

# *Fractional Stochastic Differential Equations Satisfying Fluctuation-Dissipation Theorem*

**Lei Li, Jian-Guo Liu & Jianfeng Lu**

**Journal of Statistical Physics**

1

ISSN 0022-4715

Volume 169

Number 2

J Stat Phys (2017) 169:316-339

DOI 10.1007/s10955-017-1866-z

Volume 169 • Number 2 • October 2017

Journal of  
Statistical  
Physics

10955 • ISSN 0022-4715  
169(2) 223-440 (2017)

 Springer

 Springer

**Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Fractional Stochastic Differential Equations Satisfying Fluctuation-Dissipation Theorem

Lei Li<sup>1</sup>  · Jian-Guo Liu<sup>2</sup>  · Jianfeng Lu<sup>3</sup> 

Received: 16 January 2017 / Accepted: 23 August 2017 / Published online: 2 September 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** We propose in this work a fractional stochastic differential equation (FSDE) model consistent with the over-damped limit of the generalized Langevin equation model. As a result of the ‘fluctuation-dissipation theorem’, the differential equations driven by fractional Brownian noise to model memory effects should be paired with Caputo derivatives, and this FSDE model should be understood in an integral form. We establish the existence of strong solutions for such equations and discuss the ergodicity and convergence to Gibbs measure. In the linear forcing regime, we show rigorously the algebraic convergence to Gibbs measure when the ‘fluctuation-dissipation theorem’ is satisfied, and this verifies that satisfying ‘fluctuation-dissipation theorem’ indeed leads to the correct physical behavior. We further discuss possible approaches to analyze the ergodicity and convergence to Gibbs measure in the nonlinear forcing regime, while leave the rigorous analysis for future works. The FSDE model proposed is suitable for systems in contact with heat bath with power-law kernel and subdiffusion behaviors.

**Keywords** Fractional SDE · Fluctuation-dissipation-theorem · Caputo derivative · Fractional Brownian motion · Generalized Langevin equation · Subdiffusion

---

✉ Lei Li  
leili@math.duke.edu

Jian-Guo Liu  
jliu@phy.duke.edu

Jianfeng Lu  
jianfeng@math.duke.edu

<sup>1</sup> Department of Mathematics, Duke University, Durham, NC 27708, USA

<sup>2</sup> Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA

<sup>3</sup> Department of Mathematics, Department of Physics, and Department of Chemistry, Duke University, Box 90320, Durham, NC 27708, USA

### 1 Introduction

For a particle in contact with a heat bath (such as a heavy particle surrounded by light particles), the following stochastic equation is often used to describe the evolution of the velocity of the particle

$$m\dot{v} = -\gamma v + \eta,$$

where dot denotes derivative on time,  $-\gamma v$  counts for friction and  $\eta$  is a Gaussian white noise which could be understood as the distributional derivative of the Brownian motion (or Wiener process) up to a constant factor. This equation should be understood in the SDE form

$$m dv = -\gamma v dt + \sqrt{2D_x} dW,$$

where  $W$  is a standard Brownian motion and  $D_x$  is some constant to be determined. Adding the equation for position and considering external force, one has the Langevin equation:

$$\dot{x} = v, \quad m\dot{v} = -\nabla V(x) - \gamma v + \eta. \tag{1.1}$$

Since the friction coefficient  $\gamma$  and random force  $\eta$  both stem from interactions between the particle and the environment, they should be related. The ‘fluctuation-dissipation theorem’<sup>1</sup> [2,33] provides a precise connection between them, such that the covariance satisfies

$$\mathbb{E}(\eta(t_1)\eta(t_2)) = 2kT\gamma\delta(t_1 - t_2), \tag{1.2}$$

where  $k$  is the Boltzmann constant and  $T$  is the absolute temperature, leading to  $D_x = kT\gamma$ .  $\mathbb{E}$  is the ‘ensemble average’ in physical language and it is ‘expectation’ over some underlying probability space in mathematical language. Relation (1.2) was formulated by Nyquist in [33] and then justified by Callen and Welton in [2]. The physical meaning of this relation is that the fluctuating forces must restore the energy dissipated by the friction so that the balance is achieved and the temperature of the heavy particle can reach the correct value. To see this in another view point, one may derive, either using Ito’s formula or using Green-Kubo formula, that  $D_x$  is actually the diffusion constant for position  $x$ , and  $D_x = kT\gamma$  is called the Einstein-Smoluchowski relation [25]. This relation also says that the fluctuation and dissipation must be related.

In the ‘overdamped’ regime where the inertia can be neglected ( $m \ll 1$ ), the Langevin equation is reduced to the following well-known SDE [10]:

$$\gamma dx = -\nabla V(x) dt + \sqrt{2D_x} dW. \tag{1.3}$$

In [20,30], the generalized Langevin equation (GLE) was proposed to model particle motion in contact with a heat bath when the random force is no longer memoryless:

$$\dot{x} = v, \quad m\dot{v} = -\nabla V - \int_{t_0}^t \gamma(t-s)v(s) ds + R(t), \tag{1.4}$$

where  $R(t)$  is some random force. Now the friction is the convolution between a kernel function  $\gamma$  and the velocity  $v(s)$  so that there is memory in dissipation in this model. For the particle to achieve equilibrium at the prescribed temperature, the fluctuating force  $R(t)$  and the friction kernel  $\gamma$  must be related. Without the external force (i.e.  $\nabla V = 0$ ), Kubo assumed that  $\mathbb{E}(v(t_0)R(t)) = 0, t > t_0$  and that  $v$  is a stationary process. He derived formally

<sup>1</sup> Note that we are putting quotes for the physical theorems as they are critical claims from physics compared with mathematical theorems that are rigorously justified.

(though he used the existence of the one-sided Fourier transform of  $\gamma$ , the formal derivation still holds if  $\gamma \notin L^1[0, \infty)$  as we can understand the transform in the distribution sense or replace the one-sided Fourier transform with Laplace transform) that

$$\mathbb{E}(R(t_0)R(t_0 + t)) = m\mathbb{E}(v(t_0)^2)\gamma(|t|) = kT\gamma(|t|). \tag{1.5}$$

There are other formal derivations as well (e.g. [8]). These derivations are not fully convincing though on the mathematical rigorous level. In [20], Kubo assumed the relation  $\mathbb{E}(v(t_0)R(t)) = 0, t > t_0$  arguing using causality. The issue is though  $R(t)$  does not affect  $v(t_0), v(t_0)$  can affect  $R(t)$ . In [8], Felderhoff obtained this relation from ‘Nyquist’s theorem’, while no justification is given to the latter.

For a more convincing and rigorous derivation of the GLE (1.4) and relation (1.5), one could start from a system of interacting particles as the Kac–Zwanzig model (see [9,11,19,42]). In this model, the surrounding particles in the heat bath have harmonic interactions with the particle under consideration, which is a good approximation if the configuration is near equilibrium. The whole system evolves under the total Hamiltonian. If the initial data satisfy the Gibbs measure, then after integrating out the variables for the surrounding particles, one obtains the GLE where the relation (1.5) is satisfied. From the Kac–Zwanzig model, we may find that in GLE the random force  $R(t)$  is not necessarily independent of  $x(0)$ .

Relation (1.5) is called the ‘fluctuation-dissipation theorem’ for GLE. This relation simply says the random force must balance the friction so that the system has a nontrivial equilibrium corresponds to the prescribed temperature. Note that if the kernel  $\gamma(t)$  tends to  $\gamma\delta(t)$ , the relation (1.2) can be recovered. The coefficient ‘2’ comes from the fact that

$$\int_{-\infty}^{\infty} \mathbb{E}(R(t_0)R(t_0 + t)) dt = 2kT \int_0^{\infty} \gamma(t) dt.$$

There are few rigorous mathematical justifications of the ‘fluctuation-dissipation theorem’ in the context of generalized Langevin equations. In [35], the author tried to rephrase the ‘fluctuation-dissipation theorems’ and the related linear response theory in mathematical language. Hairer and Majda in [14] developed a framework to justify the use the linear response theory through the ‘fluctuation-dissipation theorem’ for studying climate models.

In this work, through a scaling argument, we find it reasonable to consider the over-damped limit of GLE driven by fractional Brownian noise, and obtain the following fractional SDE (FSDE) model (Eq. 2.11)

$$D_c^\alpha x = -V'(x) + C_H \dot{B}_H,$$

where the fractional derivative is in Caputo sense while  $\dot{B}_H$  is the fractional Brownian noise (the distributional derivative of fractional Brownian motion). This differential form can be rewritten as an integral form (Eq. 2.14):

$$x(t) = x(0) - \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} V'(x(s)) ds + \frac{C_H}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} dB_H,$$

which is viewed as the rigorous definition of our FSDE model. After proving that the stochastic integral is a continuous process in Sect. 4, the existence and uniqueness of strong solutions become straightforward.

Let us remark that using fractional Brownian motion as a model for long range correlations is quite common: for example, waves in random media [26], subdiffusion process in complex system [4, 18, 19, 22, 29]. It has been observed in [18, 19, 29] that the systems of protein molecules have power-law memory kernel and subdiffusion behavior. Remarkably, Kou and Xie [18, 19] showed that incorporating fractional Brownian noise into the generalized

Langevin equation yields a model with a power-law kernel for subdiffusion and the results had excellent agreement with a single-molecule experiments from biological science.

If  $\alpha = \alpha^* := 2 - 2H$ , the ‘fluctuation-dissipation theorem’ is satisfied. When the force is linear, we show rigorously that the process has ergodicity and converges algebraically to the Gibbs measure (see Theorem 2). When the force is nonlinear, studying the ergodicity and asymptotic behavior is challenging. We think it is promising to solve this problem by rewriting the FSDE model into Markovian processes. So, we propose two possible approaches. The first approach is to rewrite our FSDE model as an infinite-dimensional Ornstein–Uhlenbeck (OU) process with mixing. We hope this infinite-dimensional OU with mixing can be a possible framework for proving the convergence to equilibrium satisfying Gibbs measure. Another approach is to take the limit in a heat bath model. In summary, satisfying the ‘fluctuation-dissipation theorem’ leads to the correct physical behavior: there is balance between the dissipation and fluctuation effects from the random forcing such that the Gibbs measure is the final equilibrium distribution. This means that in the correct physical FSDE models fractional Brownian noise must be paired with Caputo derivatives.

