



# Asymptotic Dynamics of Young Differential Equations

Luu Hoang Duc<sup>1,2</sup> · Phan Thanh Hong<sup>3</sup>

Dedicated to Nguyen Dinh Cong on the occasion of his 60th birthday

Received: 31 July 2020 / Revised: 20 July 2021 / Accepted: 3 October 2021 /

Published online: 1 November 2021

© The Author(s) 2021

## Abstract

We provide a unified analytic approach to study the asymptotic dynamics of Young differential equations, using the framework of random dynamical systems and random attractors. Our method helps to generalize recent results (Duc et al. in *J Differ Equ* 264:1119–1145, 2018, *SIAM J Control Optim* 57(4):3046–3071, 2019; Garrido-Atienza et al. in *Int J Bifurc Chaos* 20(9):2761–2782, 2010) on the existence of the global pullback attractors for the generated random dynamical systems. We also prove sufficient conditions for the attractor to be a singleton, thus the pathwise convergence is in both pullback and forward senses.

**Keywords** Stochastic differential equations (SDE) · Young differential equations · Young integrals · Exponential stability · Random dynamical systems · Random attractors

## 1 Introduction

This paper studies the asymptotic behavior of the stochastic differential equation

$$dy_t = [Ay_t + f(y_t)]dt + g(y_t)dZ_t \quad (1.1)$$

where  $A \in \mathbb{R}^{d \times d}$ ,  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $g : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$  are globally Lipschitz continuous functions, and  $Z$  is a two-sided stochastic process with stationary increments such that almost sure all realizations of  $Z$  are in the space  $\mathcal{C}^{p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$  of continuous paths with finite  $p$ -variation norm, for some  $1 \leq p < 2$ . An example for such a process  $Z$  is a fractional Brownian motion  $B^H$  [20] with Hurst index  $H > \frac{1}{2}$ . It is well known that Eq. (1.1) can be solved in the path-wise approach by taking a realization  $x \in \mathcal{C}^{p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$  (which is also

---

✉ Luu Hoang Duc  
duc.luu@mis.mpg.de ; lhduc@math.ac.vn

Phan Thanh Hong  
hongpt@thanglong.edu.vn

<sup>1</sup> Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany

<sup>2</sup> Institute of Mathematics, Viet Nam Academy of Science and Technology, Hanoi, Vietnam

<sup>3</sup> Thang Long University, Hanoi, Vietnam

called a driving path) and considering the Young differential equation

$$dy_t = [Ay_t + f(y_t)]dt + g(y_t)dx_t, \quad t \in \mathbb{R}_+, y_0 \in \mathbb{R}^d. \quad (1.2)$$

This way, system (1.2) is understood in the integral form

$$y_t = y_0 + \int_0^t [Ay_s + f(y_s)]ds + \int_0^t g(y_s)dx_s, \quad \forall t \in \mathbb{R}_+, y_0 \in \mathbb{R}^d, \quad (1.3)$$

where the second integral is understood in the Young sense [24]. The existence and uniqueness theorem for Young differential equations is proved in many versions, e.g. [6,17–19,21,25].

Our aim is to investigate the role of the driving noise in the longterm behavior of system (1.1). Namely we impose assumptions for the drift coefficient so that there exists a unique equilibrium for the deterministic system  $\dot{\mu} = A\mu + f(\mu)$  which is asymptotically stable; and then raise the questions on the asymptotic dynamics of the perturbed system, in particular the existence of stationary states and their asymptotic (stochastic) stability [14] with respect to almost sure convergence.

These questions could be studied in the framework of random dynamical systems [3]. Specifically, results in [11] and recently in [4,9] reveal that the stochastic Young system (1.1) generates a random dynamical system, 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.

We remind the reader of a well-known technique in [15,16,23] to generate RDS and to study random attractors of system (1.2) by a conjugacy transformation  $y_t = \psi_g(\eta_t, z_t)$ , where the semigroup  $\psi_g$  generated by the equation  $\dot{u} = g(u)$  and  $\eta$  is the unique stationary solution of the Langevin equation  $d\eta = -\eta dt + dZ_t$ . The transformed system

$$\dot{z}_t = \left( \frac{\partial \psi_g}{\partial u} \right)^{-1} (\eta_t, z_t) \left[ A\psi_g(\eta_t, z_t) + f(\psi_g(\eta_t, z_t)) + \eta C \psi_g(\eta_t, z_t) \right] \quad (1.4)$$

can then be solved in the pathwise sense and the existence of random attractor for (1.4) is equivalent to the existence of random attractor for the original system. This conjugacy method works in some special cases, particularly if  $g(\cdot)$  is the identity matrix, or more general if  $g(y) = Cy$  for some matrix  $C$  that commutes with  $A$  (see further details in Remark 3.10). For more general cases, the reader is referred to [8,9] and the references therein for recent methods in studying the asymptotic behavior of Young differential equations.

Another approach in [8,11] uses the semigroup technique to estimate the solution norms, which then proves the existence of a random absorbing set that contains a random attractor for the generated random dynamical system. Specifically, thanks to the rule of integration by parts for Young integrals, the "variation of constants" formula for Young differential equations holds (see e.g. [25] or [8]), so that  $y_t$  in Eq. (1.3) satisfies

$$y_t = \Phi(t)y_0 + \int_0^t \Phi(t-s)f(y_s)ds + \int_0^t \Phi(t-s)g(y_s)dx_s, \quad \forall t \geq 0, \quad (1.5)$$

where  $\Phi(t)$  is the semigroup generated by  $A$ . By constructing a suitable stopping times  $\{\tau_k\}_{k \in \mathbb{N}}$ , one can estimate the Hölder norm of  $y$  on interval  $[\tau_n, \tau_{n+1}]$  in the left hand side of (1.5) by the same norm on previous intervals  $[\tau_k, \tau_{k+1}]$  for  $k < n$  following a recurrent relation, thereby can apply the discrete Gronwall Lemma (see Lemma 3.12 in the Appendix). However this construction of stopping times only works under the assumption that the noise is small in the sense that its Hölder seminorm is integrable and can be controlled to be sufficiently small.

In this paper, we propose a different approach in Lemma 3.3, which first estimates the Euclidean norm  $\|y_n\|$  of  $y$  at time  $n$  in (1.5) by applying the continuous Gronwall Lemma, and then estimate Young integrals in the right hand side by the  $p$ -variation norms  $\|y\|_{p\text{-var}, [k, k+1]}$  using Proposition 3.2. Thanks to Theorem 2.4 and its corollaries, these norms  $\|y\|_{p\text{-var}, [k, k+1]}$  are estimated by  $\|y_k\|$ , which leads to a recurrent relation between  $\|y_n\|$  and previous terms  $\|y_k\|$ . As a consequence, one can apply the discrete Gronwall Lemma 3.12 and yield a stability criterion in Theorem 3.4. Therefore, the method works for a general source of noises, and the stability criterion matches the classical one for ordinary differential equations when the effect of driving noise is cleared. Moreover, the same arguments can be applied for stochastic process  $Z$  with lower regularity (for instance  $Z$  is a fractional Brownian motion  $B^H$  with  $\frac{1}{3} < H < \frac{1}{2}$ ), in that case equation (1.3) is no longer a Young equation but should be understood as a rough differential equation and can be solved by Lyon’s rough path theory [18] (see also [10,22]).

The paper is organized as follows. Section 2 is devoted to present preliminaries and main results of the paper, where the norm estimates of the solution of (1.2) is then presented in Sect. 2.1. In Sect. 3.1, we introduce the generation of random dynamical system from the equation (1.1). Using Lemma 3.3, we prove the existence of a global random pullback attractor in Theorem 3.4. Finally in Sect. 3.3, we prove that the attractor is both a pullback and forward singleton attractor if  $g$  is a linear map in Theorem 3.9, or if  $g \in C_b^2$  for small enough Lipschitz constant  $C_g$  in Theorem 3.11.

## 2 Preliminaries and Main Results

Let us first briefly make a survey on Young integrals. Denote by  $\mathcal{C}([a, b], \mathbb{R}^r)$ , for  $r \geq 1$ , the space of all continuous paths  $x : [a, b] \rightarrow \mathbb{R}^r$  equipped with supremum norm  $\|x\|_{\infty, [a, b]} = \sup_{t \in [a, b]} \|x_t\|$ , where  $\|\cdot\|$  is the Euclidean norm of a vector in  $\mathbb{R}^r$ . For  $p \geq 1$  and  $[a, b] \subset \mathbb{R}$ , denote by  $\mathcal{C}^{p\text{-var}}([a, b], \mathbb{R}^r)$  the space of all continuous paths  $x \in \mathcal{C}([a, b], \mathbb{R}^r)$  which is of finite  $p$ -variation, i.e.  $\|x\|_{p\text{-var}, [a, b]} := (\sup_{\Pi(a, b)} \sum_{i=1}^n \|x_{t_{i+1}} - x_{t_i}\|^p)^{1/p} < \infty$  where the supremum is taken over the whole class  $\Pi(a, b)$  of finite partitions  $\Pi = \{a = t_0 < t_1 < \dots < t_n = b\}$  of  $[a, b]$  (see e.g. [10]). Then  $\mathcal{C}^{p\text{-var}}([a, b], \mathbb{R}^r)$ , equipped with the  $p$ -var norm  $\|x\|_{p\text{-var}, [a, b]} := \|x_a\| + \|x\|_{p\text{-var}, [a, b]}$ , is a nonseparable Banach space [10, Theorem 5.25, p. 92]. Also for each  $0 < \alpha < 1$ , denote by  $\mathcal{C}^{\alpha\text{-Hol}}([a, b], \mathbb{R}^r)$  the space of Hölder continuous paths with exponent  $\alpha$  on  $[a, b]$ , and equip it with the norm  $\|x\|_{\alpha\text{-Hol}, [a, b]} := \|x_a\| + \sup_{a \leq s < t \leq b} \frac{\|x_t - x_s\|}{(t-s)^\alpha}$ . Note that for  $\alpha > \frac{1}{p}$ , it holds that  $\mathcal{C}^{\alpha\text{-Hol}}([a, b], \mathbb{R}^r) \subset \mathcal{C}^{p\text{-var}}([a, b], \mathbb{R}^r)$ .

We recall here a result from [6, Lemma 2.1].

**Lemma 2.1** *Let  $x \in \mathcal{C}^{p\text{-var}}([a, b], \mathbb{R}^d)$ ,  $p \geq 1$ . If  $a = a_1 < a_2 < \dots < a_k = b$ , then*

$$\sum_{i=1}^{k-1} \|x\|_{p\text{-var}, [a_i, a_{i+1}]}^p \leq \|x\|_{p\text{-var}, [a_1, a_k]}^p \leq (k-1)^{p-1} \sum_{i=1}^{k-1} \|x\|_{p\text{-var}, [a_i, a_{i+1}]}^p.$$

Now, for  $y \in \mathcal{C}^{q\text{-var}}([a, b], \mathbb{R}^{d \times m})$  and  $x \in \mathcal{C}^{p\text{-var}}([a, b], \mathbb{R}^m)$  with  $\frac{1}{p} + \frac{1}{q} > 1$ , the Young integral  $\int_a^b y_t dx_t$  can be defined as  $\int_a^b y_s dx_s := \lim_{|\Pi| \rightarrow 0} \sum_{[u, v] \in \Pi} y_u(x_v - x_u)$ , where the limit is taken over all the finite partitions  $\Pi$  of  $[a, b]$  with  $|\Pi| := \max_{[u, v] \in \Pi} |v - u|$  (see [24, pp. 264–265]). This integral satisfies the additive property and the so-called *Young-Loeve*

estimate [10, Theorem 6.8, p. 116]

$$\begin{aligned} & \left\| \int_s^t y_u dx_u - y_s [x_t - x_s] \right\| \\ & \leq (1 - 2^{1-\frac{1}{p}-\frac{1}{q}})^{-1} \|y\|_{q\text{-var},[s,t]} \|x\|_{p\text{-var},[s,t]}, \quad \forall [s, t] \subset [a, b]. \end{aligned} \tag{2.1}$$

From now on, we only consider  $q = p$  for convenience. We impose the following assumptions on the coefficients  $A$ ,  $f$  and  $g$  and the driving path  $x$ .

**Assumptions**

- (H<sub>1</sub>)  $A \in \mathbb{R}^{d \times d}$  is a matrix which has all eigenvalues of negative real parts;
- (H<sub>2</sub>)  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $g : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ , are globally Lipschitz continuous functions. In addition,  $g \in C^1$  such that  $D_g$  is also globally Lipschitz continuous. Denote by  $C_f, C_g$  the Lipschitz constants of  $f$  and  $g$  respectively;
- (H<sub>3</sub>) for a given  $p \in (1, 2)$ ,  $Z_t$  is a two-sided stochastic process with stationary increments such that almost sure all realizations belong to the space  $C^{p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$  and that

$$\Gamma(p) := \left( E \|Z\|_{p\text{-var},[-1,1]}^p \right)^{\frac{1}{p}} < \infty. \tag{2.2}$$

For instance,  $Z$  could be an  $m$ -dimensional fractional Brownian motion  $B^H$  [20] with Hurst exponent  $H \in (\frac{1}{2}, 1)$ , i.e. a family of centered Gaussian processes  $B^H = \{B_t^H\}$ ,  $t \in \mathbb{R}$  or  $\mathbb{R}_+$  with continuous sample paths and the covariance function

$$R_H(s, t) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}), \quad \forall t, s \in \mathbb{R}.$$

Assumption (H<sub>1</sub>) ensures that the semigroup  $\Phi(t) = e^{At}$ ,  $t \in \mathbb{R}$  generated by  $A$  satisfies the following properties.

