb}}}\Big) = \frac{1}{\sqrt{2\pi}} \int_{a_\lambda + \frac{\lambda n^{\frac{2}{p}}}{C_{\mathrm{emb}}}}^{\infty} e^{-\frac{\chi^2}{2}} d\chi.$$

This yields

$$\limsup_{\lambda \to 0} 2\bar{C}_p\lambda \mathbb{E}N(2\bar{C}_p\lambda, \mathbf{X}(\omega), [0,1]) \leq \limsup_{\lambda \to 0} \frac{2\bar{C}_p\lambda}{\sqrt{2\pi}} \sum_{n=1}^{\infty} \int_{a_\lambda + \frac{\lambda n^{\frac{2}{p}}}{C_{\mathrm{emb}}}}^{\infty} e^{-\frac{\chi^2}{2}} d\chi. \tag{3.44}$$

Unfortunately, the right hand side of (3.44) is infinity for $p \geq 2$. Indeed, observe that $\mathbb{P}(B_\lambda) \to 0$ as $\lambda \to 0$, which implies $a_\lambda \to -\infty$ as $\lambda \to 0$. Thus $a_\lambda + \frac{\lambda n^{\frac{2}{p}}}{C_{\rm emb}} \leq 0$ as long as $n \leq b_\lambda := \Big(C_{\rm emb}\frac{-a_\lambda}{\lambda}\Big)^{\frac{p}{2}}$. As a result,

$$\frac{\lambda}{\sqrt{2\pi}}\sum_{n=1}^\infty \int_{a_\lambda+\frac{\lambda n^{\frac{2}{p}}}{C_{\rm emb}}}^\infty e^{-\frac{\chi^2}{2}}d\chi \geq \frac{\lambda}{\sqrt{2\pi}}\sum_{n=1}^{b_\lambda}\int_0^\infty e^{-\frac{\chi^2}{2}}d\chi \approx \frac{1}{2}\lambda b_\lambda \approx \frac{1}{2}C_{\rm emb}^{\frac{p}{2}}(-a_\lambda)^{\frac{p}{2}}\lambda^{1-\frac{p}{2}} \to \infty$$

as $\lambda \to 0$ provided that $p \geq 2$. Similarly, an attempt to apply the estimate

$$\mathbb{P}\{\omega : N(2\bar{C}_p\lambda, \mathbf{X}(\omega), [0,1]) > n\} \leq \exp\Big\{\delta(\mathbb{P}(B_\lambda))\frac{\lambda n^{\frac{2}{p}}}{C_{\rm emb}} - \frac{\lambda^2 n^{\frac{4}{p}}}{2C_{\rm emb}^2}\Big\}$$

where $\delta(v) = \int_0^\infty \min(1-v, e^{-\frac{t^2v^2}{2}})dt$ as suggested in [29, Formula (4.6), p. 210] also leads to the divergence of a series similar to the one in the right hand side of (3.44). It is therefore a challenging problem on how to estimate the interesting limit (3.40).

Acknowledgments

The author would like to thank the anonymous referee for valuable remarks which led to an improvement of the paper at this present form. This work was supported by the Max Planck Institute for Mathematics in the Science (MIS-Leipzig). A part of this work was done while the author visited Vietnam Institute for Advanced Studies in Mathematics (VIASM) for a three month research stay in summer 2019.

References

[1] L. Arnold. *Random dynamical systems.* Springer, Berlin Heidelberg New York, (1998).

[2] I. Bailleul, S. Riedel, M. Scheutzow. Random dynamical systems, rough paths and rough flows. *J. Differential Equations*, Vol. **262**, (2017), 5792–5823.

[3] T. Caraballo, J. Langa, J. Robinson. Upper semicontinuity of attractors for small random perturbations of dynamical systems. *Commu. Partial Differential Equations*, **23**, (1998), 1557–1581.

[4] A. N. Carvalho, J. A. Langa. An extension of the concept of gradient semigroups which is stable under perturbation. *J. Differential Equations*, **246**, (2009), 2646-2668.

[5] T. Cass, C. Litterer, T. Lyons. Integrability and tail estimates for Gaussian rough differential equations. *Annals of Probability*, Vol. **14**, No. 4, (2013), 3026–3050.

[6] N. D. Cong, L. H. Duc, P. T. Hong. Young differential equations revisited. *J. Dyn. Diff. Equat.*, Vol. **30**, Iss. **4**, (2018), 1921–1943.

[7] N. D. Cong, L. H. Duc, P. T. Hong. Pullback attractors for stochastic Young differential delay equations. *J. Dyn. Diff. Equat.*, (2020), DOI: 10.1007/s10884-020-09894-9.

[8] H. Crauel, F. Flandoli, *Attractors for random dynamical systems.* Probab. Theory Related Fields **100**, No. 3, (1994), 365–393.

[9] H. Crauel, P. Kloeden, Nonautonomous and random attractors. *Jahresber Dtsch. Math-Ver.* **117** (2015), 173–206.

[10] L. H. Duc. Controlled differential equations as rough integrals. *to appear in Pure and Applied Functional Analysis*, (2020). Preprint arXiv:2007.06295.