While FSDEs have been discussed in some previous works already, our FSDE model (2.14) motivated by the ‘fluctuation-dissipation theorem’ seems to be new. The authors of [13, 15, 32] discussed FSDEs driven by fractional Brownian motions but used the usual first order derivative, which means that the convolution kernel for friction is a Dirac delta and there is no memory in the dissipating term, while the fluctuation term is given by fractional Brownian motions that have memory so that there is no balance. In [39], the Caputo derivative is used but they used the usual white noise to drive the process. According to the above formal derivation, when modeling a particle in contact with a heat bath with memory effects, the natural noise associated with the Caputo derivative should be the fractional noise. This means we will probably require  $\alpha = \alpha^*$  for the correct model from physical concerns. We admit however that it is possible that the models with  $\alpha \neq \alpha^*$  may be used to describe some other situations instead of the physical case we consider here.

The rest of the paper is organized as follows. In Sect. 2, we give a brief introduction to fractional Brownian motions and propose our FSDE model (2.14). In Sect. 3, we study the stochastic integral in our FSDE model (2.14) in detail and prove that it is continuous. Using the continuity of the stochastic integral, we obtain in Sect. 4 the existence and uniqueness of strong solutions for FSDE (2.14) on the interval  $[0, \infty)$  provided  $V'(\cdot)$  is Lipschitz continuous. In Sect. 5, we focus on the asymptotic behavior of the strong solutions of (2.14). In particular, in the linear regimes, (i.e.  $V'(\cdot)$  is a linear function), we compute the solutions exactly and show that the solution converges in distribution to a stationary process satisfying Gibbs measure. In the nonlinear regime, we provide two possible frameworks for studying the asymptotic behaviors when the ‘fluctuation-dissipation theorem’ is satisfied. We argue formally that the FSDE can be reduced from some Markovian processes in infinite dimensions. The rigorous study of the nonlinear regimes is left for future works.

## 2 The FSDE Model

In this section, we propose the fractional SDE model from the GLE with fractional Brownian noise. By a scaling argument in GLE, we argue that in the regimes where the environment is viscous or the mass is small, we can consider the over-damped limit of GLE driven by fractional Brownian noise (i.e., the (distributional) derivative of fractional Brownian motion) and obtain the fractional SDE model, in which the Caputo derivative is associated with

the fractional Brownian noise. This model is new. It recovers the subdiffusion discussed in [18,22] and satisfies the ‘fluctuation-dissipation theorem’.

### 2.1 Fractional Brownian Noise in Complex Systems

The studies in [18,19,22,29] indicate that fractional Brownian noise commonly arises in complex physical systems. We now give a brief introduction to fractional Brownian motion and present the GLE with fractional Brownian noise.

The fractional Brownian motion  $B_H$  (see [24,31] for more detailed discussions) with Hurst parameter  $H \in (0, 1)$  is a Gaussian process (i.e., the joint distribution for  $(B_H(t_1), \dots, B_H(t_d))$  is a  $d$ -dimensional normal distribution for any  $(t_1, \dots, t_d) \in \mathbb{R}_+^d$ ) defined on some probability space  $(\Omega, \mathcal{F}, P)$  with mean zero and covariance

$$\mathbb{E} \left( B_t^H B_s^H \right) = R_H(s, t) = \frac{1}{2} \left( s^{2H} + t^{2H} - |t - s|^{2H} \right), \tag{2.1}$$

where  $\mathbb{E}$  means the expectation over the underlying probability space. By definition,  $B_H$  has stationary increments which are normal distributions with  $\mathbb{E}((B_H(t) - B_H(s))^2) = (t - s)^{2H}$ . By the Kolmogorov continuity theorem,  $B_H$  is Hölder continuous with order  $H - \epsilon$  for any  $\epsilon \in (0, H)$ .  $B_H$  has finite  $1/H$ -variation. Besides, it is self similar:  $B_H(t) \stackrel{d}{=} a^{-H} B_H(at)$  where ‘ $\stackrel{d}{=}$ ’ means they have the same distribution. It is non-Markovian except for  $H = 1/2$  when it is reduced to the Brownian motion (i.e., Wiener process).

The existence of fractional Brownian motion can be proved by some explicit representations. In [24], the following representation is given

$$\begin{aligned} B_H(t) &= C_1(H) \left( \int_0^t (t - s)^{H-\frac{1}{2}} dW(s) + \int_{-\infty}^0 \left( (t - s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right) dW(s) \right) \\ &= C_1(H) \int_{-\infty}^0 (-r)^{H-\frac{1}{2}} (dW(r + t) - dW(r)), \end{aligned} \tag{2.2}$$

where  $W$  is a normal Brownian motion and  $C_1(H)$  is a constant to make (2.1) valid. This is also used in [13]. In [3,32], one uses

$$B_H(t) = C_2(H) \int_0^t (t - s)^{H-\frac{1}{2}} F \left( H - \frac{1}{2}, \frac{1}{2} - H, H + \frac{1}{2}, 1 - \frac{t}{s} \right) dW(s), \tag{2.3}$$

where  $F$  is the Gauss hypergeometric function. Another representation in [37] using fractional integrals might be useful sometimes, which we choose to omit here.

One can show that  $(B_H(t + h) - B_H(t))/h$  converges in distribution (i.e. under the topology of the dual of  $C_c^\infty(0, \infty)$ ) to  $\dot{B}_H(t)$  where the dot represents distributional time derivative. We check that

$$\begin{aligned} &\lim_{h \rightarrow 0^+, h_1 \rightarrow 0} \mathbb{E} \left( \frac{B_H(h) B_H(t + h_1) - B_H(t)}{h h_1} \right) \\ &= \lim_{h \rightarrow 0^+, h_1 \rightarrow 0} \frac{1}{2hh_1} \left( (t + h_1)^{2H} - (t + h_1 - h)^{2H} - t^{2H} + (t - h)^{2H} \right) \\ &= H(2H - 1)t^{2H-2}. \end{aligned} \tag{2.4}$$

If we pick the initial time in (1.4) as  $t_0 = 0$  and consider the random noises corresponding to fractional Brownian motion as discussed:

$$R_H(t) = \frac{\sqrt{kT\gamma_0}}{\sqrt{H(2H - 1)\Gamma(2H - 1)}} \dot{B}_H(t), \tag{2.5}$$

where  $\gamma_0$  is a constant representing the typical scale of friction, we then have the GLE model

$$m\dot{v} = -\nabla V(x) - \frac{\gamma_0}{\Gamma(2H - 1)} \int_0^t (t - s)^{2H-2} v(s) ds + R_H(t) \tag{2.6}$$

following the ‘fluctuation-dissipation theorem’.

We will assume throughout the paper that

$$H \in \left(\frac{1}{2}, 1\right), \tag{2.7}$$

as they are the physically most realistic regimes [19] and consequently  $2 - 2H \in (0, 1)$ .

### 2.2 Over-Damped Limit and the FSDE Model

Assume that we consider the fractional diffusion regime with time scale  $T_t$ , the length scale  $L = \sqrt{kT/\gamma_0} T_t^{1-H}$ , and velocity scale  $L/T_t = \sqrt{kT/\gamma_0} T_t^{-H}$ . We then scale the energy with  $kT$ , fractional Brownian motion with  $T_t^H$  and scale the noise with  $\sqrt{kT\gamma_0} T_t^{H-1}$ . The dimensionless GLE reads

$$\frac{mT_t^{2H}}{\gamma_0} \dot{v} = -\nabla V - \frac{1}{\Gamma(2H - 1)} \int_0^t (t - s)^{2H-2} v(s) ds + R_H(t)$$

In the regimes where  $mT_t^{2H}/\gamma_0$  is small (viscous environment, particle is small etc), the  $m\dot{v}$  term in (2.6) can be neglected, and we have the following dimensionless over-damped equation with fractional noise:

$$\frac{1}{\Gamma(2H - 1)} \int_0^t (t - s)^{2H-2} v(s) ds = -\nabla V(x) + R_H(t). \tag{2.8}$$

In the following discussion, we will always assume the equations are dimensionless while the variables  $k$  and  $T$  might be used to denote other quantities.

Recall that the Caputo derivative [12, 17] starting from  $t = 0$  for a  $C^1$  function is given by

$$D_c^\alpha w = \frac{1}{\Gamma(1 - \alpha)} \int_0^t \frac{\dot{w}(s)}{(t - s)^\alpha} ds. \tag{2.9}$$

Note that  $v(s) = \dot{x}(s)$ . The left hand side of Eq. (2.8) formally becomes the Caputo derivative of  $x$  with  $\alpha = 2 - 2H$  and the equation becomes a fractional SDE:

$$D_c^{2-2H} x = -\nabla V(x) + R_H(t). \tag{2.10}$$

This means that the power-law memory kernel yields the Caputo derivative of the trajectory naturally. This over-damped fractional SDE model is simpler compared with the GLE model (2.6) and we expect it to contain the essential physics (the subdiffusion and ‘fluctuation-dissipation theorem’) as we will study.

From here on, we will only consider 1D case ( $x \in \mathbb{R}$ ) for convenience while the general dimension is similar. The above discussion then motivates us to consider the fractional stochastic differential equation (FSDE) where we relax the constraint between  $H$  and  $\alpha$ :

$$D_c^\alpha x = -V'(x) + C_H \dot{B}_H, \tag{2.11}$$

where

$$C_H = \frac{1}{\sqrt{H(2H - 1)\Gamma(1 - \alpha)}} \tag{2.12}$$



for  $\alpha \in (1 - H, 1)$ . The index obtained from the ‘fluctuation-dissipation theorem’ is denoted as  $\alpha^* = 2 - 2H$ . We will also denote the (one-sided) kernel associated with the Caputo derivative as

$$\gamma(t) = \frac{\theta(t)}{\Gamma(1 - \alpha)} t^{-\alpha}, \tag{2.13}$$

where  $\theta(t)$  is the standard Heaviside step function.