**Proposition 2.2** *Assume that  $A$  has all eigenvalues of negative real parts. Then there exist constants  $C_A \geq 1, \lambda_A > 0$  such that the generated semigroup  $\Phi(t) = e^{At}$  satisfies*

$$\|\Phi\|_{\infty,[a,b]} \leq C_A e^{-\lambda_A a}, \tag{2.3}$$

$$\|\Phi\|_{p\text{-var},[a,b]} \leq |A| C_A e^{-\lambda_A a} (b - a), \quad \forall 0 \leq a < b, \tag{2.4}$$

in which  $|A| := \sup_{\|x\|=1} \frac{\|Ax\|}{\|x\|}$ .

**Proof** The first inequality is due to [1, Chapter 1, §3]. The second one is followed from the mean value theorem

$$\|\Phi(u) - \Phi(v)\| = \left\| \int_u^v A\Phi(s)ds \right\| \leq \int_u^v |A| C_A e^{-\lambda_A s} ds \leq |A| C_A e^{-\lambda_A u} (v - u),$$

for any  $u < v$  in  $[a, b]$  where  $e^{-\lambda_A \cdot}$  is a decreasing function. □

Our main results (Theorems 3.4, 3.9, 3.11) could be summarized as follows.

**Theorem 2.3** *Assume that the system (1.1) satisfies the assumptions H<sub>1</sub> – H<sub>3</sub>, and further that  $\lambda_A > C_f C_A$ , where  $\lambda_A$  and  $C_A$  are given from (2.3), (2.4). If*

$$\lambda_A - C_A C_f > C_A (1 + |A|) e^{\lambda_A + 2(|A| + C_f)} \left\{ \left[ 2(K + 1) C_g \Gamma(p) \right]^p + \left[ 2(K + 1) C_g \Gamma(p) \right] \right\}, \tag{2.5}$$

where  $\Gamma(p)$  is defined in (2.2) and  $K$  in (2.7), then the generated random dynamical system  $\varphi$  of (1.1) possesses a pullback attractor  $\mathcal{A}(x)$ . Moreover, in case  $g(y) = Cy + g(0)$  is a linear map satisfying (2.5) or in case  $g \in C_b^2$  with the Lipschitz constant  $C_g$  small enough, this attractor is a singleton, i.e.  $\mathcal{A}(x) = \{a(x)\}$  a.s., thus the pathwise convergence is in both the pullback and forward directions.

For the convenience of the readers, we introduce some notations and constants which are used throughout the paper.

$$L := |A| + C_f, \quad L_f := C_A C_f, \quad \lambda := \lambda_A - L_f; \tag{2.6}$$

$$K := (1 - 2^{1-\frac{2}{p}})^{-1}, \quad \alpha := \log\left(1 + \frac{1}{K+1}\right); \tag{2.7}$$

$$M_0 := \frac{\|f(0)\|}{L} + \frac{\|g(0)\|}{(K+1)C_g}; \tag{2.8}$$

$$M_1 := K C_A e^{\lambda A} (1 + |A|); \tag{2.9}$$

$$M_2 := \max\left\{C_A \frac{e^\lambda - 1}{\lambda}, M_1 C_g \left(\frac{1}{L} + \frac{1}{(K+1)C_g}\right), M_1 C_g\right\} \max\{\|f(0)\|, \|g(0)\|\}; \tag{2.10}$$

$$\hat{G} := C_A e^{\lambda A} (1 + |A|) e^{4L} \left\{ \left[2(K+1)C_g \Gamma(p)\right]^p + \left[2(K+1)C_g \Gamma(p)\right] \right\}. \tag{2.11}$$

### 2.1 Solution Estimates

In this preparatory subsection we are going to estimate several norms of the solution. To do that, the idea is to evaluate the norms of the solution on a number of consecutive small intervals. Here we would like to construct, for any  $\gamma > 0$  and any given interval  $[a, b]$ , a sequence of greedy times  $\{\tau_k(\gamma)\}_{k \in \mathbb{N}}$  as follows (see e.g. [5,6,8])

$$\tau_0 = a, \quad \tau_{k+1}(\gamma) := \inf\{t > \tau_k(\gamma) : \|x\|_{p\text{-var}, [\tau_k(\gamma), t]} = \gamma\} \wedge b. \tag{2.12}$$

Define

$$N = N_{\gamma, [a, b]}(x) := \sup\{k \in \mathbb{N}, \tau_k(\gamma) \leq b\}, \tag{2.13}$$

then due to the superadditivity of  $\|x\|_{p\text{-var}, [s, t]}^p$

$$\begin{aligned} N - 1 &\leq \sum_{k=0}^{N-2} \gamma^{-p} \|x\|_{p\text{-var}, [\tau_k, \tau_{k+1}]}^p \leq \gamma^{-p} \|x\|_{p\text{-var}, [\tau_0, \tau_{N-1}]}^p \\ &\leq \gamma^{-p} \|x\|_{p\text{-var}, [a, b]}^p, \end{aligned}$$

$$\text{which yields } N \leq 1 + \gamma^{-p} \|x\|_{p\text{-var}, [a, b]}^p. \tag{2.14}$$

From now on, we fix  $p \in (1, 2)$  and  $\gamma := \frac{1}{2(K+1)C_g}$ , and write in short  $N_{[a, b]}(x)$  to specify the dependence of  $N$  on  $x$  and the interval  $[a, b]$ .

We assume throughout this section that the assumption **(H<sub>2</sub>)**, **(H<sub>3</sub>)** are satisfied. The following theorem presents a standard method to estimate the  $p$ -variation and the supremum norms of the solution of (1.2), by using the continuous Gronwall lemma and a discretization scheme with the greedy times (2.12).

**Theorem 2.4** *There exists a unique solution to (1.2) for any initial value, whose supremum and  $p$ -variation norms are estimated as follows*

$$\|y\|_{\infty,[a,b]} \leq \left[ \|y_a\| + M_0 N_{[a,b]}(x) \right] e^{\alpha N_{[a,b]}(x) + 2L(b-a)}, \tag{2.15}$$

$$\|y\|_{p\text{-var},[a,b]} \leq \left[ \|y_a\| + M_0 N_{[a,b]}(x) \right] e^{\alpha N_{[a,b]}(x) + 2L(b-a)} N_{[a,b]}^{\frac{p-1}{p}}(x), \tag{2.16}$$

where  $L, \alpha$  and  $M_0$  are given by (2.6), (2.7) and (2.8) respectively.

**Proof** There are similar versions of Theorem 2.4 in [19, Proposition 1] for Young equations or in [5, Lemma 4.5] and [22, Theorem 3.1] for rough differential equations with bounded  $g$ , thus we will only sketch out the proof here for the benefit of the readers. To prove (2.15), we use the fact that  $\|g(y)\|_{p\text{-var},[s,t]} \leq C_g \|y\|_{p\text{-var},[s,t]}$  and apply (2.1) with  $K$  in (2.7) to derive

$$\|y_t - y_s\| \leq \int_s^t (L \|y_u\| + \|f(0)\|) du + \|x\|_{p\text{-var},[s,t]} \left( \|g(y_s)\| + K C_g \|y\|_{p\text{-var},[s,t]} \right)$$

so that 
$$\|y\|_{p\text{-var},[s,t]} \leq \int_s^t L \|y\|_{p\text{-var},[s,u]} du + (\|f(0)\| + L \|y_s\|)(t - s) + \|x\|_{p\text{-var},[s,t]} \left( \|g(y_s)\| + (K + 1) C_g \|y\|_{p\text{-var},[s,t]} \right).$$

As a result,

$$\|y\|_{p\text{-var},[s,t]} \leq \int_s^t 2L \|y\|_{p\text{-var},[s,u]} du + 2(\|f(0)\| + L \|y_s\|)(t - s) + 2 \|x\|_{p\text{-var},[s,t]} \|g(y_s)\| \tag{2.17}$$

whenever  $(K + 1)C_g \|x\|_{p\text{-var},[s,t]} \leq \frac{1}{2}$ . Applying the continuous Gronwall Lemma [2, Lemma 6.1, p 89] for  $\|y\|_{p\text{-var},[s,t]}$  yields

$$\begin{aligned} \|y\|_{p\text{-var},[s,t]} &\leq 2(\|f(0)\| + L \|y_s\|)(t - s) + 2 \|x\|_{p\text{-var},[s,t]} \|g(y_s)\| \\ &\quad + \int_s^t 2L e^{2L(t-u)} \left[ 2(\|f(0)\| + L \|y_s\|)(u - s) + 2 \|x\|_{p\text{-var},[s,u]} \|g(y_s)\| \right] du \\ &\leq \left( M_0 + \frac{K + 2}{K + 1} \|y_s\| \right) e^{2L(t-s)} - \|y_s\| \end{aligned} \tag{2.18}$$

whenever  $\|x\|_{p\text{-var},[s,t]} \leq \gamma = \frac{1}{2(K+1)C_g}$ . Now construct the sequence of greedy times  $\{\tau_k = \tau_k(\gamma)\}_{k \in \mathbb{N}}$  on interval  $[a, b]$  as in (2.12), it follows from induction that

$$\begin{aligned} \|y_{\tau_{k+1}}\| &\leq \|y\|_{\infty, [\tau_k, \tau_{k+1}]} \leq \|y\|_{p\text{-var}, [\tau_k, \tau_{k+1}]} \leq \left( e^{\alpha \|y_{\tau_k}\|} + M_0 \right) e^{2L(\tau_{k+1} - \tau_k)} \\ &\leq \left[ \|y_a\| + M_0(k + 1) \right] e^{\alpha(k+1) + 2L(\tau_{k+1} - \tau_0)}, \quad \forall k = 0, \dots, N_{[a,b]}(x) - 1, \end{aligned}$$

which proves (2.15) since  $\tau_{N_{[a,b]}(x)} = b$ . On the other hand, it follows from inequality of  $p$ -variation seminorm in Lemmas 2.1 and (2.18) that for all  $k = 0, \dots, N_{[a,b]}(x) - 1$ ,

$$\begin{aligned} \|y\|_{p\text{-var},[a,b]} &\leq N_{[a,b]}^{\frac{p-1}{p}}(x) \sum_{k=0}^{N_{[a,b]}(x)-1} \left\{ \|y_{\tau_k}\| \left( e^{\alpha + 2L(\tau_{k+1} - \tau_k)} - 1 \right) + M_0 e^{2L(\tau_{k+1} - \tau_k)} \right\} \\ &\leq N_{[a,b]}^{\frac{p-1}{p}}(x) \left[ \|y_a\| + M_0 N_{[a,b]}(x) \right] e^{\alpha N_{[a,b]}(x) + 2L(b-a)} - \|y_a\| \end{aligned}$$

which proves (2.16). □

By the same arguments, we can prove the following results.

**Corollary 2.5** *If in addition  $g$  is bounded by  $\|g\|_\infty < \infty$ , then*

$$\|y\|_{p\text{-var},[a,b]} \leq \left[ \|y_a\| + \left( \frac{\|f(0)\|}{L} \vee 2\|g\|_\infty \right) (1 + \|x\|_{p\text{-var},[a,b]}) N_{[a,b]}(x) \right] e^{2L(b-a)} N_{[a,b]}^{\frac{p-1}{p}}(x), \tag{2.19}$$

in which  $a \vee b := \max\{a, b\}$ .

**Corollary 2.6** *The following estimate holds*

$$\begin{aligned} \|y\|_{p\text{-var},[a,b]} \leq & \left[ \|y_a\| + \left( \frac{\|f(0)\|}{L} \vee 2\|g(0)\| \right) (1 + \|x\|_{p\text{-var},[a,b]}) N_{[a,b]}(x) \right] \times \\ & \times e^{\alpha N_{[a,b]}(x) + 2L(b-a)} N_{[a,b]}^{\frac{p-1}{p}}(x). \end{aligned} \tag{2.20}$$

The lemma below is useful in evaluating the difference of two solutions of Eq. (1.2). The proof is similar to [6, Lemma 3.1] and will be omitted here.

**Lemma 2.7** *Let  $y^1, y^2$  be two solution of (1.2). Assign*

$$Q(t) = Q(t, y^1, y^2) := g(y_t^1) - g(y_t^2), \quad t \geq 0$$

where  $g$  satisfies  $(H_2)$ .