[11] L. H. Duc, M. J. Garrido-Atienza, A. Neuenkirch, B. Schmalfuß. Exponential stability of stochastic evolution equations driven by small fractional Brownian motion with Hurst parameter in $(\frac{1}{2}, 1)$. *J. Differential Equations*, 264 (2018), 1119-1145.

[12] L. H. Duc, P. T. Hong, N. D. Cong. Asymptotic stability for stochastic dissipative systems with a Hölder noise. *SIAM Journal on Control and Optimization*, **57** (4), (2019), 3046-3071.

[13] P. Friz, M. Hairer. *A course on rough path with an introduction to regularity structure.* Universitext, Vol. **XIV**, Springer, Berlin, 2014.

[14] P. Friz, N. Victoir. *Multidimensional stochastic processes as rough paths: theory and applications.* Cambridge Studies in Advanced Mathematics, 120. Cambridge Unversity Press, Cambridge, (2010).

[15] M. Garrido-Atienza, B. Maslowski, B. Schmalfuß. Random attractors for stochastic equations driven by a fractional Brownian motion. *International Journal of Bifurcation and Chaos*, Vol. 20, No. 9, (2010), 2761–2782.

[16] M. Garrido-Atienza, K. Lu, B. Schmalfuss. *Local pathwise solutions to stochastic evolution equations driven by fractional Brownian motions with Hurst parameters* $H \in (\frac{1}{3}, \frac{1}{2}]$. SIAM Appl. Dyn. Syst., **15**(1), (2016), 625–654.

[17] M. Gubinelli. Controlling rough paths. *J. Funtional Analysis*, **216** (1), (2004), 86–140.

[18] M. Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *The Annals of Probability*, Vol. **33**, (2005), 703–758.

[19] M. Hairer, A. Ohashi. Ergodic theory for sdes with extrinsic memory. *The Annals of Probability*, Vol. **35**, (2007), 1950–1977.

[20] M. Hairer, N. Pillai. Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré*, Vol. **47**, (2011), 601–628

[21] M. Hairer, N. Pillai. Regularity of laws and ergodicity of hypoelliptic stochastic differential equations driven by rough paths. *The Annals of Probability*, Vol. **41**, (2013), 2544–2598.

[22] J. Hale, H. Koçak. *Dynamics and Bifurcations.* Springer-Verlag, (1991).

[23] J. Hale, G. Raugel. Upper semicontinuity of the attractor for a singularly perturbedhyperbolic equation. *J. Differential Equations*, **73**, (1988), 197-214.

[24] R. Hesse, A. Neamtu. Local mild solutions for rough stochastic partial differential equations. *J. Differential Equations*, **267**, (2019), 6480–6538.

[25] Y. Hu, D. Nualart. Rough path analysis via fractional calculus. *Trans. Amer. Math. Soc.*, **361**(5), (2009), 2689–2718.

[26] R. Khasminskii. *Stochastic stability of differential equations.* (2nd edition), Vol. 66, 2011.

[27] H. Keller, B. Schmalfuss. Attractors for stochastic differential equations with nontrivial noise. *Bul. Acad. Ştiinţe Repub. Mold. Mat.* **26**, No. 1, (1998), 43–54.

[28] P. Imkeller, B. Schmalfuss. The conjugacy of stochastic and random differential equations and the existence of global attractors. *J. Dyn. Diff. Equat.* **13**, No. 2, (2001), 215–249.

[29] M. Ledoux. Isoperimetry and Gaussian analysis. In *Lectures on Probability Theory and Statistics (Saint-Flour). Lecture Notes in Math.*, **1648**, Springer, Berlin, (1994), 165–294.

[30] T. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.*, Vol. **14** (2), (1998), 215–310.

[31] T. Lyons, M. Caruana, Th. Lévy. *Differential equations driven by rough paths.* Lecture Notes in Mathematics, Vol. **1908**, Springer, Berlin 2007.

[32] B. Mandelbrot, J. van Ness. Fractional Brownian motion, fractional noises and applications. *SIAM Review*, **4**, No. 10, (1968), 422–437.

[33] X. Mao. *Exponential stability of stochastic differential equations.* NewYork: Marcel Dekker, 1994.

[34] I. Nourdin. *Selected aspects of fractional Brownian motion.* Bocconi University Press, Springer, 2012.

[35] S. Riedel, M. Scheutzow. Rough differential equations with unbounded drift terms. *J. Differential Equations*, Vol. **262**, (2017), 283–312.

[36] A. V. Skorohod. *Asymptotic methods in the theory of stochastic differential equations.* Translations of Mathematical Monographs, American Mathematical Society, Vol. **78**, 1989.

[37] H. J. Sussmann. On the gap between deterministic and stochastic ordinary differential equations. *The Annals of Probability.* **6**, No. 1, (1978), 19–41.

[38] B. Wang. Existence and upper semicontinuity of attractors for stochastic equations with deterministic non-autonomous terms. *Stoch. Dyn.*, **14**(4), (2014), 1450009.