In [21], a definition of the Caputo derivative based on a convolution group was proposed, which agrees with (2.9) when the function is absolutely continuous on  $(0, t)$ . The observation of the underlying convolution group structure allows us to de-convolve and change the Caputo derivative to integral form as

$$\begin{aligned} x(t) &= x(0) + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} D_c^\alpha x(s) ds \\ &= x(0) - \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} V'(x(s)) ds + \frac{C_H}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} dB_H, \end{aligned} \tag{2.14}$$

where we formally used  $R_H ds = C_H \dot{B}_H ds = C_H dB_H$ . This integral will then be understood as the rigorous definition of the FSDE (2.11). The last term in (2.14) is an integral with respect to fractional Brownian motion, which we will make the meaning precise later. We will study FSDE (2.14) and try to understand the role of the ‘fluctuation-dissipation theorem’. For convenience, we denote

$$G(t) = \frac{C_H}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} dB_H(s) = \int_0^\infty f_t(s) dB_H(s), \tag{2.15}$$

where  $f_t(s) = \frac{C_H}{\Gamma(\alpha)} ((t - s)^+)^{\alpha-1}$  and  $\alpha \in (1 - H, 1)$ . We shall study the process  $G$  in Sect. 3.

### 3 The Process $G$ as a Stochastic Integral

To make the meaning of the FSDE precise, we must understand the process  $G$ . In this section, we first review the stochastic integrals with respect to fractional Brownian motions and then study some properties of  $G$ .

#### 3.1 Stochastic Integrals Driven by Fractional Brownian Motions

The stochastic integrals with respect to fractional Brownian motions have been thoroughly discussed in literature [3, 7, 28, 41]. In [28, 41], the stochastic integrals are defined pathwise using the Riemann–Stieltjes integrals by making use of certain properties of the paths. In [3, 7], the so-called Malliavin calculus is used to define the stochastic integrals (Wick–Ito–Skorohod integrals, or the ‘divergence’) and the Ito’s formula is established, which connects both definitions. For a review, one can refer to [1, 31]. In the case that the integrand is deterministic, those two definitions agree. By (2.14), we only need the integrals of deterministic processes with respect to fractional Brownian motion. We shall give a brief introduction to the theory for deterministic processes and the readers can turn to the references listed here for general processes.

Let us fix  $T > 0$  and define the stochastic integrals on the interval  $[0, T]$ . The definition of integration of deterministic processes on  $[0, T]$  starts with the step functions. Let  $\mathcal{E}$  be the set of all step functions on  $[0, T]$ , i.e.  $\varphi \in \mathcal{E}$  is given by

$$\varphi = \sum_{j=1}^m a_j 1_{(t_{j-1}, t_j]}(t), \tag{3.1}$$

where  $1_E(t)$  is the characteristic function of set  $E$ . The integral  $B^H(\varphi)$  is defined by

$$B^H(\varphi) = \int_0^T \varphi dB_H(t) = \sum_{j=1}^m a_j (B_H(t_j) - B_H(t_{j-1})). \tag{3.2}$$

Consider the inner product

$$\langle \varphi_1, \varphi_2 \rangle_{\mathcal{H}} = \mathbb{E}(B^H(\varphi_1)B^H(\varphi_2)). \tag{3.3}$$

It is easily verified that  $\forall \varphi_1, \varphi_2 \in \mathcal{E}$ ,

$$\begin{aligned} \langle \varphi_1, \varphi_2 \rangle_{\mathcal{H}} &= H(2H - 1) \int_0^T \int_0^T |r - u|^{2H-2} \varphi_1(r)\varphi_2(u) dudr \\ &= \frac{\pi\kappa(2\kappa + 1)}{\Gamma(1 - 2\kappa) \sin(\pi\kappa)} \int_0^T s^{-2\kappa} (I^\kappa u^\kappa f)(s)(I^\kappa u^\kappa g)(s) ds, \end{aligned} \tag{3.4}$$

where  $\kappa = H - \frac{1}{2}$  and  $I^\kappa$  is the right Riemann–Liouville fractional calculus, given by [37]:

$$(I^\kappa f)(s) = \begin{cases} \frac{1}{\Gamma(\kappa)} \int_s^T f(u)(u - s)^{\kappa-1} du, & \kappa > 0, \\ -\frac{1}{\Gamma(1 - \kappa)} \frac{d}{ds} \int_s^T f(u)(u - s)^{-\kappa} du. & \kappa < 0. \end{cases}$$

This then motivates the definition of

$$\mathcal{H}_0 = \left\{ \varphi \in L^1_{loc}[0, T] : \int_0^T \int_0^T |r - u|^{2H-2} |\varphi(r)| |\varphi(u)| dr du < \infty \right\} \tag{3.5}$$

and

$$\Lambda = \left\{ f \in L^1_{loc}[0, T] : \int_0^T s^{-2\kappa} (I^\kappa u^\kappa f)^2(s) ds < \infty \right\}. \tag{3.6}$$

Clearly,  $\mathcal{H}_0 \subset \Lambda$ . The integral  $B^H(\varphi)$  can then be defined for  $\varphi \in \Lambda$  by approximating them with step functions. In [36,37], it is shown that both inner product spaces  $\mathcal{H}_0$  and  $\Lambda$  are not complete and therefore not Hilbert spaces. However, the space  $B^H(\mathcal{E})$  clearly has a closure in  $L^2(\Omega, P)$ . This means some elements in the closure corresponds to distributions that are not in  $L^1_{loc}[0, T]$ . Let  $\mathcal{H}$  be the space of the closure of  $\mathcal{E}$  under the inner product (3.3) and thus  $\mathcal{H}$  contains some distributions.  $\forall \varphi_1, \varphi_2 \in \mathcal{H}_0 \subset \mathcal{H}$ ,

$$\langle \varphi_1, \varphi_2 \rangle_{\mathcal{H}} = \mathbb{E}(B^H(\varphi_1)B^H(\varphi_2)) = H(2H - 1) \int_0^T \int_0^T |r - u|^{2H-2} \varphi_1(r)\varphi_2(u) dudr. \tag{3.7}$$

The following lemma provides a convenient way to check that some deterministic processes can be integrated by fractional Brownian motion [27,31]:

**Lemma 1** *If  $H > 1/2$ , then there exists  $b_H > 0$  such that for any  $\varphi \in L^{1/H}[0, T]$  such that*

$$\|\varphi\|_{\mathcal{H}_0} \leq b_H \|\varphi\|_{L^{1/H}[0, T]}. \tag{3.8}$$

where

$$\|\varphi\|_{\mathcal{H}_0}^2 = \int_0^T \int_0^T |r - u|^{2H-2} |\varphi(r)| |\varphi(u)| dr du.$$

### 3.2 Some Basic Properties of $G$

We can easily verify that  $f_t \in L^{1/H}[0, T]$  whenever  $t \leq T$ , and hence the integral on  $[0, T]$  is well defined. Further, for any  $T_1 > t, T_2 > t$ , the integrals of  $f_t$  over  $[0, T_1]$  and  $[0, T_2]$  agree on  $[0, \min(T_1, T_2)]$ . In this sense, the integral  $\int_0^\infty f_t(s) dB_H(s)$  can then be understood as in  $[0, T]$  for any  $T > t$ .

Roughly speaking, since  $B_H$  is  $H - \epsilon$  Hölder continuous for any  $\epsilon \in (0, H)$ ,  $G(t)$  should be like  $\alpha + H - 1 - \epsilon$  Hölder continuous for any  $\epsilon \in (0, \alpha + H - 1)$  by the regularity of  $B_H$ . We shall make this precise in this subsection.

**Lemma 2**  $G(t)$  is a Gaussian process with mean zero and covariance given by

$$\begin{aligned} \phi(t_1, t_2) &= \mathbb{E}(G(t_1)G(t_2)) = \frac{B(2H - 1, \alpha)}{B(\alpha, 1 - \alpha)\Gamma(\alpha)} \\ &\times \int_0^{\min(t_1, t_2)} dr \left( (t_1 - r)^{\alpha-1} (t_2 - r)^{2H-2+\alpha} + (t_2 - r)^{\alpha-1} (t_1 - r)^{2H-2+\alpha} \right). \end{aligned} \tag{3.9}$$

In particular, if  $\alpha = \alpha^*$ ,  $G(t) \stackrel{d}{=} \beta_H B_{1-H}$  where  $\stackrel{d}{=}$  means they have the same distribution, and

$$\beta_H = \frac{\sqrt{2}}{\sqrt{\Gamma(3 - 2H)}}. \tag{3.10}$$

In other words,  $G(t)$  is a fractional Brownian motion with Hurst parameter  $1 - H$  up to a constant  $\beta_H$  if  $\alpha = \alpha^*$ .

*Proof* Clearly,  $G(t)$  is a Gaussian process with mean zero because any linear operation of Gaussian process is again Gaussian.