(i) *If in addition,  $D_g$  is of Lipschitz continuity with Lipschitz constant  $C'_g$ , then*

$$\|Q\|_{p\text{-var},[u,v]} \leq C_g \|y^1 - y^2\|_{p\text{-var},[u,v]} + C'_g \|y^1 - y^2\|_\infty \|y^1\|_{p\text{-var},[u,v]}. \tag{2.21}$$

(ii) *If  $g$  is a linear map, then  $\|Q\|_{p\text{-var},[u,v]} \leq C_g \|y^1 - y^2\|_{p\text{-var},[u,v]}$ .*

Thanks to Lemma 2.7, the difference of two solutions of (1.2) can be estimated in  $p$ -var norm as follows.

**Corollary 2.8** *Let  $y^1, y^2$  be two solutions of (1.2) and assign  $z_t = y_t^2 - y_t^1$  for all  $t \geq 0$ .*

(i) *If  $D_g$  is of Lipschitz continuity with Lipschitz constant  $C'_g$  then*

$$\|z\|_{p\text{-var},[a,b]} \leq \|z_a\| (N'_{[a,b]}(x))^{\frac{p-1}{p}} 2^{N'_{[a,b]}(x)} e^{2L(b-a)}, \quad \forall a \leq b \tag{2.22}$$

in which

$$N'_{[a,b]}(x) \leq 1 + [2(K + 1)(C_g \vee C'_g)]^p \|x\|_{p\text{-var},[a,b]}^p (1 + \|y^1\|_{p\text{-var},[a,b]})^p. \tag{2.23}$$

(ii) *If in addition  $g$  is a linear map then*

$$\|z\|_{p\text{-var},[a,b]} \leq \|z_a\| e^{\alpha N_{[a,b]}(x) + 2L(b-a)}. \tag{2.24}$$

**Proof** The proof use similar arguments to the proof of Theorem 2.4, thus it will be omitted here. The readers are referred to [19, Proposition 1], [6, Theorem 3.9] for similar versions. □

**Remark 2.9** It follows from (2.14) that

$$N_{[a,b]}^{\frac{p-1}{p}}(x) \leq \left( 1 + [2(K + 1)C_g]^p \|x\|_{p\text{-var},[a,b]}^p \right)^{\frac{p-1}{p}}$$

$$\begin{aligned} &\leq 1 + [2(K + 1)C_g]^{p-1} \|x\|_{p\text{-var},[a,b]}^{p-1}, \\ N_{[a,b]}^{\frac{2p-1}{p}}(x) &\leq \left(1 + [2(K + 1)C_g]^p \|x\|_{p\text{-var},[a,b]}^p\right)^{\frac{2p-1}{p}} \\ &\leq 2^{\frac{p-1}{p}} \left(1 + [2(K + 1)C_g]^{2p-1} \|x\|_{p\text{-var},[a,b]}^{2p-1}\right). \end{aligned}$$

As a result, the norm estimates (2.16), (2.19), (2.20) have the same form

$$\|y\|_{p\text{-var},[a,b]} \leq \|y_a\| \Lambda_1(x, [a, b]) + \Lambda_2(x, [a, b]) \tag{2.25}$$

in which  $\Lambda_i(x, [a, b])$  are functions of  $\|x\|_{p\text{-var},[a,b]}$ . Similarly, (2.22) (for a fixed solution  $y^1$ ) and (2.24) can also be rewritten in the form (2.25) with  $\Lambda_2 \equiv 0$ .

In the following, let  $y$  be a solution of (1.2) on  $[a, b] \subset \mathbb{R}^+$  and  $\mu$  be the solution of the corresponding deterministic system, i.e

$$\dot{\mu}_t = A\mu_t + f(\mu_t), t \in [a, b] \tag{2.26}$$

with the same initial condition  $\mu_a = y_a$ . Assign  $h_t := y_t - \mu_t$ . The following result, which is used in studying singleton attractors in Theorem 3.11, estimates the norms of  $h$  with the initial condition  $\|y_a\|$ , up to a fractional order.

**Corollary 2.10** *Assume that  $g$  is bounded. Then for a fixed constant  $\beta = \frac{1}{p} \in (\frac{1}{2}, 1)$ , there exists for each interval  $[a, b]$  a constant  $D$  depending on  $b - a$  such that*

$$\|h\|_{\infty,[a,b]} \leq D \left(\|y_a\|^\beta + 1\right) \|x\|_{p\text{-var},[a,b]} N_{[a,b]}(x), \tag{2.27}$$

$$\|h\|_{p\text{-var},[a,b]} \leq D \left(\|y_a\|^\beta + 1\right) \|x\|_{p\text{-var},[a,b]} N_{[a,b]}^{\frac{2p-1}{p}}(x). \tag{2.28}$$

**Proof** The proof follows similar steps to [13, Proposition 4.6] with only small modifications in estimates for  $p$ -variation norms and in usage of the continuous Gronwall lemma. To sketch out the proof, we first observe from (2.26) with  $r = b - a$  that

$$\|\mu_t - \mu_s\| \leq \int_s^t (\|f(0)\| + L\|\mu_u\|) du \leq e^{Lr} (L\|\mu_a\| + \|f(0)\|)(t - s), \quad a \leq s \leq t \leq b. \tag{2.29}$$

Next, it follows from  $\mathbf{H}_2$  and the boundedness of  $g$  by  $\|g\|_\infty$  that

$$\|h_t - h_s\| \leq \int_s^t L\|h_u\| du + \|g\|_\infty \|x\|_{p\text{-var},[s,t]} + K \|x\|_{p\text{-var},[s,t]} \|g(h + \mu)\|_{p\text{-var},[s,t]}. \tag{2.30}$$

Observe that due to the boundedness of  $g$ ,

$$\begin{aligned} \|g(h_t + \mu_t) - g(h_s + \mu_s)\| &\leq C_g \|h_t - h_s\| + (2\|g\|_\infty \vee C_g) \|\mu_t - \mu_s\|^\beta \\ &\leq C_g \|h_t - h_s\| + D(1 + \|\mu_a\|^\beta)(t - s)^\beta, \quad \forall a \leq s < t \leq b \end{aligned}$$

where the last estimate is due to (2.29) and  $D$  is a generic constant depending on  $b - a$ . This leads to

$$\|g(h + \mu)\|_{p\text{-var},[s,t]} \leq C_g \|h\|_{p\text{-var},[s,t]} + D(1 + \|\mu_a\|^\beta)(t - s)^\beta. \tag{2.31}$$

Replacing (2.31) into (2.30) yields



$$\begin{aligned} & \|h\|_{p\text{-var},[s,t]} \\ & \leq \int_s^t L \|h\|_{p\text{-var},[s,u]} du + L \|h_s\|(t-s) + \left[ \|g\|_\infty \vee D(1 + \|\mu_a\|^\beta) \right] \|x\|_{p\text{-var},[s,t]} \\ & \quad + KC_g \|x\|_{p\text{-var},[s,t]} \|h\|_{p\text{-var},[s,t]} \end{aligned}$$

which is similar to (2.17). Using similar arguments to the proof of Theorem 2.4 and taking into account (2.19), we conclude that

$$\|h\|_{\infty,[a,b]} \leq e^{2Lr} \left[ \|h_a\| + D(1 + \|\mu_a\|^\beta) \|x\|_{p\text{-var},[a,b]} N_{[a,b]}(x) \right],$$

for a generic constant  $D$ . Finally, (2.27) is derived since  $h_a = 0$ . The estimate (2.28) is obtained similarly. □

### 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 Young equation (1.1). Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space equipped with a so-called *metric dynamical system*  $\theta$ , which is a measurable mapping  $\theta : \mathbb{R} \times \Omega \rightarrow \Omega$  such that  $\theta_t : \Omega \rightarrow \Omega$  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*  $\varphi : \mathbb{R} \times \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d, (t, \omega, y_0) \mapsto \varphi(t, \omega)y_0$  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 [3].

In our scenario, denote by  $C^{0,p\text{-var}}([a, b], \mathbb{R}^m)$  the closure of  $C^\infty([a, b], \mathbb{R}^m)$  in  $C^{p\text{-var}}([a, b], \mathbb{R}^m)$ , and by  $C^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$  the space of all  $x : \mathbb{R} \rightarrow \mathbb{R}^m$  such that  $x|_I \in C^{0,p\text{-var}}(I, \mathbb{R}^m)$  for each compact interval  $I \subset \mathbb{R}$ . Then equip  $C^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$  with the compact open topology given by the  $p$ -variation norm, i.e. the topology generated by the metric:

$$d_p(x_1, x_2) := \sum_{k \geq 1} \frac{1}{2^k} (\|x_1 - x_2\|_{p\text{-var},[-k,k]} \wedge 1).$$

Assign

$$\Omega := C_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m) := \{x \in C^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m) \mid x_0 = 0\},$$

and equip with the Borel  $\sigma$ -algebra  $\mathcal{F}$ . Note that for  $x \in C_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m), \|x\|_{p\text{-var},I}$  and  $\|x\|_{p\text{-var},I}$  are equivalent norms for every compact interval  $I$  containing 0.

To equip this measurable space  $(\Omega, \mathcal{F})$  with a metric dynamical system, consider a stochastic process  $\bar{Z}$  defined on a probability space  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$  with realizations in  $(C_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m), \mathcal{F})$ . Assume further that  $\bar{Z}$  has stationary increments. Denote by  $\theta$  the *Wiener shift*

$$(\theta_t x)_\cdot = x_{t+\cdot} - x_t, \quad \forall t \in \mathbb{R}, x \in C_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m).$$

It is easy to check that  $\theta$  forms a continuous (and thus measurable) dynamical system  $(\theta_t)_{t \in \mathbb{R}}$  on  $(C_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m), \mathcal{F})$ . Moreover, the Young integral satisfies the shift property with respect

to  $\theta$ , i.e.

$$\int_a^b y_u dx_u = \int_{a-r}^{b-r} y_{r+u} d(\theta_r x)_u \tag{3.2}$$

(see details in [6, p. 1941]). It follows from [4, Theorem 5] that, there exists a probability  $\mathbb{P}$  on  $(\Omega, \mathcal{F}) = (\mathcal{C}_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m), \mathcal{F})$  that is invariant under  $\theta$ , and the so-called *diagonal process*  $Z : \mathbb{R} \times \Omega \rightarrow \mathbb{R}^m, Z(t, x) = x_t$  for all  $t \in \mathbb{R}, x \in \Omega$ , such that  $Z$  has the same law with  $\bar{Z}$  and satisfies the *helix property*:

$$Z_{t+s}(x) = Z_s(x) + Z_t(\theta_s x), \forall x \in \Omega, t, s \in \mathbb{R}.$$

Such stochastic process  $Z$  has also stationary increments and almost all of its realizations belongs to  $\mathcal{C}_0^{0,p\text{-var}}(\mathbb{R}, \mathbb{R}^m)$ . It is important to note that the existence of  $\bar{Z}$  is necessary to construct the diagonal process  $Z$ .

When dealing with fractional Brownian motion [20], we can start with the space  $\mathcal{C}_0(\mathbb{R}, \mathbb{R}^m)$  of continuous functions on  $\mathbb{R}$  vanishing at zero, with the Borel  $\sigma$ -algebra  $\mathcal{F}$ , and the Wiener shift and the Wiener probability  $\mathbb{P}$ , and then follow [12, Theorem 1] to construct an invariant probability measure  $\mathbb{P}^H = B^H \mathbb{P}$  on the subspace  $\mathcal{C}^v$  such that  $B^H \circ \theta = \theta \circ B^H$ . It can be proved that  $\theta$  is ergodic (see [12]).

Under this circumstance, if we assume further that (2.2) is satisfied, then it follows from Birkhorff ergodic theorem that

$$\Gamma(x, p) := \limsup_{n \rightarrow \infty} \left( \frac{1}{n} \sum_{k=1}^n \|\theta_{-k} x\|_{p\text{-var}, [-1,1]}^p \right)^{\frac{1}{p}} = \Gamma(p) \tag{3.3}$$

for almost all realizations  $x_t = Z_t(x)$  of  $Z$ . Particularly, in case  $Z = B^H = (B_1^H, \dots, B_m^H)$  where  $B_i^H$  are scalar fractional Brownian motions (not necessarily independent), we can apply Lemma 2.1 in [6] with the estimate in [9, Lemma 4.1 (iii), p.14] to obtain that  $\Gamma(p) < \infty$ .