Without loss of generality, we can assume  $t_2 \geq t_1 \geq 0$ . The covariance can be computed using the isometry (3.7)

$$\begin{aligned} \mathbb{E}(G(t_1)G(t_2)) &= \langle f_{t_1}, f_{t_2} \rangle_{\mathcal{H}} \\ &= \frac{1}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \int_0^{t_1} \int_0^{t_2} |r - u|^{2H-2} (t_1 - r)^{\alpha-1} (t_2 - u)^{\alpha-1} du dr. \end{aligned}$$

We break the integral into two parts  $I_1 + I_2$ , where

$$I_1 = \frac{1}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \iint_{u \geq r} \dots du dr, \quad I_2 = \frac{1}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \iint_{r \geq u} \dots du dr.$$

By explicit computation,

$$\begin{aligned} I_1 &= \frac{1}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \int_0^{t_1} dr (t_1 - r)^{\alpha-1} \int_r^{t_2} du (u - r)^{2H-2} (t_2 - u)^{\alpha-1} \\ &= \frac{B(2H - 1, \alpha)}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \int_0^{t_1} (t_1 - r)^{\alpha-1} (t_2 - r)^{2H-2+\alpha} dr. \end{aligned}$$

This can further be written in terms of the so-called hypergeometric functions but we choose not to do it. Similarly,

$$\begin{aligned}
 I_2 &= \frac{1}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \int_0^{t_1} du (t_2 - u)^{\alpha-1} \int_u^{t_1} dr (r - u)^{2H-2} (t_1 - r)^{\alpha-1} \\
 &= \frac{B(2H - 1, \alpha)}{\Gamma(\alpha)B(\alpha, 1 - \alpha)} \int_0^{t_1} (t_2 - u)^{\alpha-1} (t_1 - u)^{2H-2+\alpha} du.
 \end{aligned}$$

If  $\alpha = \alpha^* = 2 - 2H$ , the integrals  $I_1$  and  $I_2$  can be evaluated exactly:

$$I_1 + I_2 = \frac{1}{\Gamma(3 - 2H)} \left( t_1^{2-2H} + t_2^{2-2H} - (t_2 - t_1)^{2-2H} \right),$$

which shows the last claim. □

The above computation shows trivially that

**Corollary 1** *In the case  $V(x)$  is a constant, the solution of FSDE  $D_c^{2-2H} x = R_H(t)$  satisfies  $\text{Var}(x(t)) \propto t^{2-2H}$ . In other words, we have subdiffusion.*

*Remark 1* This agrees with the Langevin model in [19, Theorem 2.2], though the author was discussing the case with mass. Further, this corollary shows that the solution to our model usually is a fractional Brownian motion with a Hurst parameter  $1 - H \in (0, 1/2)$ , and this agrees with the data analysis in [18,22,29], implying that our model makes physical sense.

**Proposition 1** *There exists  $C > 0$  such that  $\mathbb{E}|G(t_2) - G(t_1)|^2 \leq C|t_2 - t_1|^{2H+2\alpha-2}$  and therefore  $G(t)$  is  $H + \alpha - 1 - \epsilon$  Hölder continuous for any  $\epsilon \in (0, H + \alpha - 1)$ .*

*Proof*

$$\mathbb{E}|G(t_2) - G(t_1)|^2 = \phi(t_2, t_2) + \phi(t_1, t_1) - 2\phi(t_1, t_2).$$

To be notationally convenient, let us define

$$\varphi(s, t) = \frac{B(\alpha, 1 - \alpha)\Gamma(\alpha)}{B(2H - 1, \alpha)} \phi(s, t).$$

Without loss of generality, we assume  $t_2 \geq t_1$ . Applying  $a + b \geq 2\sqrt{ab}$  whenever  $a \geq 0, b \geq 0$ , we have

$$\varphi(t_1, t_2) \geq 2 \int_0^{t_1} (t_2 - r)^{H+\alpha-3/2} (t_1 - r)^{H+\alpha-3/2} dr$$

If  $H + \alpha - 3/2 \leq 0$ , then,

$$\varphi(t_1, t_2) \geq 2 \int_0^{t_1} (t_2 - r)^{2H+2\alpha-3} dr = \frac{2}{2H + 2\alpha - 2} \left( t_2^{2H+2\alpha-2} - (t_2 - t_1)^{2H+2\alpha-2} \right).$$

Hence,

$$\begin{aligned}
 \mathbb{E}|G(t_2) - G(t_1)|^2 &\leq C_1 \left( t_1^{2H+2\alpha-2} - t_2^{2H+2\alpha-2} + 2(t_2 - t_1)^{2H+2\alpha-2} \right) \\
 &\leq 2C_1 (t_2 - t_1)^{2H+2\alpha-2},
 \end{aligned}$$

since  $0 < 2H + 2\alpha - 2 \leq 1, t_1^{2H+2\alpha-2} - t_2^{2H+2\alpha-2} \leq 0$ .

If  $H + \alpha - 3/2 > 0$ , then

$$\begin{aligned} \varphi(t_2, t_2) + \varphi(t_1, t_1) - 2\varphi(t_1, t_2) &= \int_{t_1}^{t_2} (t_2 - r)^{2H+2\alpha-3} dr \\ &\quad + \int_0^{t_1} \left( (t_2 - r)^{H+\alpha-3/2} - (t_1 - r)^{H+\alpha-3/2} \right)^2 dr. \end{aligned}$$

The first integral is easily seen to be bounded by  $C|t_2 - t_1|^{2H+2\alpha-2}$  for some constant  $C$ . For the second term, we have:

$$\left( (t_2 - r)^{H+\alpha-\frac{3}{2}} - (t_1 - r)^{H+\alpha-\frac{3}{2}} \right)^2 = \left( H + \alpha - \frac{3}{2} \right)^2 \left( \int_{t_1}^{t_2} (s - r)^{H+\alpha-\frac{5}{2}} ds \right)^2.$$

Let  $I_\epsilon = \left( \int_{t_1}^{t_2} (s - r + \epsilon)^{H+\alpha-5/2} ds \right)^2$  with  $r \leq t_1$ . Then,

$$\begin{aligned} \int_0^{t_1} I_\epsilon dr &\leq (t_2 - t_1) \int_0^{t_1} \int_{t_1}^{t_2} (s - r + \epsilon)^{2H+2\alpha-5} ds dr \\ &= (t_2 - t_1) \int_{t_1}^{t_2} \int_0^{t_1} \dots dr ds = \frac{(t_2 - t_1)}{|2H + 2\alpha - 4|(2H + 2\alpha - 3)} \\ &\quad \times \left( (t_2 - t_1 + \epsilon)^{2H+2\alpha-3} - \epsilon^{2H+2\alpha-3} - (s + \epsilon)^{2H+2\alpha-3} \Big|_{t_1}^{t_2} \right) \\ &\leq C_{\alpha,H} (t_2 - t_1) (t_2 - t_1 + \epsilon)^{2H+2\alpha-3}. \end{aligned}$$

Note that  $2H + 2\alpha - 4 < 0$ . Taking  $\epsilon \rightarrow 0$  shows that the second term is bounded by  $C(t_2 - t_1)^{2H+2\alpha-2}$ .

The Kolmogorov continuity criteria shows that  $G(t)$  is  $H + \alpha - 1 - \epsilon$  Hölder continuous for any  $\epsilon \in (0, H + \alpha - 1)$  almost surely, ending the proof. □

**Lemma 3** Let  $\{g_\alpha\}$  be the convolution group in [21]. In particular, for  $\alpha > -1$

$$g_\alpha(t) = \begin{cases} \frac{\theta(t)}{\Gamma(\alpha)} t^{\alpha-1}, & \alpha > 0, \\ \delta(t), & \alpha = 0, \\ \frac{1}{\Gamma(1+\alpha)} D(\theta(t)t^\alpha), & \alpha \in (-1, 0). \end{cases}$$

Here,  $\theta(t)$  is the Heaviside step function while  $D$  means the distributional derivative with respect to  $t$ . Let  $G_\alpha$  be the process  $G$  with the kernel indexed by  $\alpha$ . Let  $\alpha_1 \in (1 - H, 1)$  and  $\alpha_2 + \alpha_1 \in (1 - H, 1)$ . Then, it holds that  $g_{\alpha_2} * G_{\alpha_1} = G_{\alpha_1+\alpha_2}$ .

*Proof* It suffices to look at a continuous path of  $B_H$ . For such a path, we can mollify to  $B_H^\epsilon = B_H * \eta_\epsilon$  where  $\eta_\epsilon = \frac{1}{\epsilon} \eta(\frac{\cdot}{\epsilon})$  with  $\eta \in C_c^\infty(-\infty, 0)$ ,  $0 \leq \eta \leq 1$  and  $\int_{-\infty}^\infty \eta dt = 1$ . Then,  $g_{\alpha_2} * (g_{\alpha_1} * \frac{d}{dt} B_H^\epsilon) = g_{\alpha_1+\alpha_2} * \frac{d}{dt} B_H^\epsilon$  by [21]. Taking  $\epsilon \rightarrow 0$  and using the Hölder continuity of  $B_H$ , we arrive at the conclusion. □

### 4 Existence of the Strong Solutions

For the discussion on existence of solutions of a class of SDEs driven by fractional Brownian motion, one may refer to [13,32]. Our FSDEs are different from those studied in [13,32], as we have both the Caputo derivatives and fractional Brownian motions.

Mathematically, for fractional differential equations with Caputo derivative of order  $\alpha \in (0, 1)$ , we only need to specify the initial value at one point  $t = 0$  [5,6,21]. For our fractional

SDE, this is clear from Eq. (2.14). Intuitively, the system is activated at  $t = 0$  and one starts to count the memory effect from  $t = 0$ . For better understanding, our model is the over-damped case of the generalized Langevin equation, which is derived from Kac–Zwanzig model (see [11, 19]). In the Kac–Zwanzig model, specifying the value at  $t = 0$  is enough. Hence, to make the FSDE solvable, we only need to specify the data at  $t = 0$ .

We first define the so-called strong solution:

**Definition 1** Given a probability space  $(\Omega, \mathcal{F}, P)$  and a random variable  $x_0$  on this space, suppose  $B_H$  is a fractional Brownian motion over this space, which may be coupled to  $x_0$ . A Strong solution of the fractional stochastic differential equation (2.11) with initial condition  $x_0$  on the interval  $[0, T)$  ( $T > 0$ ) is a process  $x(t)$  that is continuous and adapted to the filtration  $(\mathcal{G}_t)$  with  $\mathcal{G}_t = \cap_{s>t} (\sigma(B_H(\tau), 0 \leq \tau \leq s) \cup \sigma(x_0))$ ,  $\forall t \in [0, T)$ , satisfying

- (1)  $P(x(0) = x_0) = 1$ .
- (2) With probability one, we have  $\forall t \in [0, T)$ , Eq.(2.14) holds.

It is standard to prove that the strong solution exists and is unique given the initial data since we have shown that process  $G$  is continuous. We state the theorem and put the proof in the ‘‘Appendix’’ for a reference:

**Theorem 1** Let  $H > 1/2$  and  $\alpha \in (1 - H, 1)$ . Assume that  $V'(\cdot)$  is Lipschitz continuous. Then, there exists a unique strong solution on  $[0, \infty)$  to the FSDE (2.11) for a given fractional Brownian motion and initial distribution in the sense of Definition 1.

If  $V'(x)$  is only locally Lipschitz, we probably need  $V$  to be confining, or in other words,  $\lim_{|x| \rightarrow \infty} V(x) = \infty$  and  $e^{-\beta V(x)} \in L^1(\mathbb{R})$  for any  $\beta > 0$  for the global existence of the solution. We are not going to pursue this issue any further in this work.

## 5 Asymptotic Analysis

In this section, we discuss the ergodicity and convergence to equilibrium satisfying Gibbs measure. This is important for a physical system. When the force is linear, we show rigorously that the process is ergodic and converges algebraically to the Gibbs measure if the ‘fluctuation-dissipation theorem’ is satisfied (see Theorem 2). When the force is nonlinear, we believe this problem must be solved by rewriting the FSDE model into Markovian processes and we propose two possible approaches for this.

### 5.1 Linear Force Case

Consider that  $V'(x) = kx$  for some  $k > 0$ . By Theorem 1, the solution exists and is unique. In this section, we will show rigorously for the linear force case that our model indeed has physical meaning.

For the convenience, we introduce the function

$$e_{\alpha,k}(t) = E_{\alpha}(-kt^{\alpha}), \tag{5.1}$$

where

$$E_{\alpha}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\alpha + 1)} \tag{5.2}$$

is the Mittag–Leffler function. Note that  $e_{\alpha,k}$  solves the equation

$$D_c^\alpha e_{\alpha,k} = -k e_{\alpha,k}, \quad e_{\alpha,k}(0) = 1.$$

See for example [23].

We can now state the result for asymptotic behaviors:

**Theorem 2** *Let  $V = \frac{1}{2}kx^2$ . As  $t \rightarrow \infty$ , the solution  $x(t)$  to the FSDE (2.11) converges in distribution to a normal distribution, i.e.,  $x(t)$  tends to a stationary Gaussian process:  $x_\infty(t)$ . The covariance  $h(\tau) = \mathbb{E}(x_\infty(t)x_\infty(t + \tau))$  of this stationary process satisfies*

$$\mathcal{F}(h(\tau)) = \frac{2\Gamma(2H + 1) \sin(H\pi)}{\Gamma(1 - \alpha)} \frac{|\omega|^{1-2H}}{|(i\omega)^\alpha + k|^2}, \tag{5.3}$$

where  $\mathcal{F}(\cdot)$  is the Fourier transform operator for tempered distributions. If  $\alpha = \alpha^*$  so that the ‘fluctuation-dissipation theorem’ is satisfied, the covariance is given exactly by

$$h(\tau) = \frac{1}{k} e_{\alpha,k}(\tau). \tag{5.4}$$

In particular,  $x_\infty(t)$  satisfies the Gibbs measure

$$\mu(dx) \sim \exp\left(-\frac{1}{2}kx^2\right) dx.$$

To prove this theorem, our approach is to find out the exact formulas for the solutions:

$$\begin{aligned} x(t) &= x_0 e_{\alpha,k}(t) + \left( G(t) + \int_0^t G(t-s) \dot{e}_{\alpha,k}(s) ds \right) \\ &= x_0 e_{\alpha,k}(t) - \frac{C_H}{k} \int_0^t \dot{e}_{\alpha,k}(t-\tau) dB_H(\tau) =: X_1 + X_2. \end{aligned} \tag{5.5}$$

Recall again that the dot means derivative on time. We will analyze the asymptotic behaviors of this stochastic process to conclude our claim.

Before we give the proof, we remark that the variance of the first term is  $\sim t^{-2\alpha}$  while the variance of the second term increases to the stationary variance with rate  $t^{2H-2-\alpha}$ . The loss of the variance of the first term can be balanced by the gain of the second term only if  $-2\alpha = 2H - 2 - \alpha$  or  $\alpha = \alpha^*$ . If  $\alpha$  is too small, then, the effect of initial data dampens slowly, or the dissipation caused by viscosity is small, which cannot balance the fluctuation. If  $\alpha$  is too big, then the effect of initial data dampens too fast due to strong dissipation. Hence, the fluctuation-dissipation theorem must be satisfied to model a true physical system so that there is balance. We also remark that, as we have seen, even if there is no balance between fluctuation and dissipation, the whole process will still tends to a normal distribution, though it might not be the correct physical equilibrium.

We now move to the proof of Theorem 2. To do that, we prove several auxiliary lemmas. We first of all introduce a lemma regarding the behavior of  $e_{\alpha,k}$ :

**Lemma 4**  $e_{\alpha,k}$ , the solution to the initial value problem

$$D_c^\alpha e_{\alpha,k} = -k e_{\alpha,k}, \quad e_{\alpha,k}(0) = 1,$$

is continuous on  $[0, \infty)$  and smooth on  $(0, \infty)$ . As  $t \rightarrow \infty$ ,

$$e_{\alpha,k}(t) = O(t^{-\alpha}).$$

The derivative is negative:  $\dot{e}_{\alpha,k}(t) < 0$ . As  $t \rightarrow 0^+$ ,  $\dot{e}_{\alpha,k}(t) \sim Ct^{\alpha-1}$ . Further, there exist  $C_1 > 0, C_2 > 0$  such that for  $t \geq 1$ ,

$$C_1 t^{-\alpha-1} \leq |\dot{e}_{\alpha,k}(t)| \leq C_2 t^{-\alpha-1}. \tag{5.6}$$

We have for  $t > 0$ ,

$$\dot{e}_{\alpha,k}(t) = -kg_{\alpha}(t) - \frac{k}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \dot{e}_{\alpha,k}(s) ds. \tag{5.7}$$

*Proof* The fact that  $e_{\alpha,k}$  is the solution to the IVP is well-known [23]. Denoting the Heaviside step function as  $\theta(t)$  and  $g_{\alpha}(t) = \frac{\theta(t)}{\Gamma(\alpha)} t^{\alpha-1}$ . Using the group technique and the inverse formula introduced in [21], we find that

$$\theta(t)e_{\alpha,k} = \theta(t) (1 + g_{\alpha} * (-k\theta(t)e_{\alpha,k})).$$

Taking the distributional derivative on both sides, we find that

$$\theta(t)\dot{e}_{\alpha,k} = -kg_{\alpha} - kg_{\alpha} * (\theta(t)\dot{e}_{\alpha,k}).$$

Since all distributions are locally integrable, the convolution can be written as Lebesgue integral and we have the equality (5.7).

By the series expansion of Mittag–Leffler functions (Eq. 5.2), we find the local behavior of  $\dot{e}_{\alpha,k}$  near  $t = 0$ . From the series expansion, it is seen that  $e_{\alpha,k}$  is strictly decreasing on  $(0, \infty)$ . The asymptotic behavior at  $t \rightarrow \infty$ , is obtained by Tauberian analysis [40] using the Laplace transforms of  $\dot{e}_{\alpha,k}(t)$  and  $\ddot{e}_{\alpha,k}(s)$  (noting the Laplace transform  $\mathcal{L}(\dot{e}_{\alpha,k}) = -\frac{k}{s^{\alpha+k}}$ ), or the asymptotic behavior of Mittag–Leffler function directly.  $\square$

**Lemma 5** *The solution to the FSDE (2.11) with  $V'(x) = kx$  is given by (5.5).*

*Proof* Since  $G(t)$  is almost surely continuous, we just solve the equation for each continuous sample path of  $G(t)$ . Let  $T > 0$ . We set  $\tilde{G}(t) = G(T)$  when  $t > T$  so that the Laplace transform of  $\tilde{G}$  exists. Consider the equation

$$y(t) = x(0) - \frac{k}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} y(s) ds + \tilde{G}(t).$$

Take the Laplace transform (denoted by  $\mathcal{L}$ ) on both sides. Since  $\mathcal{L}(t^{\alpha-1}) = \Gamma(\alpha)s^{-\alpha}$ , we find

$$\mathcal{L}(y) = \frac{x_0 s^{\alpha-1}}{s^{\alpha} + k} + \mathcal{L}(\tilde{G}) \left( 1 - \frac{k}{s^{\alpha} + k} \right).$$

We have by the Laplace transform of  $e_{\alpha,k}$  that

$$y(t) = x_0 e_{\alpha,k}(t) + \tilde{G}(t) + \int_0^t \tilde{G}(t-s) \dot{e}_{\alpha,k}(s) ds.$$

Clearly,

$$x(t) = y(t), \quad t \leq T.$$

Since  $T$  is arbitrary, then, we have for any  $t \geq 0$  that the unique solution  $x(t)$  is given by

$$x(t) = x_0 e_{\alpha,k}(t) + \left( G(t) + \int_0^t G(t-s) \dot{e}_{\alpha,k}(s) ds \right).$$



Now, we argue that  $X_2$  in (5.5) can also be written as

$$X_2(t) = -\frac{C_H}{k} \int_0^t \dot{e}_{\alpha,k}(t - \tau) dB_H(\tau).$$

We first of all rewrite

$$\int_0^t G(t - s) \dot{e}_{\alpha,k}(s) ds = \frac{C_H}{\Gamma(\alpha)} \int_0^t \int_0^{t-s} (t - s - \tau)^{\alpha-1} dB_H(\tau) \dot{e}_{\alpha,k}(s) ds.$$

As in the proof of Lemma 3, we may mollify the random path. Then, we can change the order of integration. Taking the mollifying parameter to zero, we get

$$\int_0^t G(t - s) \dot{e}_{\alpha,k}(s) ds = \frac{C_H}{\Gamma(\alpha)} \int_0^t \int_0^{t-\tau} (t - s - \tau)^{\alpha-1} \dot{e}_{\alpha,k}(s) ds dB_H(\tau).$$

By the identity for  $\dot{e}_{\alpha,k}$  (Eq. 5.7), we have

$$\frac{1}{\Gamma(\alpha)} \int_0^{t-\tau} (t - s - \tau)^{\alpha-1} \dot{e}_{\alpha,k}(s) ds = -g_\alpha(t - \tau) - \frac{1}{k} \dot{e}_{\alpha,k}(t - \tau).$$

This then yields

$$\int_0^t G(t - s) \dot{e}_{\alpha,k}(s) ds = -\frac{C_H}{\Gamma(\alpha)} \int_0^t (t - \tau)^{\alpha-1} dB_H(\tau) - \frac{C_H}{k} \int_0^t \dot{e}_{\alpha,k}(t - \tau) dB_H(\tau).$$

This then shows the claim. □

$X_2$  is Gaussian process since  $B_H$  is a Gaussian process. The mean of  $X_2$  is clearly zero. We can investigate the variance to see its asymptotic behavior.