**Proposition 3.1** *The system (1.1) generates a continuous random dynamical system.*

**Proof** The proof follows directly from [4] and [6, Section 4.2], so we only sketch out the proof here. First, for each fixed driving path  $x \in \Omega$ , Eq. (1.1) is solved in the path-wise sense by Young equations (1.2) and (1.3) for the starting point  $y_0$  at time 0 in the forward time if  $t > 0$ , or in the backward time for  $t < 0$  by the backward Young equation

$$y_t = y_0 + \int_t^0 [Ay_s + f(y_s)] ds + \int_t^0 g(y_s) dx_s, \quad \forall t \leq 0, y_0 \in \mathbb{R}^d,$$

(see details in [6, Theorem 3.8]). Define the mapping  $\varphi(t, x)y_0 := y_t(x, y_0)$ , for  $t \in \mathbb{R}, x \in \Omega, y_0 \in \mathbb{R}^d$ , which is the pathwise solution of (1.1), then it follows from the existence and uniqueness theorem and the Wiener shift property (3.2) that  $\varphi$  satisfies the cocycle property (3.1) (see [6, Subsection 4.2] and [11] for more details). Also, it is proved in [6, Theorem 3.9] that the solution  $y_t(x, y_0)$  is continuous w.r.t.  $(t, x, y_0)$ , hence given a probability structure on  $\Omega$ ,  $\varphi$  is a continuous random dynamical system. □

### 3.2 Existence of Pullback Attractors

Given a random dynamical system  $\varphi$  on  $\mathbb{R}^d$ , we follow [7], [3, Chapter 9] to present the notion of random pullback attractor. Recall that a set  $\hat{M} := \{M(x)\}_{x \in \Omega}$  a *random set*, if  $y \mapsto d(y|M(x))$  is  $\mathcal{F}$ -measurable for each  $y \in \mathbb{R}^d$ , where  $d(E|F) = \sup\{\inf\{d(y, z)|z \in$

$F\}|y \in E\}$  for  $E, F$  are nonempty subset of  $\mathbb{R}^d$  and  $d(y|E) = d(\{y\}|E)$ . An universe  $\mathcal{D}$  is a family of random sets which is closed w.r.t. inclusions (i.e. if  $\hat{D}_1 \in \mathcal{D}$  and  $\hat{D}_2 \subset \hat{D}_1$  then  $\hat{D}_2 \in \mathcal{D}$ ).

In our setting, we define the universe  $\mathcal{D}$  to be a family of *tempered* random sets  $D(x)$ , which means the following: A random variable  $\rho(x) > 0$  is called *tempered* if it satisfies

$$\lim_{t \rightarrow \pm\infty} \frac{1}{t} \log^+ \rho(\theta_t x) = 0, \quad \text{a.s.} \tag{3.4}$$

(see e.g. [3, pp. 164, 386]) which, by [15, p. 220]), is equivalent to the sub-exponential growth

$$\lim_{t \rightarrow \pm\infty} e^{-ct|t|} \rho(\theta_t x) = 0 \quad \text{a.s.} \quad \forall c > 0.$$

A random set  $D(x)$  is called *tempered* if it is contained in a ball  $B(0, \rho(x))$  a.s., where the radius  $\rho(x)$  is a tempered random variable.

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

$$\lim_{t \rightarrow \infty} d(\varphi(t, \theta_{-t} x)\hat{D}(\theta_{-t} x)|\mathcal{A}(x)) = 0. \tag{3.5}$$

$\mathcal{A}$  is called a *forward random attractor* in  $\mathcal{D}$ , if  $\mathcal{A}$  is invariant and attracts any closed random set  $\hat{D} \in \mathcal{D}$  in the forward sense, i.e.

$$\lim_{t \rightarrow \infty} d(\varphi(t, x)\hat{D}(x)|\mathcal{A}(\theta_t x)) = 0. \tag{3.6}$$

The existence of a random pullback attractor follows from the existence of a random pullback absorbing set (see [7, Theorem 3]). A random set  $\mathcal{B} \in \mathcal{D}$  is called *pullback absorbing* in a universe  $\mathcal{D}$  if  $\mathcal{B}$  absorbs all sets in  $\mathcal{D}$ , i.e. for any  $\hat{D} \in \mathcal{D}$ , there exists a time  $t_0 = t_0(x, \hat{D})$  such that

$$\varphi(t, \theta_{-t} x)\hat{D}(\theta_{-t} x) \subset \mathcal{B}(x), \quad \text{for all } t \geq t_0. \tag{3.7}$$

Given a universe  $\mathcal{D}$  and a random compact pullback absorbing set  $\mathcal{B} \in \mathcal{D}$ , there exists a unique random pullback attractor in  $\mathcal{D}$ , given by

$$\mathcal{A}(x) = \bigcap_{s \geq 0} \overline{\bigcup_{t \geq s} \varphi(t, \theta_{-t} x)\mathcal{B}(\theta_{-t} x)}. \tag{3.8}$$

Since the rule of integration by parts for Young integral is proved in [25], the "variation of constants" formula for Young differential equations holds (see e.g. [8]), i.e.  $y_t$  satisfies

$$y_t = \Phi(t - a)y_a + \int_a^t \Phi(t - s)f(y_s)ds + \int_a^t \Phi(t - s)g(y_s)dx_s, \quad \forall t \geq a. \tag{3.9}$$

We need the following auxiliary results.

**Proposition 3.2** *Given (2.3) and (2.4), the following estimate holds: for any  $0 \leq a < b \leq c$*

$$\begin{aligned} & \left\| \int_a^b \Phi(c - s)g(y_s)dx_s \right\| \\ & \leq KC_A \left[ 1 + |A|(b - a) \right] \|x\|_{p\text{-var}, [a, b]} e^{-\lambda_A(c-b)} \left[ C_g \|y\|_{p\text{-var}, [a, b]} + \|g(0)\| \right]. \end{aligned} \tag{3.10}$$

**Proof** The proof follows directly from (2.3) and (2.4) as follows

$$\begin{aligned} & \left\| \int_a^b \Phi(c-s)g(y_s)dx_s \right\| \\ & \leq \|x\|_{p\text{-var},[a,b]} \left( \|\Phi(c-a)g(y_a)\| + K \|\Phi(c-\cdot)g(y)\|_{p\text{-var},[a,b]} \right) \\ & \leq \|x\|_{p\text{-var},[a,b]} \left\{ \|\Phi(c-a)\| \|g(y_a)\| \right. \\ & \quad \left. + K \left( \|\Phi(c-\cdot)\|_{p\text{-var},[a,b]} \|g(y)\|_{\infty,[a,b]} + \|\Phi(c-\cdot)\|_{\infty,[a,b]} \|g(y)\|_{p\text{-var},[a,b]} \right) \right\} \\ & \leq KC_A \|x\|_{p\text{-var},[a,b]} e^{-\lambda_A(c-b)} \times \\ & \quad \times \left[ C_g \|y_a\| + \|g(0)\| + |A|(b-a)(C_g \|y\|_{\infty,[a,b]} + \|g(0)\|) + C_g \|y\|_{p\text{-var},[a,b]} \right] \\ & \leq KC_A \left[ 1 + |A|(b-a) \right] \|x\|_{p\text{-var},[a,b]} e^{-\lambda_A(c-b)} \left[ C_g \|y\|_{p\text{-var},[a,b]} + \|g(0)\| \right]. \end{aligned}$$

□

The following lemma is the crucial technique of this paper.

**Lemma 3.3** Assume that  $y_t$  satisfies (3.9). Then for any  $n \geq 0$ ,

$$\begin{aligned} \|y_t\|e^{\lambda t} & \leq C_A \|y_0\| + \frac{C_A}{\lambda_A - L_f} \|f(0)\| \left( e^{\lambda t} - 1 \right) \\ & \quad + \sum_{k=0}^n e^{\lambda k} KC_A (1 + |A|) \|x\|_{p\text{-var},\Delta_k} e^{\lambda k} \left[ C_g \|y\|_{p\text{-var},\Delta_k} + \|g(0)\| \right], \forall t \in \Delta_n, \end{aligned} \tag{3.11}$$

where  $\Delta_k := [k, k + 1]$ ,  $L_f$  and  $\lambda$  are defined in (2.6).

**Proof** First, for any  $t \in [n, n + 1]$ , it follows from (2.3) and the global Lipschitz continuity of  $f$  that

$$\begin{aligned} \|y_t\| & \leq \|\Phi(t)y_0\| + \int_0^t \|\Phi(t-s)f(y_s)\|ds + \left\| \int_0^t \Phi(t-s)g(y_s)dx_s \right\| \\ & \leq C_A e^{-\lambda_A t} \|y_0\| + \int_0^t C_A e^{-\lambda_A(t-s)} \left( C_f \|y_s\| + \|f(0)\| \right) ds + \left\| \int_0^t \Phi(t-s)g(y_s)dx_s \right\| \\ & \leq C_A e^{-\lambda_A t} \|y_0\| + \frac{C_A}{\lambda_A} \|f(0)\| (1 - e^{-\lambda_A t}) + C_A C_f \int_0^t e^{-\lambda_A(t-s)} \|y_s\| ds + \beta_t, \end{aligned}$$

where  $\beta_t := \left\| \int_0^t \Phi(t-s)g(y_s)dx_s \right\|$ . Multiplying both sides of the above inequality with  $e^{\lambda_A t}$  yields

$$\|y_t\|e^{\lambda_A t} \leq C_A \|y_0\| + \frac{C_A}{\lambda_A} \|f(0)\| (e^{\lambda_A t} - 1) + \beta_t e^{\lambda_A t} + C_A C_f \int_0^t e^{\lambda_A s} \|y_s\| ds.$$

By applying the continuous Gronwall Lemma [2, Lemma 6.1, p 89], we obtain

$$\begin{aligned} \|y_t\|e^{\lambda_A t} & \leq C_A \|y_0\| + \frac{C_A}{\lambda_A} \|f(0)\| (e^{\lambda_A t} - 1) + \beta_t e^{\lambda_A t} \\ & \quad + \int_0^t L_f e^{L_f(t-s)} \left[ C_A \|y_0\| + \frac{C_A}{\lambda_A} \|f(0)\| (e^{\lambda_A s} - 1) + \beta_s e^{\lambda_A s} \right] ds. \end{aligned}$$

Once again, multiplying both sides of the above inequality with  $e^{-L_f t}$  yields

$$\begin{aligned} \|y_t\|e^{\lambda t} &\leq C_A \|y_0\|e^{-L_f t} + \frac{C_A}{\lambda_A} \|f(0)\| \left( e^{\lambda t} - e^{-L_f t} \right) + \beta_t e^{\lambda t} \\ &\quad + \int_0^t L_f e^{-L_f s} \left[ C_A \|y_0\| + \frac{C_A}{\lambda_A} \|f(0)\| (e^{\lambda s} - 1) + \beta_s e^{\lambda s} \right] ds \\ &\leq C_A \|y_0\| + \frac{C_A}{\lambda} \|f(0)\| \left( e^{\lambda t} - 1 \right) + \beta_t e^{\lambda t} + \int_0^t L_f \beta_s e^{\lambda s} ds. \end{aligned} \tag{3.12}$$

Next, assign  $p([a, b]) := KC_A \left[ 1 + |A|(b - a) \right] \|x\|_{p\text{-var}, [a, b]} \left[ C_g \|y\|_{p\text{-var}, [a, b]} + \|g(0)\| \right]$ , and apply (3.10) in Proposition 3.2, it follows that for all  $s \leq t$

$$\begin{aligned} \beta_s e^{\lambda s} &= e^{\lambda s} \left\| \int_0^s \Phi(s - u)g(y_u)dx_u \right\| \\ &\leq e^{\lambda s} \sum_{k=0}^{\lfloor s \rfloor - 1} \left\| \int_{\Delta_k} \Phi(s - u)g(y_u)dx_u \right\| + e^{\lambda s} \left\| \int_{\lfloor s \rfloor}^s \Phi(s - u)g(y_u)dx_u \right\| \\ &\leq e^{\lambda s} \sum_{k=0}^{\lfloor s \rfloor - 1} e^{-\lambda_A(s-k-1)} p(\Delta_k) + e^{\lambda s} p(\lfloor s \rfloor, s) \\ &\leq \sum_{k=0}^{\lfloor s \rfloor} e^{\lambda s} e^{-\lambda_A(s-k-1)} p(\Delta_k) \leq \sum_{k=0}^{\lfloor s \rfloor} e^{\lambda_A} e^{\lambda k} e^{-L_f(s-k)} p(\Delta_k). \end{aligned} \tag{3.13}$$

By replacing (3.13) into (3.12) we obtain

$$\begin{aligned} \|y_t\|e^{\lambda t} &\leq C_A \|y_0\| + \frac{C_A}{\lambda} \|f(0)\| \left( e^{\lambda t} - 1 \right) \\ &\quad + \sum_{k=0}^n e^{\lambda_A} e^{\lambda k} e^{-L_f(t-k)} p(\Delta_k) + L_f \int_0^t \sum_{k=0}^{\lfloor s \rfloor} e^{\lambda_A} e^{\lambda k} e^{-L_f(s-k)} p(\Delta_k) ds \\ &\leq C_A \|y_0\| + \frac{C_A}{\lambda} \|f(0)\| \left( e^{\lambda t} - 1 \right) + \sum_{k=0}^n e^{\lambda_A} e^{\lambda k} p(\Delta_k) \left( e^{-L_f(t-k)} \right. \\ &\quad \left. + \int_k^t L_f e^{-L_f(s-k)} ds \right) \\ &\leq C_A \|y_0\| + \frac{C_A}{\lambda} \|f(0)\| \left( e^{\lambda t} - 1 \right) + \sum_{k=0}^n e^{\lambda_A} e^{\lambda k} p(\Delta_k), \quad \forall t \in [n, n + 1) \end{aligned}$$

where we use the fact that

$$e^{-L_f(t-a)} + \int_a^t L_f e^{-L_f(s-a)} ds = 1, \quad \forall t \geq a. \tag{3.14}$$

The continuity of  $y$  at  $t = n + 1$  then proves (3.11). □

We now formulate the first main result of the paper.

**Theorem 3.4** *Under the assumptions  $(H_1) - (H_3)$ , assume further that  $\lambda_A > C_f C_A$ , where  $\lambda_A$  and  $C_A$  are given from (2.3), (2.4). If the criterion (2.5)*

$$\lambda_A - C_A C_f > C_A (1 + |A|) e^{\lambda_A + 2(|A| + C_f)} \left\{ \left[ 2(K + 1)C_g \Gamma(p) \right]^p + \left[ 2(K + 1)C_g \Gamma(p) \right] \right\}$$

holds, where  $\Gamma(p)$  is defined in (2.2), then the generated random dynamical system  $\varphi$  of system (1.1) possesses a pullback attractor  $\mathcal{A}$ .

**Proof Step 1.** To begin, we rewrite the estimate (2.15) in the short form, using (2.25) in Remark 2.9

$$\|y\|_{p\text{-var}, \Delta_k} \leq \|y_k\| \Lambda_1(x, \Delta_k) + M_0 \Lambda_2(x, \Delta_k) \tag{3.15}$$

where  $\Delta_k = [k, k + 1]$ ,  $M_0$  is given by (2.8) and

$$\begin{aligned} \Lambda_1(x, [a, b]) &:= \left( 1 + [2(K + 1)C_g]^{p-1} \|x\|_{p\text{-var}, [a, b]}^{p-1} \right) F(x, [a, b]), \\ \Lambda_2(x, [a, b]) &:= 2^{\frac{p-1}{p}} \left( 1 + [2(K + 1)C_g]^{2p-1} \|x\|_{p\text{-var}, [a, b]}^{2p-1} \right) F(x, [a, b]), \\ F(x, [a, b]) &:= \exp \left\{ \alpha \left( 1 + [2(K + 1)C_g]^p \|x\|_{p\text{-var}, [a, b]}^p \right) + 2L(b - a) \right\}, \end{aligned} \tag{3.16}$$

for  $L, \alpha$  in (2.6) and (2.7) respectively. Replacing (3.15) into (3.11) in Lemma 3.3 and using  $M_1, M_2$  in (2.9) and (2.10), we obtain

$$\begin{aligned} \|y_n\| e^{\lambda n} &\leq C_A \|y_0\| + (e^{\lambda n} - 1) \frac{C_A \|f(0)\|}{\lambda} \\ &\quad + e^{\lambda A} K C_A (1 + |A|) \sum_{k=0}^{n-1} e^{\lambda k} \|x\|_{p\text{-var}, \Delta_k} \left[ C_g \left( \|y_k\| \Lambda_1(x, \Delta_k) \right. \right. \\ &\quad \left. \left. + M_0 \Lambda_2(x, \Delta_k) \right) + \|g(0)\| \right] \\ &\leq C_A \|y_0\| + M_1 C_g \sum_{k=0}^{n-1} \|x\|_{p\text{-var}, \Delta_k} \Lambda_1(x, \Delta_k) e^{\lambda k} \|y_k\| \\ &\quad + M_2 \sum_{k=0}^{n-1} e^{\lambda k} \left[ 1 + \|x\|_{p\text{-var}, \Delta_k} \left( 1 + \Lambda_2(x, \Delta_k) \right) \right]. \end{aligned} \tag{3.17}$$

Assign  $a := C_A \|y_0\|$ ,  $u_k := e^{\lambda k} \|y_k\|$  and

$$G(x, [a, b]) := \|x\|_{p\text{-var}, [a, b]} \Lambda_1(x, [a, b]), \tag{3.18}$$

$$H(x, [a, b]) := 1 + \|x\|_{p\text{-var}, [a, b]} \left( 1 + \Lambda_2(x, [a, b]) \right) \tag{3.19}$$

for all  $k \geq 0$ , where  $\Lambda_1, \Lambda_2$  are given by (3.16). Observe that (3.17) has the form

$$u_n \leq a + M_1 C_g \sum_{k=0}^{n-1} G(x, \Delta_k) u_k + M_2 \sum_{k=0}^{n-1} e^{\lambda k} H(x, \Delta_k).$$

We are in the position to apply the discrete Gronwall Lemma 3.12, so that

$$\begin{aligned} \|y_n(x, y_0)\| &\leq C_A \|y_0\| e^{-\lambda n} \prod_{k=0}^{n-1} \left[ 1 + M_1 C_g G(x, \Delta_k) \right] \\ &\quad + M_2 \sum_{k=0}^{n-1} e^{-\lambda(n-k)} H(x, \Delta_k) \prod_{j=k+1}^{n-1} \left[ 1 + M_1 C_g G(x, \Delta_j) \right], \quad \forall n \geq 1. \end{aligned} \tag{3.20}$$

**Step 2.** Next, for any  $t \in [n, n + 1]$ , due to (2.15) and (2.14), we can write

$$\|y_t(x, y_0)\| \leq \|y_n(x, y_0)\|F(x, \Delta_n) + M_0\Lambda_0(x, \Delta_n) \tag{3.21}$$

$$\text{where } \Lambda_0(x, [a, b]) := \left(1 + [2(K + 1)C_g]^p \|x\|_{p\text{-var}, [a, b]}^p\right) F(x, [a, b]). \tag{3.22}$$

Consequently, replacing  $x$  with  $\theta_{-t}x$  in (3.21) and using (3.20) yields

$$\begin{aligned} & \|y_t(\theta_{-t}x, y_0(\theta_{-t}x))\| \\ & \leq C_A \|y_0(\theta_{-t}x)\| e^{-\lambda n} F(\theta_{-t}x, \Delta_n) \prod_{k=0}^{n-1} \left[1 + M_1 C_g G(\theta_{k-t}x, [0, 1])\right] + M_0 \Lambda_0(\theta_{-t}x, \Delta_n) \\ & \quad + M_2 \sum_{k=0}^{n-1} e^{-\lambda(n-k)} F(\theta_{-t}x, \Delta_n) H(\theta_{k-t}x, [0, 1]) \prod_{j=k+1}^{n-1} \left[1 + M_1 C_g G(\theta_{j-t}x, [0, 1])\right] \\ & \leq C_A F(x, [-1, 1]) \|y_0(\theta_{-t}x)\| e^{-\lambda n} \sup_{\epsilon \in [0, 1]} \prod_{k=1}^n \left[1 + M_1 C_g G(\theta_{-k}x, [-\epsilon, 1 - \epsilon])\right] \\ & \quad + M_0 \Lambda_0(x, [-1, 1]) \\ & \quad + M_2 F(x, [-1, 1]) \sup_{\epsilon \in [0, 1]} \sum_{k=1}^n e^{-\lambda k} H(\theta_{-k}x, [-\epsilon, 1 - \epsilon]) \\ & \quad \quad \prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-j}x, [-\epsilon, 1 - \epsilon])\right] \\ & \leq C_A F(x, [-1, 1]) \|y_0(\theta_{-t}x)\| e^{-\lambda n} \sup_{\epsilon \in [0, 1]} \prod_{k=1}^n \left[1 + M_1 C_g G(\theta_{-k}x, [-\epsilon, 1 - \epsilon])\right] \\ & \quad + M_0 \Lambda_0(x, [-1, 1]) + M_2 F(x, [-1, 1]) b(x), \tag{3.23} \end{aligned}$$

$$\text{where } b(x) := \sup_{\epsilon \in [0, 1]} \sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-k}x, [-\epsilon, 1 - \epsilon])$$

$$\prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-j}x, [-\epsilon, 1 - \epsilon])\right], \tag{3.24}$$

( $b(x)$  can take value  $\infty$ ). Applying the inequality  $\log(1 + ae^b) \leq a + b$  for  $a, b \geq 0$  and using (3.18), (3.15), we obtain

$$\begin{aligned} & \log \left(1 + M_1 C_g G(x, [-\epsilon, 1 - \epsilon])\right) \\ & \leq M_1 C_g e^{\alpha + 2L} \left[ \|x\|_{p\text{-var}, [-1, 1]} + [2(K + 1)C_g]^{p-1} \|x\|_{p\text{-var}, [-1, 1]}^p \right] \\ & \quad + \frac{[2(K + 1)C_g]^p \|x\|_{p\text{-var}, [-1, 1]}^p}{K + 1} \\ & \leq \left[ M_1 e^{\alpha + 2L} + 2 \right] [2(K + 1)]^{p-1} C_g^p \|x\|_{p\text{-var}, [-1, 1]}^p + M_1 e^{\alpha + 4L} C_g \|x\|_{p\text{-var}, [-1, 1]} \\ & \leq C_A e^{\lambda A + 2L} (1 + |A|) \left\{ \left[2(K + 1)C_g \|x\|_{p\text{-var}, [-1, 1]} \right]^p + \left[2(K + 1)C_g \|x\|_{p\text{-var}, [-1, 1]} \right] \right\}. \end{aligned}$$

Together with (3.3) and (2.9), it follows that for a.s. all  $x$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \sup_{\epsilon \in [0, 1]} \prod_{k=1}^n \left[1 + M_1 C_g G(\theta_{-k}x, [-\epsilon, 1 - \epsilon])\right] \right\}$$

$$\leq C_A e^{\lambda_A + 2L} (1 + |A|) \left\{ \left[ 2(K + 1)C_g \Gamma(p) \right]^p + \left[ 2(K + 1)C_g \Gamma(p) \right] \right\} = \hat{G}, \tag{3.25}$$

where  $\hat{G}$  is defined in (2.11) and is also the right hand side of (2.5), and  $L$  is defined in (2.6). Hence for  $t \in \Delta_n$  with  $0 < \delta < \frac{1}{2}(\lambda - \hat{G})$  and  $n \geq n_0$  large enough

$$e^{(-\delta + \hat{G})n} \leq \sup_{\varepsilon \in [0,1]} \prod_{k=0}^{n-1} \left[ 1 + M_1 C_g G(\theta_{-k}x, [-\varepsilon, 1 - \varepsilon]) \right] \leq e^{(\delta + \hat{G})n}.$$

Starting from any point  $y_0(\theta_{-t}x) \in D(\theta_{-t}x)$  which is tempered, there exists, due to (3.4), an  $n_0$  independent of  $y_0$  large enough such that for any  $n \geq n_0$  and any  $t \in [n, n + 1]$

$$\|y_t(\theta_{-t}x, y_0(\theta_{-t}x))\| \leq C_A e^{\lambda_A} \|y_0(\theta_{-t}x)\| F(x, [-1, 1]) \exp \left\{ - \left( \lambda - \hat{G} - \delta \right) n \right\} \tag{3.26}$$

$$\begin{aligned} &+ M_0 \Lambda_0(x, [-1, 1]) + M_2 F(x, [-1, 1]) b(x) \\ &\leq 1 + M_2 F(x, [-1, 1]) b(x) + M_0 \Lambda_0(x, [-1, 1]) =: \hat{b}(x) \end{aligned} \tag{3.27}$$

where  $F, \Lambda_0$  are given in (3.16) and (3.22). In addition, it follows from (2.14) and the inequality  $\log(1 + ab) \leq \log(1 + a) + \log b$  for all  $a \geq 0, b \geq 1$ , that

$$\begin{aligned} \log \hat{b}(x) &\leq \log[1 + M_2 F(x, [-1, 1])] + \log[1 + b(x)] \\ &\quad + \log \left\{ 1 + M_0 \left[ 1 + \left( 2(K + 1)C_g \right)^p \|x\|_{p\text{-var},[-1,1]}^p F(x, [-1, 1]) \right] \right\} \\ &\leq D + \log[1 + b(x)] + [2(K + 1)C_g]^p \|x\|_{p\text{-var},[-1,1]}^p + 2 \log F(x, [-1, 1]) \\ &\leq D(1 + \|x\|_{p\text{-var},[-1,1]}^p) + \log[1 + b(x)] \end{aligned} \tag{3.28}$$

where  $D$  is a constant.

**Step 3.** Notice that (3.3) implies  $\lim_{n \rightarrow \infty} \frac{\|\theta_{-n}x\|_{p\text{-var},[-1,1]}}{n} = 0$ . The proof would be complete if one can prove Proposition 3.5 below, that  $b(x)$  is finite and tempered a.s. Indeed, assume that Proposition 3.5 holds, then by applying [3, Lemma 4.1.2], we obtain the temperedness of  $\hat{b}(x)$  in the sense of (3.4). We conclude that there exists a compact absorbing set  $\mathcal{B}(x) = \bar{B}(0, \hat{b}(x))$  and thus a pullback attractor  $\mathcal{A}(x)$  for system (1.2) in the form of (3.8).  $\square$

To complete the proof of Theorem 3.4, we now formulate and prove Proposition 3.5.