*Remark 2*  $X_2$  never converges in  $L^p$  or almost surely, so we only consider convergence in distribution.

For notational convenience, let us denote

$$r(t) = -\dot{e}_{\alpha,k}(t) \geq 0. \tag{5.8}$$

By the isometry (3.7), we can compute that

$$\begin{aligned} \Sigma(t) &= \text{Var}(X_2(t)) = H(2H - 1) \frac{C_H^2}{k^2} \int_0^t \int_0^t r(t - s)r(t - \tau) |s - \tau|^{2H-2} d\tau ds \\ &= \frac{1}{k^2 \Gamma(1 - \alpha)} \int_0^t \int_0^t r(s)r(\tau) |\tau - s|^{2H-2} d\tau ds. \end{aligned} \tag{5.9}$$

**Lemma 6** *Let  $\alpha \in (1 - H, 1)$ .  $\Sigma = \lim_{t \rightarrow \infty} \Sigma(t)$  exists and there exist  $C_1 > 0, C_2 > 0$  such that*

$$C_1 t^{2H-2-\alpha} < \Sigma - \Sigma(t) < C_2 t^{2H-2-\alpha}. \tag{5.10}$$

*Proof* By Lemma 4 and (5.8),  $r$  is positive and

$$\int_0^\infty r(t) dt = 1.$$

By Lemma 4, there exist  $C_1 > 0, C_2 > 0$  such that for  $t \geq 1$

$$C_1 t^{-\alpha-1} \leq r \leq C_2 t^{-\alpha-1}.$$

Then, that  $\Sigma = \lim_{t \rightarrow \infty} \Sigma(t)$  exists is clear.

Consider the remainder  $\Sigma - \Sigma(t)$ , which is an integral over the region  $\mathbb{R}_{\geq 0}^2 \setminus [0, t] \times [0, t]$ . Due to the symmetry, we have

$$k^2 \Gamma(1 - \alpha)(\Sigma - \Sigma(t)) = 2 \int_t^\infty ds r(s) \int_0^s r(\tau)(s - \tau)^{2H-2} d\tau.$$

Consider that  $t$  is large and therefore  $s \geq t > 1$ . Below, the letter  $C$  denotes a generic constant which is independent of  $s$  and  $t$  but the concrete value could change from line to line. Denote the inside of the above integral as

$$J(s) = \int_0^s r(\tau)(s - \tau)^{2H-2} d\tau \leq \int_0^1 r(\tau)(s - \tau)^{2H-2} d\tau + \int_1^s C\tau^{-\alpha-1}|s - \tau|^{2H-2} d\tau.$$

The first term is controlled by  $(s - 1)^{2H-2} \int_0^1 r(\tau) d\tau$ . The second term is estimated as:

$$\begin{aligned} &Cs^{2H-2-\alpha} \left( \int_{1/s}^{1/2} z^{-1-\alpha}(1 - z)^{2H-2} dz + \int_{1/2}^1 z^{-1-\alpha}(1 - z)^{2H-2} dz \right) \\ &\leq Cs^{2H-2-\alpha} \left( 2^{2-2H} \frac{1}{\alpha}(s^\alpha - 2^\alpha) + \bar{C} \right) \leq Cs^{2H-2}, \end{aligned}$$

where  $\bar{C} = \int_{1/2}^1 z^{-1-\alpha}(1 - z)^{2H-2} dz$  independent of  $s$ . Hence by the asymptotic behavior of  $r$ ,

$$\Sigma - \Sigma(t) \leq C \int_t^\infty |r(s)|s^{2H-2} du \leq Ct^{2H-2-\alpha}.$$

For the other direction, we just note  $J(s) \geq s^{2H-2} \int_0^1 |r(\tau)| d\tau$ . □

*Proof of Theorem 2* By inspection of the solution (5.5), it is clear that  $X_1 \rightarrow 0$  almost surely and in  $L^2$  as  $t \rightarrow \infty$ . We only have to focus on  $X_2$ .

Since  $X_2$  is a Gaussian process with mean zero, by Lemma 6,  $\text{Var}(X_2)$  converges and thus  $X_2$  converges in distribution. We can now show the convergence of the covariance

$$\begin{aligned} h(\tau; t) &= \mathbb{E}(X_2(t)X_2(t + \tau)) \\ &= \frac{1}{k^2 \Gamma(1 - \alpha)} \int_0^t \int_0^{t+\tau} r(t - u)r(t + \tau - v)|u - v|^{2H-2} dv du \\ &= \frac{1}{k^2 \Gamma(1 - \alpha)} \int_0^t \int_0^{t+\tau} r(u)r(v)|v - \tau - u|^{2H-2} dv du. \end{aligned}$$

Denote

$$h(\tau) = \frac{1}{k^2 \Gamma(1 - \alpha)} \int_0^\infty \int_0^\infty r(u)r(v)|v - \tau - u|^{2H-2} dv du \geq 0. \tag{5.11}$$

We argue that there exists  $C > 0$  depending on  $H, \alpha, k$  such that  $h(\tau) \leq C$ . To do this, we use  $\tilde{r}(t)$  to represent the even extension of  $r$ . Then, we find

$$\begin{aligned} \int_0^\infty \int_0^\infty r(u)r(v)|v - \tau - u|^{2H-2} dv du &= \int_\tau^\infty \int_0^\infty r(u - \tau)r(v)|v - u|^{2H-2} dv du \\ &\leq \int_0^\infty \int_0^\infty \tilde{r}(u - \tau)r(v)|v - u|^{2H-2} dv du \\ &\leq \frac{1}{H(2H - 1)} \|\tilde{r}(\cdot - \tau)\|_{\mathcal{H}_0} \|r\|_{\mathcal{H}_0}. \end{aligned}$$

The last inequality follows from the fact that (3.7) gives the inner product. According to Lemma 4, the asymptotic behavior of  $r(t) = -\dot{e}_{\alpha,k}$  implies that

$$\|\tilde{r}(\cdot - \tau)\|_{1/H}^{1/H} \leq 2\|r\|_{1/H}^{1/H} < \infty$$

since  $\alpha \in (1 - H, 1)$ . Lemma 1 implies that  $h(\tau)$  is bounded.

This means that the integral in (5.11) converges as  $t \rightarrow \infty$ , for  $-\infty < \tau < \infty$ :

$$h(\tau; t) \rightarrow h(\tau) = \frac{1}{k^2 \Gamma(1 - \alpha)} \int_0^\infty \int_0^\infty r(u)r(v)|v - \tau - u|^{2H-2} dv du \leq C.$$

Consequently,  $X_2$  converges in distribution to a stationary process. The limit process  $x_\infty(t)$  has the covariance  $h(\tau)$  to be bounded.

Since  $h(\tau)$  is bounded, it is a tempered distribution. The Fourier transform exists. The following formal computation can be justified by considering  $h(\tau)e^{-\epsilon\tau^2}$  and then taking the limit  $\epsilon \rightarrow 0$  under the topology of the tempered distribution.

$$\begin{aligned} & \int_{-\infty}^\infty e^{-i\omega\tau} \int_0^\infty \int_\tau^\infty r(u - \tau)r(v)|u - v|^{2H-2} dudv d\tau \\ &= \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^u r(u - \tau)e^{-i\omega\tau} d\tau |u - v|^{2H-2} r(v) dudv. \end{aligned}$$

The inner most integral turns out to be

$$e^{-i\omega u} \int_0^\infty e^{i\omega\tau} r(\tau) d\tau = I(-i\omega)e^{-i\omega u},$$

with

$$I(s) = \frac{k}{s^\alpha + k}.$$

The whole thing turns out to be

$$I(-i\omega)I(i\omega) \int_{-\infty}^\infty e^{-i\omega z} |z|^{2H-2} dz = \frac{k^2}{|(i\omega)^\alpha + k|^2} (2\Gamma(2H + 1) \sin(H\pi)) |\omega|^{1-2H}.$$

This shows the first claim.