**Proposition 3.5** *Assume that (2.5) holds. Then  $b(x)$  defined in (3.24)*

$$b(x) := \sup_{\varepsilon \in [0,1]} \sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-k}x, [-\varepsilon, 1 - \varepsilon]) \prod_{j=1}^{k-1} \left[ 1 + M_1 C_g G(\theta_{-j}x, [-\varepsilon, 1 - \varepsilon]) \right], \quad \forall x \in \Omega$$

*is finite and tempered a.s. [in the sense of (3.4)].*

**Proof** From the definition of  $H$  in (3.19), by similar computations as in (3.28) using the integrability of  $\|x\|_{p\text{-var},[-1,1]}$  it is easy to prove that  $\log H(x, [-1, 1])$  is integrable, thus

$$\limsup_{n \rightarrow \infty} \frac{\log H(\theta_{-n}x, [-1, 1])}{n} = 0 \quad \text{a.s.}$$

Hence, under condition (2.5), there exists for each  $0 < 2\delta < \lambda - \hat{G}$  an  $n_0 = n_0(\delta, x)$  such that for all  $n \geq n_0$ ,

$$e^{-\delta n} \leq H(\theta_{-n}x, [-1, 1]) \leq e^{\delta n}.$$



Consequently,

$$\begin{aligned}
 b(x) &\leq \sum_{k=1}^{n_0-1} e^{-\lambda k} H(\theta_{-k}x, [-1, 1]) \sup_{\varepsilon \in [0,1]} \prod_{j=1}^{k-1} \left(1 + M_1 C_g G(\theta_{-j}x, [-\varepsilon, 1 - \varepsilon])\right) \\
 &\quad + \sum_{k=n_0}^{\infty} e^{-(\lambda-2\delta-\hat{G})k} \\
 &\leq \sum_{k=1}^{n_0-1} e^{-\lambda k} H(\theta_{-k}x, [-1, 1]) \prod_{j=1}^{k-1} \left(1 + M_1 C_g G(\theta_{-j}x, [-1, 1])\right) + \frac{e^{-(\lambda-2\delta-\hat{G})n_0}}{1 - e^{-(\lambda-2\delta-\hat{G})}}
 \end{aligned}$$

which is finite a.s. Moreover, for each fixed  $x$ ,  $\|x\|_{p\text{-var},[s,t]}$  is continuous w.r.t  $(s, t)$  on  $\{(s, t) \in \mathbb{R}^2 | s \leq t\}$  (see [10, Proposition 5.8, p. 80]). Therefore,  $G(x, [-\varepsilon, 1 - \varepsilon])$  and  $H(x, [-\varepsilon, 1 - \varepsilon])$  are continuous functions of  $\varepsilon$ . Since  $b(x) \leq b^*(x)$ , the series

$$\sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-k}x, [-\varepsilon, 1 - \varepsilon]) \prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-j}x, [-\varepsilon, 1 - \varepsilon])\right]$$

converges uniformly w.r.t.  $\varepsilon \in [0, 1]$ , thus the series is also continuous w.r.t.  $\varepsilon \in [0, 1]$ . Hence the supremum in the definition of  $b(x)$  in (3.24) can be taken for rational  $\varepsilon$ , which proves  $b(x)$  to be a random variable on  $\Omega$ .

To prove the temperdness of  $b$ , observe that for each  $t \in [n, n + 1]$

$$\begin{aligned}
 b(\theta_t x) &= b(\theta_n \theta_{t-n} x) \\
 &= \sup_{\varepsilon \in [0,1]} \sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-k+n} \theta_{-\varepsilon+t-n} x, [0, 1]) \prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-j+n} \theta_{-\varepsilon+t-n} x, [0, 1])\right] \\
 &\leq \sup_{\varepsilon \in [-1,1]} \sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-k+n} x, [-\varepsilon, 1 - \varepsilon]) \prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-j+n} x, [-\varepsilon, 1 - \varepsilon])\right] \\
 &\leq \max\{b(\theta_n x), b(\theta_{n+1} x)\}.
 \end{aligned}$$

For  $n > 0$ ,

$$\begin{aligned}
 b(\theta_{-n} x) &= \sup_{\varepsilon \in [0,1]} \sum_{k=1}^{\infty} e^{-\lambda k} H(\theta_{-(k+n)} x, [-\varepsilon, 1 - \varepsilon]) \prod_{j=1}^{k-1} \left[1 + M_1 C_g G(\theta_{-(j+n)} x, [-\varepsilon, 1 - \varepsilon])\right] \\
 &\leq e^{\lambda n} \sup_{\varepsilon \in [0,1]} \sum_{k=1}^{\infty} e^{-\lambda(n+k)} H(\theta_{-(k+n)} x, [-\varepsilon, 1 - \varepsilon]) \\
 &\quad \prod_{j=1}^{n+k-1} \left[1 + M_1 C_g G(\theta_{-(j+n)} x, [-\varepsilon, 1 - \varepsilon])\right] \\
 &\leq e^{\lambda n} b(x).
 \end{aligned}$$

Therefore

$$\limsup_{n \rightarrow +\infty} \frac{\log b(\theta_{-n} x)}{n} \leq \lambda < \infty.$$

Applying [3, Proposition 4.1.3(i), p. 165] we conclude that

$$\limsup_{t \rightarrow -\infty} \frac{\log b(\theta_t x)}{|t|} = \limsup_{t \rightarrow +\infty} \frac{\log b(\theta_t x)}{t} = 0,$$

i.e.  $b$  is tempered. □

**Corollary 3.6** *Assume that  $f(0) = g(0) = 0$  so that  $y \equiv 0$  is a solution of (1.1). Then under the assumptions of Theorem 3.4 with condition (2.5), the random attractor  $\mathcal{A}(x)$  is the fixed point 0.*

**Proof** Using (3.26) and the fact that  $M_0 = M_2 = 0$  if  $f(0) = g(0) = 0$ , we obtain

$$\|y_t(\theta_{-t}x, y_0)\| \leq C_A e^{\lambda_A} F(x, [-1, 1]) \|y_0(\theta_{-t}x)\| \exp \left\{ - \left( \lambda - \hat{G} - \delta \right) n \right\} \quad (3.29)$$

for  $t \in \Delta_n$ . It follows that all other solutions converge exponentially in the pullback sense to the trivial solution, which plays the role of the global pullback attractor. □

**Remark 3.7** In [8,11] the authors consider a Hilbert space  $V$  together with a covariance operator  $Q$  on  $V$  such that  $Q$  is of a trace-class, i.e. for a complete orthonormal basis  $(e_i)_{i \in \mathbb{N}}$  of  $V$ , there exists a sequence of nonnegative numbers  $(q_i)_{i \in \mathbb{N}}$  such that  $\text{tr}(Q) := \sum_{i=1}^{\infty} q_i < \infty$ . A  $V$ -valued fractional Brownian motion  $B^H = \sum_{i=1}^{\infty} \sqrt{q_i} \beta_i^H e_i$  is then considered, where  $(\beta_i^H)_{i \in \mathbb{N}}$  are stochastically independent scalar fractional Brownian motions of the same Hurst exponent  $H$ . The authors then develop the semigroup method to estimate the Hölder norm of  $y$  on intervals  $\tau_k, \tau_{k+1}$  where  $\tau_k$  is a sequence of stopping times

$$\tau_0 = 0, \tau_{k+1} - \tau_k + \|x\|_{\beta, [\tau_k, \tau_{k+1}]} = \mu$$

for some  $\mu \in (0, 1)$  and  $\beta > \frac{1}{p}$ , which leads to the estimate of the exponent as

$$- \left( \lambda_A - C(C_A, \mu) e^{\lambda_A \mu} \max\{C_f, C_g\} \frac{n}{\tau_n} \right) \tau_n, \quad (3.30)$$

where  $C(C_A, \mu)$  is a constant depending on  $C_A, \mu$ . It is then proved that there exists  $\liminf_{n \rightarrow \infty} \frac{\tau_n}{n} = \frac{1}{d}$ , where  $d = d(\mu, \text{tr}(Q))$  is a constant depending on the moment of the stochastic noise. As such the exponent is estimated as

$$- \left( \lambda_A - C(C_A, \mu) e^{\lambda_A \mu} \max\{C_f, C_g\} d(\mu, \text{tr}(Q)) \right). \quad (3.31)$$

However, it is technically required from the stopping time analysis (see [8, Section 4]) that the stochastic noise has to be small in the sense that the trace  $\text{tr}(Q) = \sum_{i=1}^{\infty} q_i$  must be controlled as small as possible. In addition, in case the noise is diminished, i.e.  $g \equiv 0$ , (3.31) reduces to a very rough criterion for exponential stability of the ordinary differential equation

$$C_f \leq \frac{1}{C(C_A, \mu) d(\mu, \text{tr}(Q))} \lambda_A e^{-\lambda_A}.$$

In contrast, our method uses the greedy time sequence in Theorem 2.4, so that later we can work with the simpler (regular) discretization scheme without constructing additional stopping time sequence. Also in Lemma 3.3 we apply first the continuous Gronwall lemma in (3.12) in order to clear the role of the drift coefficient  $f$ . Then by using (2.15) to give a direct estimate of  $y_k$ , we are able apply the discrete Gronwall Lemma directly and obtain a very explicit criterion.

The left and the right hand sides of criterion (2.5)

$$\lambda_A - C_A C_f > C_A (1 + |A|) e^{\lambda_A + 2(|A| + C_f)} \left\{ \left[ 2(K + 1) C_g \Gamma(p) \right]^p + \left[ 2(K + 1) C_g \Gamma(p) \right] \right\}$$

can be interpreted as, respectively, the decay rate of the drift term and the intensity of the diffusion term, where the term  $e^{\lambda_A + 4(|A| + C_f)}$  is the unavoidable effect of the discretization

scheme. Criterion (2.5) is therefore a better generalization of the classical criterion for stability of ordinary differential equations, and is satisfied if either  $C_g$  or  $\Gamma(p)$  is sufficiently small. In particular, when  $g \equiv 0$ , (2.5) reduces to  $\lambda_A > C_A C_f$ , which matches to the classical result.

### 3.3 Singleton Attractors

In the rest of the paper, we would like to study sufficient conditions for the global attractor to consist of only one point, as seen, for instance, in Corollary 3.6. First, the answer is affirmative for  $g$  of linear form, as proved in [9] for dissipative systems. Here we also present a similar version using the semigroup method.

To begin, let  $y^1, y^2$  be two solutions of (1.2) and assign  $z_t = y_t^2 - y_t^1$  for all  $t \geq 0$ . Similar to (3.9),  $z$  satisfies

$$z_t = \Phi(t-a)z_a + \int_a^t \Phi(t-s) \left[ f(z_s + y_s^1) - f(y_s^1) \right] ds + \int_a^t \Phi(t-s) Q(s, z_s) dx_s, \quad \forall t \geq a, \tag{3.32}$$

where  $Q(s, z_s) = g(z_s + y_s^1) - g(y_s^1)$ . Observe that by similar computations to (3.10), it is easy to prove that

$$\begin{aligned} \left\| \int_a^b \Phi(c-s) Q(s, z_s) dx_s \right\| &\leq K C_A (C_g \vee C'_g) \left[ 1 + |A|(b-a) \right] e^{-\lambda_A(c-b)} \|x\|_{p\text{-var},[a,b]} \times \\ &\quad \times (1 + \|y^1\|_{p\text{-var},[a,b]}) \|z\|_{p\text{-var},[a,b]}. \end{aligned} \tag{3.33}$$

We need the following auxiliary result.