If  $\alpha = \alpha^* = 2 - 2H$ , we find that

$$\mathcal{F}(h(\tau)) = \frac{2 \sin(H\pi) |\omega|^{1-2H}}{|(i\omega)^\alpha + k|^2}.$$

Recall that we have the identity

$$\int_0^\infty e^{-ts} E_\alpha(-kt^\alpha) dt = \frac{s^{\alpha-1}}{s^\alpha + k}.$$

It follows that

$$\int_{-\infty}^\infty e^{-i\omega t} E_\alpha(-k|t|^\alpha) dt = 2 \frac{\operatorname{Re}((i\omega)^{\alpha-1})(k + (-i\omega)^\alpha)}{|k + (i\omega)^\alpha|^2} = \frac{2k \sin(\alpha\pi/2) |\omega|^{\alpha-1}}{|k + (i\omega)^\alpha|^2}.$$

Hence, we find in this case

$$h(\tau) = \frac{1}{k} e_{\alpha,k}(\tau).$$

It follows that the final equilibrium is a normal distribution with variance

$$\Sigma = h(0) = \frac{1}{k}$$

and the last claim follows. □

### 5.2 The General Case

We have proved that for linear regimes, when  $\alpha = \alpha^*$  is considered, the distribution converges to the Gibbs measure with algebraic rate. The linear forcing case is special, but it shows that our model makes physical meaning. For general forcing regimes with the ‘fluctuation-dissipation theorem’ satisfied ( $\alpha = \alpha^*$ ), proving the ergodicity and that the distribution converges to the Gibbs measure algebraically seems hard. We believe this problem can be solved by figuring out some Markovian representations. In the following, we propose two such possible Markovian embedding approaches that may be helpful for studying the asymptotic behavior.

#### 5.2.1 Infinitely Dimensional Ornstein–Uhlenbeck Process with Mixing

If the kernel  $\gamma(t)$  is the sum of finitely many exponentials, it is well known the GLE has a Markovian representation with a particular mixing (see [34] for the details) so that the Gibbs measure is an invariant measure. However, the result corresponding to a general kernel with fat tail is yet unknown. In our FSDE, the kernel  $\gamma(t) = \frac{\theta(t)}{\Gamma(1-\alpha)}t^{-\alpha}$  is of fat tail but it is completely monotone. A completely monotone function is the Laplace transform of a Radon measure on  $[0, \infty)$  by the famous Bernstein theorem [40]. In other words, the kernel  $\gamma(\cdot)$  can be written as superpositions of infinitely many exponentials. Based on this observation, we can formally rewrite our FSDE model to an infinite-dimensional OU process with mixing. We hope the techniques in [34] may be generalized to this infinite OU process to discuss the ergodicity of our FSDE model. This seems beyond the scope of this paper and we leave the rigorous discussion to future.

To understand the idea, we first of all consider the deterministic equation

$$D_c^\alpha x = \gamma(t) * (\theta(t)\dot{x}) = x, \quad x(0) = x_0. \tag{5.12}$$

It is well-known that the solution of this equation is  $x(t) = x_0 E_\alpha(t^\alpha)$ , which is continuous on  $[0, \infty)$  and smooth on  $(0, \infty)$ , and further  $\dot{x} \geq 0$  [21].

The kernel  $\gamma(t)$  is completely monotone and it can be written as

$$\gamma(t) = \int_0^\infty e^{-\lambda t} \rho(\lambda) d\lambda, \quad \rho(\lambda) = \frac{1}{B(\alpha, 1-\alpha)} \lambda^{\alpha-1}. \tag{5.13}$$

Here  $B(\cdot, \cdot)$  is the Beta function. We then decouple the fractional ODE (5.12) as an infinitely dimensional Markovian process with a mixing effect:

$$\begin{cases} 0 = x(t) + \xi(t), & t > 0, & x(0+) = x_0, \\ \dot{\xi}_\lambda(t) = -\lambda \xi_\lambda(t) - \sqrt{\rho(\lambda)} \dot{x}(t), & & \xi_\lambda(0) = 0, \\ \xi(t) = \lim_{\epsilon \rightarrow 0} \int_0^\infty e^{-\lambda \epsilon} \sqrt{\rho} \xi_\lambda(t) d\lambda. \end{cases} \tag{5.14}$$

We solve the second equation in (5.14) as

$$\xi_\lambda(t) = - \int_0^t \sqrt{\rho(\lambda)} e^{-\lambda(t-s)} \dot{x}(s) ds, \tag{5.15}$$

which implies that  $\xi$  in the third equation is well-defined. Provided  $\dot{x} \geq 0$ , we switch the order of integration for  $\xi$  and applying monotone convergence theorem,

$$\begin{aligned} \xi(t) &= -\lim_{\epsilon \rightarrow 0} \int_0^\infty \int_0^t \rho(\lambda) e^{-\lambda(t-s+\epsilon)} \dot{x}(s) ds d\lambda \\ &= -\lim_{\epsilon \rightarrow 0} \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s+\epsilon)^{-\alpha} \dot{x}(s) ds = -D_c^\alpha x(t). \end{aligned} \tag{5.16}$$

The equation  $x = D_c^\alpha x$ ,  $t > 0$  then follows. This system then decouples the memory to a system of uncountable Markovian functions with the simple mixing given by the third equation in (5.14).

*Remark 3* Let us mention a subtlety of the system: it seems that the initial value of  $x$  is unimportant as one can reduce the system to

$$\begin{aligned} \dot{\xi}_\lambda(t) &= -\lambda \xi_\lambda(t) + \sqrt{\rho(\lambda)} \dot{\xi}(t), \quad t > 0. \quad \xi_\lambda(0) = 0. \\ \xi(t) &= \lim_{\epsilon \rightarrow 0} \int_0^\infty e^{-\lambda \epsilon} \sqrt{\rho(\lambda)} \xi_\lambda(t) d\lambda. \end{aligned}$$

This seems to be solvable without considering  $x_0$ . Actually, this system is not well-posed. The reason is that the equation for  $\xi_\lambda$  may not be valid at  $t = 0$  and  $\lim_{t \rightarrow 0} \xi(t) \neq \xi(0) = 0$ . (In the original system,  $\xi(0) = \xi(0+) = x(0+)$  is equivalent to  $\lim_{t \rightarrow 0} D_c^\alpha x = 0$ .) We must know  $\lim_{t \rightarrow 0} \xi(t) = \lim_{t \rightarrow 0} D_c^\alpha x$  to start the process, which is equivalent to assigning the initial value of  $x$ .

Back to our FSDE (2.11), the computation for the deterministic case then leads us to consider:

$$\begin{cases} V'(x(t)) = \xi(t), & t > 0 \\ \xi(t) = \lim_{\epsilon \rightarrow 0^+} \int_0^\infty \xi_\lambda(t) e^{-\epsilon \lambda} \rho(\lambda)^{1/2} d\lambda, & t > 0 \\ \dot{\xi}_\lambda(t) = -\lambda \xi_\lambda(t) - \sqrt{\rho(\lambda)} \dot{x}(t) + \sqrt{2\lambda} \dot{W}_\lambda(t). \end{cases} \tag{5.17}$$

Here we assume  $\xi_\alpha(0)$ 's are i.i.d, normal with variance 1. This is a random system of differential algebraic equations (DAE), and clearly Markovian. The issue is that we have an uncountable-dimensional OU process driven by an uncountable-dimensional Wiener process (normal Brownian motion).

With the random noise, we may not be able justify the computation as we did for the deterministic cases. However, a formal computation may still be illustrating, through which we argue that this DAE system is equivalent to our FSDE. By solving  $\xi_\lambda$  formally, we have

$$\begin{aligned} \xi(t) &= \lim_{\epsilon \rightarrow 0^+} \left( \int_{[0,\infty)} \xi_\lambda(0) \sqrt{\rho} e^{-\lambda(t+\epsilon)} d\lambda + \int_{[0,\infty)} \int_0^t \sqrt{2\lambda \rho} e^{-\lambda(t-s+\epsilon)} dW_\lambda(s) d\lambda \right) \\ &\quad - \lim_{\epsilon \rightarrow 0} \int_{[0,\infty)} \int_0^t \rho(\lambda) e^{-\lambda(t-s+\epsilon)} \dot{x}(s) ds d\lambda =: R(t) + K(t). \end{aligned} \tag{5.18}$$

In the case  $t > 0, \tau \geq 0$ , we have

$$\begin{aligned} \mathbb{E}(R(t)R(t+\tau)) &= \int_{[0,\infty)} \rho e^{-\lambda(2t+\tau)} \text{Var}(\xi_0) d\lambda + \int_{[0,\infty)} \int_0^t 2\lambda \rho(\lambda) e^{-\lambda(2t+\tau-2s)} ds d\lambda \\ &= \gamma(\tau + 2t) + \gamma(\tau) - \gamma(\tau + 2t) = \gamma(\tau). \end{aligned} \tag{5.19}$$

Of course, the change of order of integration and expectation is not justified rigorously, but the computation is still interesting. Since both  $R(t)$  and  $C_H \dot{B}_H$  are Gaussian process and they have the same covariance, we can then identify them.

For the term  $K(t)$  in (5.18), since  $\rho(\lambda)e^{-\epsilon\lambda} \in L^1[0, \infty)$ , we may change the order of integration and  $K(t) = -D_c^\alpha x$  for  $t > 0$ . Hence,

$$\xi = -D^\alpha x + R(t), t > 0. \tag{5.20}$$

This then formally verifies that FSDE (2.11) can be obtained from the Markovian DAE system.

The same subtlety in Remark 3 appears here.  $\xi(0) \neq -D^\alpha x|_{t=0} + R(0)$ , which allows us to specify the initial condition  $x_0$ .

Since the Gibbs measures for the GLE with the kernel to be finitely exponentials are invariant measures [34], we think it is promising to show that Gibbs measures are the final equilibrium measures for our model. The discussion here provides a possible framework for the study of general  $V(x)$ . To study the stochastic DAE system, one may have to put some structure in the space of infinite-dimensional Gaussian process, and then somehow figure out that the Gibbs measure for the whole system is an invariant measure. This will then be left for future.