**Lemma 3.8** *Assume that all the conditions in Theorem 3.4 are satisfied. Let  $y^1, y^2$  be two solutions of (1.2) and assign  $z_t = y_t^2 - y_t^1$  for all  $t \geq 0$ .*

(i) *If  $D_g$  is of Lipschitz continuity with Lipschitz constant  $C'_g$ , then*

$$\begin{aligned} e^{\lambda n} \|z_n\| &\leq C_A \|z_0\| + e^{\lambda A} K C_A (1 + |A|) (C_g \vee C'_g) \sum_{k=0}^{n-1} \|x\|_{p\text{-var}, \Delta_k} e^{\lambda k} \\ &\quad \left( 1 + \|y^1\|_{p\text{-var}, \Delta_k} \right) \|z\|_{p\text{-var}, \Delta_k}. \end{aligned} \tag{3.34}$$

(ii) *If  $g(y) = Cy + g(0)$  then*

$$e^{\lambda n} \|z_n\| \leq C_A \|z_0\| + e^{\lambda A} K C_A (1 + |A|) C_g \sum_{k=0}^{n-1} \|x\|_{p\text{-var}, \Delta_k} e^{\lambda k} \|z\|_{p\text{-var}, \Delta_k}. \tag{3.35}$$

**Proof** The arguments follow the proof of Lemma 3.3 step by step, and apply Lemma 2.7, Proposition 3.2 to obtain the estimate

$$e^{\lambda A t} \|z_t\| \leq C_A \|z_0\| + e^{\lambda A t} \beta_t + L_f \int_0^t \left( \|z_s\| + e^{\lambda A s} \beta_s \right) e^{L_f(t-s)} ds,$$

where  $\beta_t = \left\| \int_0^t \Phi(t-s) Q(s, z_s) dx_s \right\|$  is estimated using (3.33). The rest will be omitted.  $\square$

**Theorem 3.9** *Assume that  $g(y) = Cy + g(0)$  is a linear map so that  $C_g = |C|$ . Then under the condition (2.5),*

$$\lambda_A - C_A C_f > C_A (1 + |A|) e^{\lambda_A + 2(|A| + C_f)} \left\{ \left[ 2(K + 1) C_g \Gamma(p) \right]^p + \left[ 2(K + 1) C_g \Gamma(p) \right] \right\},$$

the pullback attractor is a singleton, i.e.  $\mathcal{A}(x) = \{a(x)\}$  almost surely. Moreover, it is also a forward singleton attractor.

**Proof** The existence of the pullback attractor  $\mathcal{A}$  is followed by Theorem 3.4. Take any two points  $a_1, a_2 \in \mathcal{A}(x)$ . For a given  $n \in \mathbb{N}$ , assign  $x^* := \theta_{-n}x$  and consider the equation

$$dy_t = [Ay_t + f(y_t)]dt + g(y_t)dx_t^*. \tag{3.36}$$

Due to the invariance of  $\mathcal{A}$  under the flow, there exist  $b_1, b_2 \in \mathcal{A}(x^*)$  such that  $a_i = y_n(x^*, b_i)$ . Put  $z_t = z_t(x^*) := y_t(x^*, b_1) - y_t(x^*, b_2)$  then  $z_n(x^*) = a_1 - a_2$ . By applying Lemma 3.8 with  $x$  replaced by  $x^*$ , and using Lemma 2.7 (ii), we can rewrite the estimates in (3.35) as

$$e^{\lambda n} \|z_n\| \leq C_A \|z_0\| + M_1 C_g \sum_{k=0}^{n-1} \|x^*\|_{p\text{-var}, \Delta_k} e^{\lambda k} \|z\|_{p\text{-var}, \Delta_k}. \tag{3.37}$$

Meanwhile, using Corollary 2.8(ii) and Remark 2.9, with  $M_0$  in (2.8) now equal to zero, we obtain

$$\|z\|_{p\text{-var}, [a, b]} \leq \|z_a\| \Lambda_1(x^*, [a, b]),$$

in which  $\Lambda_1$  defined in (3.16). As a result, (3.37) has the form

$$\begin{aligned} e^{\lambda n} \|z_n\| &\leq C_A \|z_0\| + M_1 C_g \sum_{k=0}^{n-1} \|x^*\|_{p\text{-var}, [0, 1]} \Lambda_1(x^*, [0, 1]) e^{\lambda k} \|z_k\| \\ &\leq C_A \|z_0\| + M_1 C_g \sum_{k=0}^{n-1} G(x^*, \Delta_k) e^{\lambda k} \|z_k\|, \end{aligned}$$

where  $G$  is defined in (3.18). Now applying the discrete Gronwall Lemma 3.12, we conclude that

$$e^{\lambda n} \|z_n\| \leq C_A \|z_0\| \prod_{k=0}^{n-1} [1 + M_1 C_g G(x^*, \Delta_k)]. \tag{3.38}$$

Similar to (3.25)

$$\begin{aligned} &\log(1 + M_1 C_g G(x^*, \Delta_k)) \\ &\leq C_A e^{\lambda_A + 2L} (1 + |A|) \left[ \left( 2(K + 1) C_g \|\theta_{k-n}x\|_{p\text{-var}, [0, 1]} \right. \right. \\ &\quad \left. \left. + \left( 2(K + 1) C_g \|\theta_{k-n}x\|_{p\text{-var}, [0, 1]} \right)^p \right) \right]. \end{aligned}$$

Therefore it follows from (3.38) that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \|z_n\| &\leq -\lambda + \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \log[1 + M_1 C_g G(x^*, \Delta_k)] \\ &\leq -\lambda + C_A e^{\lambda_A + 2L} (1 + |A|) \left[ 2(K + 1) C_g \Gamma(p) + \left( 2(K + 1) C_g \right)^p \Gamma(p)^p \right] \\ &< 0 \end{aligned}$$

under the condition (2.5). This follows that  $\lim_{n \rightarrow \infty} \|a_1 - a_2\| = 0$  and  $\mathcal{A}$  is a one point set almost surely. Similar arguments in the forward direction ( $x^*$  is replaced by  $x$ ) also prove that  $\mathcal{A}$  is a forward singleton attractor almost surely. □

**Remark 3.10** As pointed out in the Introduction section, if we use the conjugacy transformation (developed in [15,16,23]) of the form  $y_t = e^{C\eta_t} z_t$ , where the semigroup  $e^{Ct}$  is generated by the equation  $\dot{u} = Cu$  and  $\eta$  is the unique stationary solution of the Langevin equation  $d\eta = -\eta dt + dZ_t$  (with  $Z$  is a scalar stochastic process), then the transformed system has the form

$$\dot{z}_t = e^{-C\eta_t} \left[ A e^{C\eta_t} z_t + f(e^{C\eta_t} z_t) + \eta_t C e^{C\eta_t} z_t \right].$$

However, even in the simplest case  $f \equiv 0$ , there is no effective method to study the asymptotic stability of the non-autonomous linear stochastic system

$$\dot{z}_t = \left( e^{-C\eta_t} A e^{C\eta_t} + \eta_t C \right) z_t. \tag{3.39}$$

An exception is when  $A$  and  $C$  are commute, since we could reduce system (3.39) in the form

$$\dot{z}_t = \left( A + \eta_t C \right) z_t,$$

thereby solve it explicitly as

$$z_t = z_0 \exp \left\{ At + C \int_0^t \eta_s ds \right\} = z_0 \exp \left\{ At - C(\eta_t - \eta_0 - Z_t + Z_0) \right\}.$$

In this case, the exponential stability is proved using the fact that  $\exp\{-C(\eta_t - \eta_0 - Z_t + Z_0)\}$  is tempered. However, since  $A$  and  $C$  are in general not commute, we can not apply the conjugacy transformation but should instead use our method described in Theorems 3.4 and 3.9.

Next, motivated by [13], we consider the case in which  $g \in C_b^2$  and  $C_g$  is also the Lipschitz constant of  $Dg$ . Notice that our conditions for  $A$  and  $f$  can be compared similarly to the dissipativity condition in [13]. However, unlike the probabilistic conclusion of existence and uniqueness of a stationary measure in [13], we go a further step by proving that for  $C_g$  small enough, the random attractor is indeed a singleton, thus the convergence to the attractor is in the pathwise sense and of exponential rate.

**Theorem 3.11** *Assume that  $g \in C_b^2$  with  $\|g\|_\infty < \infty$ , and denote by  $C_g$  the Lipschitz constant of  $g$  and  $Dg$ . Assume further that  $\lambda_A > C_A C_f$  and*

$$E \left\| Z \right\|_{p\text{-var}, [-r, r]}^{\frac{4p(p+1)}{p-1}} < \infty, \quad \forall r > 0. \tag{3.40}$$

*Then system (1.2) possesses a pullback attractor. Moreover, there exists a  $\delta > 0$  small enough such that for any  $C_g \leq \delta$  the attractor is a singleton almost surely, thus the pathwise convergence is in both the pullback and forward directions.*

**Proof Step 1.** Similar to [13, Proposition 4.6], we prove that there exist a time  $r > 0$ , a constant  $\eta \in (0, 1)$ , and an integrable random variable  $\xi_r(x)$  such that

$$\|y_r(x, y_0)\|^p \leq \eta \|y_0\|^p + \xi_r(x). \tag{3.41}$$

First we fix  $r > 0$  and consider  $\mu, h$  as defined in Lemma 2.10 on  $[0, r]$ , i.e  $\mu_t$  is the solution of the deterministic system  $\dot{\mu} = A\mu + f(\mu)$  which starts at  $\mu_0 = y_0$  and  $h_t = y_t - \mu_t$ . Then using (3.12)

$$\|\mu_r\| \leq C_A \|y_0\| e^{-\lambda r} + C_A \frac{\|f(0)\|}{\lambda}.$$

On the other hand, due to (2.27) in Corollary 2.10 and (2.14),

$$\|h_r\| \leq \|h\|_{\infty, [0, r]} \leq \xi_0(x)(1 + \|y_0\|^\beta),$$

where  $\beta = \frac{1}{p}$  and  $\xi_0$  is a polynomial of  $\|x\|_{p\text{-var}, [0, r]}$  of the form

$$\xi_0(x) = D \|x\|_{p\text{-var}, [0, r]} \left( 1 + \|x\|_{p\text{-var}, [0, r]}^p \right),$$

where  $D$  is a constant.

Now for  $\epsilon > 0$  small enough, we apply the convex inequality and Young inequality to conclude that

$$\|y_r\|^{2p} \leq (\|h_r\| + \|\mu_r\|)^{2p} \leq (1 + \epsilon)^{2(2p-1)} \left[ (C_A e^{-\lambda r})^{2p} + \epsilon\beta \right] \|y_0\|^{2p} + \xi_r(x), \tag{3.42}$$

where

$$\xi_r(x) = \xi_r\left(\frac{1}{\epsilon}, x\right) \leq \frac{D}{\epsilon^{2p-1 + \frac{2p}{p-1}}} \left( 1 + \|x\|_{p\text{-var}, [0, r]}^{\frac{2p^2(p+1)}{p-1}} \right)$$

for some generic constant  $D$  (depends on  $r$ ). Thus  $\xi_r$  is integrable due to  $\frac{2p^2(p+1)}{p-1} \leq \frac{4p(p+1)}{p-1}$  and (3.40). By choosing  $r > 0$  large enough and  $\epsilon \in (0, 1)$  small enough such that

$$C_A e^{-\lambda r} < 1 \quad \text{and} \quad \eta := (1 + \epsilon)^{2(2p-1)} \left[ (C_A e^{-\lambda r})^{2p} + \epsilon\beta \right] < 1$$

we obtain (3.41).

**Step 2.** Next, for simplicity we only estimate  $y$  at the discrete times  $nr$  for  $n \in \mathbb{N}$ , the estimate for  $t \in [nr, (n+1)r]$  is similar to (3.21). From (3.41), it is easy to prove by induction that

$$\|y_{nr}(x, y_0)\|^{2p} \leq \eta^n \|y_0\|^{2p} + \sum_{i=0}^{n-1} \eta^i \xi_r(\theta_{(n-i)r}x), \quad \forall n \geq 1;$$

thus for  $n$  large enough

$$\|y_{nr}(\theta_{-nr}x, y_0)\|^{2p} \leq \eta^n \|y_0\|^{2p} + \sum_{i=0}^n \eta^i \xi_r(\theta_{-ir}x) \leq 1 + \sum_{i=0}^\infty \eta^i \xi_r(\theta_{-ir}x) =: R_r(x).$$

In this case we could choose  $\hat{b}(x)$  in (3.27) to be  $\hat{b}(x) = R_r(x)^{\frac{1}{2p}}$  so that there exists a pullback absorbing set  $\mathcal{B}(x) = B(0, \hat{b}(x))$  containing our random attractor  $\mathcal{A}(x)$ . Moreover, due to the integrability of  $\xi_r(x)$ ,  $R_r(x)$  is also integrable with  $\mathbb{E}R_r = 1 + \frac{\mathbb{E}\xi_r}{1-\eta}$ .

**Step 3.** Now back to the arguments in the proof of Theorem 3.9 and note that  $Dg$  is also globally Lipschitz with the same constant  $C_g$ . Using Lemma 2.7 (i) and rewriting (3.34) in Lemma 3.8 for  $x^*$  yields

$$e^{\lambda n} \|z_n\| \leq C_A \|z_0\| + M_1 C_g \sum_{k=0}^{n-1} \|x^*\|_{p\text{-var}, \Delta_k} e^{\lambda k} \left( 1 + \|y^1\|_{p\text{-var}, \Delta_k} \right) \|z\|_{p\text{-var}, \Delta_k} \tag{3.43}$$

where the  $p$ -variation norm of  $z$  can be estimated, due to Corollary 2.8(i), as

$$\|z\|_{p\text{-var}, \Delta_k} \leq \left(N'_{\Delta_k}(x^*)\right)^{\frac{p-1}{p}} 2^{N'_{\Delta_k}(x^*)} e^{2L} \|z_a\|, \tag{3.44}$$

with 
$$N'_{\Delta_k}(x^*) \leq 1 + [2(K + 1)C_g]^p \left\| \|x^*\| \right\|_{p\text{-var}, \Delta_k}^p \left(1 + \left\| \|y^1\| \right\|_{p\text{-var}, \Delta_k}\right)^p. \tag{3.45}$$

This together with (3.43) derives

$$e^{\lambda n} \|z_n\| \leq C_A \|z_0\| + \sum_{k=0}^{n-1} I_k e^{\lambda k} \|z_k\|$$

where 
$$I_k = M_1 C_g \left\| \|x^*\| \right\|_{p\text{-var}, \Delta_k} \left(1 + \left\| \|y^1\| \right\|_{p\text{-var}, \Delta_k}\right) \left(N'_{\Delta_k}(x^*)\right)^{\frac{p-1}{p}} 2^{N'_{\Delta_k}(x^*)} e^{2L}. \tag{3.46}$$

By applying the discrete Gronwall Lemma 3.12, we obtain

$$\|z_n\| \leq C_A \|z_0\| e^{-\lambda n} \prod_{k=0}^{n-1} (1 + I_k)$$

which yields 
$$\frac{1}{n} \log \|z_n\| \leq \frac{1}{n} \log \left(2C_A \hat{b}(x^*)\right) - \lambda + \frac{1}{n} \sum_{k=0}^{n-1} \log(1 + I_k). \tag{3.47}$$

On the other hand, due to (2.14), (2.19) and (3.27), it is easy to prove with a generic constant  $D$  that

$$\left\| \|y^1\| \right\|_{p\text{-var}, \Delta_k} \leq D \left(1 + \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]}^{2p-1}\right) \left(1 + \hat{b}(\theta_{(k-n)}x) + \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]}\right),$$

thus

$$\begin{aligned} & [2(K + 1)C_g] \left\| \|x^*\| \right\|_{p\text{-var}, \Delta_k} \left(1 + \left\| \|y^1\| \right\|_{p\text{-var}, \Delta_k}\right) \\ & \leq D \left\{ \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]} + \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]}^2 + \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]}^{2p} \right. \\ & \quad \left. + \left\| \|\theta_{(k-n)}x\| \right\|_{p\text{-var}, [0,1]}^{2p+1} \right\} \times \\ & \quad \times [1 + \hat{b}(\theta_{(k-n)}x)] =: \hat{F}(\theta_{(k-n)}x). \end{aligned}$$

All together,  $I_k$  is bounded from above by

$$I_k \leq DC_g \hat{F}(\theta_{(k-n)}x) \left[1 + \hat{F}^{p-1}(\theta_{(k-n)}x)\right] \exp \left\{ \log 2 \left[1 + \hat{F}^p(\theta_{(k-n)}x)\right] \right\} \tag{3.48}$$

where the right hand side of (3.48) is a function of  $\theta_{(k-n)}x$ . The ergodic Birkhorff theorem is then applied for (3.47), so that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \|z_n\| \leq -\lambda + E \log \left\{ 1 + DC_g \left[ \hat{F}(x) + \hat{F}^p(x) \right] e^{D\hat{F}^p(x) \log 2} \right\} \text{ a.s. } \tag{3.49}$$

Apply the inequalities

$$\begin{aligned} \log(1 + x + y) & \leq \log(1 + x) + \log(1 + y), \\ \log(1 + xy) & \leq \log(1 + x) + \log(1 + y), \\ \log(1 + xe^y) & \leq x + y, \quad x, y \geq 0, \end{aligned}$$

it follows that

$$\log \left\{ 1 + DC_g \left[ \hat{F}(x) + \hat{F}^p(x) \right] e^{D\hat{F}^p(x) \log 2} \right\} \leq D \left( 1 + \hat{F}(x) + \hat{F}^p(x) \right). \tag{3.50}$$

To estimate  $\hat{F}^p(x)$ , we apply Cauchy and Young inequalities to obtain, up to a generic constant  $D > 0$ ,

$$\begin{aligned} \hat{F}^p(x) &\leq D \left[ 1 + \|x\|_{p\text{-var},[0,1]}^2 + \|x\|_{p\text{-var},[0,1]}^4 + \|x\|_{p\text{-var},[0,1]}^{4p} + \|x\|_{p\text{-var},[0,1]}^{4p+2} + \hat{b}^2(x) \right]^p \\ &\leq D \left[ 1 + \|x\|_{p\text{-var},[0,1]}^{2p} + \|x\|_{p\text{-var},[0,1]}^{4p} + \|x\|_{p\text{-var},[0,1]}^{4p^2} + \|x\|_{p\text{-var},[0,1]}^{4p^2+2p} + \hat{b}^{2p}(x) \right]. \end{aligned}$$

Hence the right hand side in the last line of (3.50) is integrable due to (3.27) and the integrability of  $\|x\|_{p\text{-var},[-1,1]}^{\frac{4p(p+1)}{p-1}}$  in (3.40) and of  $\hat{b}(x)^{2p}$  in Step 2. On the other hand, the expression under the expectation of (3.49) tends to zero a.s. as  $C_g$  tends to zero. Due to the Lebesgue’s dominated convergence theorem, the expectation converges to zero as  $C_g$  tends to zero. As a result, there exists  $\delta$  small enough such that for  $C_g < \delta$  we have  $\|z_n\| = \|a_1 - a_2\| \rightarrow 0$  as  $n$  tends to infinity exponentially with the uniform convergence rate in (3.49). This proves  $a_1 \equiv a_2$  a.s. and  $\mathcal{A}$  is a singleton.

**Step 4.** Let  $y_t^1 = y(t, x, a(x))$ ,  $y_t^2 = y(t, x, y_0(x))$  be the solutions starting from  $a(x)$ ,  $y_0(x)$  respectively at  $t = 0$ . Since  $\mathcal{A}$  is invariant,  $y_t^1 = a(\theta_t x)$ . By repeating the arguments in Step 3 ( $x^*$  is replaced by  $x$ ), we conclude that  $\mathcal{A}(x) = \{a(x)\}$  is also a forward attractor.  $\square$

**Acknowledgements** The authors would like to thank the anonymous referee for valuable remarks. This work was supported by the Max Planck Institute for Mathematics in the Science (MIS-Leipzig), and the International Centre for Research and Postgraduate Training in Mathematics (ICRTM)- Institute of Mathematics, Vietnam Academy of Science and Technology under grant number ICRTM02\_2020.02. L.H. Duc would like to thank Vietnam Institute for Advanced Studies in Mathematics (VIASM) for financial support during three month research stay at the institute in 2019. P.T.Hong would like to thank the IMU Breakout Graduate Fellowship Program for the financial support.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

**Lemma 3.12** (Discrete Gronwall Lemma) *Let  $a$  be a non negative constant and  $u_n, \alpha_n, \beta_n$  be nonnegative sequences satisfying*

$$u_n \leq a + \sum_{k=0}^{n-1} \alpha_k u_k + \sum_{k=0}^{n-1} \beta_k, \quad \forall n \geq 1$$

then 
$$u_n \leq \max\{a, u_0\} \prod_{k=0}^{n-1} (1 + \alpha_k) + \sum_{k=0}^{n-1} \beta_k \prod_{j=k+1}^{n-1} (1 + \alpha_j), \quad \forall n \geq 1. \tag{3.51}$$



**Proof** Put

$$S_n := a + \sum_{k=0}^{n-1} \alpha_k u_k + \sum_{k=0}^{n-1} \beta_k, \quad T_n := \max\{a, u_0\} \prod_{k=0}^{n-1} (1 + \alpha_k) + \sum_{k=0}^{n-1} \beta_k \prod_{j=k+1}^{n-1} (1 + \alpha_j).$$

We will prove by induction that  $S_n \leq T_n$  for all  $n \geq 1$ . Namely, the statement holds for  $n = 1$  since  $S_1 = a + \alpha_0 u_0 + \beta_0 \leq \max\{a, u_0\}(1 + \alpha_0) + \beta_0 = T_1$ .

We assume that  $S_n \leq T_n$  for  $n \geq 1$ , then due to the fact that  $u_n \leq S_n$  we obtain

$$\begin{aligned} S_{n+1} &= a + \sum_{k=0}^{n-1} \alpha_k u_k + \sum_{k=0}^{n-1} \beta_k + \alpha_n u_n + \beta_n = S_n + \alpha_n u_n + \beta_n \\ &\leq S_n + \alpha_n S_n + \beta_n \leq T_n (1 + \alpha_n) + \beta_n \\ &\leq \left[ \max\{a, u_0\} \prod_{k=0}^{n-1} (1 + \alpha_k) + \sum_{k=0}^{n-1} \beta_k \prod_{j=k+1}^{n-1} (1 + \alpha_j) \right] (1 + \alpha_n) + \beta_n \\ &\leq \max\{a, u_0\} \prod_{k=0}^n (1 + \alpha_k) + \sum_{k=0}^n \beta_k \prod_{j=k+1}^{n-1} (1 + \alpha_j) = T_{n+1}. \end{aligned}$$

Since  $u_n \leq S_n$ , (3.51) holds. □

**References**

1. Adrianova, L.Y.: Introduction to linear systems of differential equations. In: Translations of Mathematical Monographs, vol. 46, Americal Mathematical Society (1995)
2. Amann, H.: Ordinary Differential Equations: An Introduction to Nonlinear Analysis. Walter de Gruyter, Berlin (1990)
3. Arnold, L.: Random Dynamical Systems. Springer, Berlin (1998)
4. Bailleul, I., Riedel, S., Scheutzow, M.: Random dynamical systems, rough paths and rough flows. *J. Differ. Equ.* **262**, 5792–5823 (2017)
5. Cass, T., Litterer, C., Lyons, T.: Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **14**(4), 3026–3050 (2013)
6. Cong, N.D., Duc, L.H., Hong, P.T.: Young differential equations revisited. *J. Dyn. Differ. Equ.* **30**(4), 1921–1943 (2018)
7. Crauel, H., Kloeden, P.: Nonautonomous and random attractors. *Jahresber Dtsch. Math. Ver.* **117**, 173–206 (2015)
8. Duc, L.H., Garrido-Atienza, M.J., Neuenkirch, A., Schmalfuß, B.: Exponential stability of stochastic evolution equations driven by small fractional Brownian motion with Hurst parameter in  $(\frac{1}{2}, 1)$ . *J. Differ. Equ.* **264**, 1119–1145 (2018)
9. Duc, L.H., Hong, P.T., Cong, N.D.: Asymptotic stability for stochastic dissipative systems with a Hölder noise. *SIAM J. Control. Optim.* **57**(4), 3046–3071 (2019)
10. Friz, P., Victoir, N.: Multidimensional stochastic processes as rough paths: theory and applications. In: Cambridge Studies in Advanced Mathematics, vo. 120, Cambridge University Press, Cambridge (2010)
11. Garrido-Atienza, M., Maslowski, B., Schmalfuß, B.: Random attractors for stochastic equations driven by a fractional Brownian motion. *Int. J. Bifurc. Chaos* **20**(9), 2761–2782 (2010)
12. Garrido-Atienza, M., Schmalfuss, B.: Ergodicity of the infinite dimensional fractional Brownian motion. *J. Dyn. Differ. Equ.* **23**, 671–681 (2011). <https://doi.org/10.1007/s10884-011-9222-5>
13. Hairer, M., Ohashi, A.: Ergodic theory for sdes with extrinsic memory. *Ann. Probab.* **35**, 1950–1977 (2007)
14. Khasminskii, R.: Stochastic Stability of Differential Equations, vol. 66. Springer, Berlin (2011)
15. Imkeller, P., Schmalfuss, B.: The conjugacy of stochastic and random differential equations and the existence of global attractors. *J. Dyn. Differ. Equ.* **13**(2), 215–249 (2001)
16. Keller, H., Schmalfuss, B.: Attractors for stochastic differential equations with nontrivial noise. *Bul. Acad. Stiinte Repub. Mold. Mat.* **26**(1), 43–54 (1998)

17. Lyons, T.: Differential equations driven by rough signals, I: an extension of an inequality of LC Young. *Math. Res. Lett.* **1**(4), 451–464 (1994)
18. Lyons, T.: Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14**(2), 215–310 (1998)
19. Lejay, A.: Controlled differential equations as Young integrals: a simple approach. *J. Differ. Equ.* **249**, 1777–1798 (2010)
20. Mandelbrot, B., van Ness, J.: Fractional Brownian motion, fractional noises and applications. *SIAM Rev.* **4**(10), 422–437 (1968)
21. Nualart, D., Răşcanu, A.: Differential equations driven by fractional Brownian motion. *Collect. Math.* **53**(1), 55–81 (2002)
22. Riedel, S., Scheutzow, M.: Rough differential equations with unbounded drift terms. *J. Differ. Equ.* **262**, 283–312 (2017)
23. Sussmann, H.J.: On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probab.* **6**(1), 19–41 (1978)
24. Young, L.C.: An integration of Hölder type, connected with Stieltjes integration. *Acta Math.* **67**, 251–282 (1936)
25. Zähle, M.: Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Rel. Fields* **111**(3), 333–374 (1998)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.