### 5.2.2 A Heat Bath Model

In this subsection, we summarize the heat bath model proposed in [16,38] for the generalized Langevin equation. The key point is that one can consider the whole dynamics of the particle together with the heat bath, which is Markovian. If one integrates out the degrees of freedom for the heat bath, one obtains the GLE. The whole heat bath model is the continuous version of the Kac–Zwanzig model mentioned in [11,19,42]. We think this heat bath model may be another promising direction to study the ergodicity and the asymptotic behavior of our FSDE model. Formally, if one takes the  $m \rightarrow 0$  limit for the special kernel  $\gamma(t) \propto |t|^{-\alpha}$  (the discussion in [16,38] is not applicable to this kernel though), our FSDE can be obtained. This limit for the classical Langevin equation (Eq. 1.1) is called the Smoluchowski–Kramers approximation [10] and the limit for generalized Langevin equation has not been studied yet to our best knowledge. We will summarize the formulation here briefly and then give a brief discussion to connect it with our FSDE model.

Assume that the particle is put in a heat bath modeled by infinitely many free phonons and the corresponding scalar field  $\varphi$  is given by the massless Klein–Gordon equation (which is a wave equation),

$$(-\partial_t^2 + \Delta)\varphi = 0. \tag{5.21}$$

The Lagrangian density is  $\mathcal{L} = -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi$ , where  $\mu$  goes over the time-spatial coordinate in relativity, and the Hamiltonian is

$$\mathcal{H}_h = \frac{1}{2} \int_{\mathbb{R}^n} (|\nabla\varphi|^2 + |\pi|^2)dx, \tag{5.22}$$

where  $\pi = \partial_t\varphi$  should be regarded as a new variable.

This Hamiltonian motivates that the correct space for the heat bath is

$$\mathcal{V} = H^1(\mathbb{R}^n) \otimes L^2(\mathbb{R}^n)$$

with the inner product given by

$$\langle f, g \rangle = \int_{\mathbb{R}^n} (\nabla f_1 \cdot \nabla g_1 + f_2 g_2) dx, \quad \forall f = (f_1, f_2) \in \mathcal{V}, \quad g = (g_1, g_2) \in \mathcal{V}. \quad (5.23)$$

Note that Gaussian measures can be constructed over this Hilbert space.  $\forall f, g \in \mathcal{V}$  and  $\xi$  is an  $\mathcal{V}$ -valued random variable satisfying a Gaussian measure  $\mu_{\phi_0}^\beta$  indexed by  $\phi_0 \in \mathcal{V}$  and  $\beta > 0$ , then,

$$\mathbb{E}(\langle f, \xi - \phi_0 \rangle \langle \xi - \phi_0, g \rangle) = \beta^{-1} \langle f, g \rangle. \quad (5.24)$$

The coupling between the particle and the heat bath is given by

$$\mathcal{H}_I = \int_{\mathbb{R}^n} \varphi(x) \rho(q - x) dx = \int_{\mathbb{R}^n} \varphi(x) \rho(x - q) dx,$$

where  $\rho$  is a radially symmetric function which can be understood as the coupling strength. In literature [16, 38],  $\rho$  is assumed to be in  $L^2$ , so that the coupling strength is finite and can be approximated by the dipole expansion:

$$\mathcal{H}_I = \int_{\mathbb{R}^n} \rho q \cdot \nabla \varphi dx + \frac{q^2}{2} \int_{\mathbb{R}^n} |\rho^2| dx. \quad (5.25)$$

The second term is some correction added to make the model clean so that the GLE can be derived from this model.

The total Hamiltonian that describes the coupling between the particle and the heat bath is given by

$$\begin{aligned} \mathcal{H} &= \frac{1}{2m} p^2 + V(q) + \frac{1}{2} \int_{\mathbb{R}^n} (|\pi|^2 + |\nabla \varphi|^2) dx + \int_{\mathbb{R}^n} \rho q \cdot \nabla \varphi dx + \frac{q^2}{2} \int_{\mathbb{R}^n} |\rho^2| dx \\ &= \frac{1}{2m} p^2 + V(q) + \frac{1}{2} \int_{\mathbb{R}^n} |\nabla \varphi + q\rho|^2 + |\pi|^2 dx. \end{aligned} \quad (5.26)$$

where  $\lim_{|q| \rightarrow \infty} V(q) = \infty$  and  $\exp(-\beta V(\cdot)) \in L^1(\mathbb{R}^n)$  for any  $\beta > 0$ .

With this coupling, the authors in [16, 38] showed that the particle satisfies the generalized Langevin equation obeying the ‘dissipation-fluctuation theorem’ provided the initial data satisfy a certain Gaussian measure. The GLE for  $n = 1$  case is written as

$$\begin{aligned} \dot{q} &= v, \quad m\dot{v} = -V'(q) - \int_0^t \gamma(t-s)\dot{q}(s) ds + R(t), \\ \gamma(t) &= \int_{\mathbb{R}} |\hat{\rho}|^2 e^{ikt} dk, \quad \mathbb{E}(R(t)R(s)) = \gamma(|t-s|). \end{aligned}$$

With this result, the authors conclude the following:

**Proposition 2** *Suppose  $R(t)$  is a 1D stationary Gaussian process with mean zero and*

$$\mathbb{E}(R(t)R(s)) = \gamma(|t-s|). \quad (5.27)$$

*If  $\gamma$  is the Fourier transform of an  $L^1(\mathbb{R})$  even nonnegative function, then there exists a coupling between  $q(0) = q_0$  and  $R(t)$  so that the equation*

$$\dot{q} = v, \quad m\dot{v} = -V'(q) - \int_0^t \gamma(t-s)\dot{q}(s) ds + R(t) \quad (5.28)$$

admits the Gibbs measure

$$\mu(dqdv) \propto \exp\left(-\frac{mv^2}{2} - V(q)\right) dqdv, \tag{5.29}$$

as the invariant measure.

For any initial distribution  $\mu^0$  that is absolutely continuous with respect to  $\mu$  and any coupling between  $q_0$  and  $R(t)$ ,  $\mu^t$  converges weakly to the Gibbs measure  $\mu$ .

Our FSDE model is similar to the problems studied in [16, 38], except that  $\gamma(t) \propto |t|^{-\alpha}$  and  $m = 0$ . Note that the kernel  $|t|^{-\alpha}$  is not the Fourier transform of an  $L^1$  kernel. One can therefore mollify  $\gamma$  by

$$\gamma_\epsilon(t) = \eta_\epsilon * \gamma(t), \tag{5.30}$$

so that  $\gamma_\epsilon$  is the Fourier transform of an  $L^1$  kernel. One can then study the GLE with kernel  $\gamma_\epsilon$ . If final equilibrium is preserved with  $\epsilon \rightarrow 0$  limit, then the Gibbs measure is the equilibrium measure for the GLE with kernel  $|t|^{-\alpha}$ . Then, formally, the Smoluchowski–Kramers approximation  $m \rightarrow 0$  limit (if valid) yields that the Gibbs measure proportional to  $\exp(-V(q))$  is the final equilibrium measure of our FSDE (2.14). This provides another possible framework for general potential  $V(x)$  and we leave the rigorous study for future.

*Remark 4* The Smoluchowski–Kramers approximation ( $m \rightarrow 0$  limit) for the usual Langevin equations has been discussed in [10]. However, for the generalized Langevin equation, the limit  $m \rightarrow 0$  is subtle. The limit equation for a general kernel  $\gamma$  may not be a good initial value problem. The initial value problem

$$\int_0^t \gamma(t-s)\dot{q}(s) ds = -V'(q) + R(t), \quad q(0) = q_0$$

generally admits no continuous solution if  $\gamma(t)$  is bounded. Hence, the possible approach is to show first that convergence to Gibbs measure is valid for the GLE when  $\gamma(t) \propto |t|^{-\alpha}$  and then show the  $m \rightarrow 0$  limit can pass to the final equilibrium measures.

**Acknowledgements** The work of J.-G Liu is partially supported by KI-Net NSF RNMS11-07444 and NSF DMS-1514826. The work of J. Lu is supported in part by National Science Foundation under Grant DMS-1454939. J. Lu would also like to thank Eric Vanden-Eijnden for helpful discussions.

### Appendix: Proof of Theorem 1

*Proof* We just consider a sample point  $x_0$  and a sample path  $G$  with  $G$  being continuous. We then construct a path that satisfies the integral equation given this sample initial data.

By Proposition 1,  $G(t)$  is continuous. Consider the sequence given by

$$x^{(0)} = x_0,$$

and  $x^{(n)}$ ,  $n \geq 1$  is given by

$$x^{(n)}(t) = x_0 - \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} V'(x^{(n-1)}(s)) ds + G(t).$$

Assume  $L$  is a Lipschitz constant for  $V'(\cdot)$ . Introducing  $g_\gamma = \frac{\theta(t)}{\Gamma(\gamma)} t^{\gamma-1}$ , we find that  $\{g_\gamma\}_{\gamma>0}$  forms a convolution semigroup (Lemma 3). We define

$$e^n = x^{(n)} - x^{(n-1)}.$$



Explicit formula tells us that

$$e^1 = -V'(x_0)g_{\alpha+1} + G(t),$$

and that

$$|e^n| = | -g_\alpha * (V'(x^{n-1}) - V'(x^{n-2}))| \leq Lg_\alpha * |e^{n-1}|, \quad n \geq 2.$$

Hence,

$$|e^n| \leq L^{n-1}g_{(n-1)\alpha} * |e^1|.$$

Direct computation shows that  $\sup_{0 \leq t \leq T} g_{(n-1)\alpha} * |e^1|$  decays exponentially in  $n$ . Hence,  $\sum_n |e^n|$  converges. It follows that  $\sum_n e^n$  converges uniformly on any interval  $[0, T]$  with  $T \in (0, \infty)$ . The limit is also a continuous function. It turns out that the limit satisfies the integral equation.

For the uniqueness, assume that both  $x(t)$  and  $y(t)$  are solutions. Then, we take a sample where both  $x(t)$  and  $y(t)$  are continuous. For this sample,  $\forall t > 0$ ,

$$|x(t) - y(t)| = \frac{1}{\Gamma(\alpha)} \left| \int_0^t (t-s)^{\alpha-1} (V'(x(s)) - V'(y(s))) ds \right| \leq L(g_\alpha * |x - y|)(t).$$

Applying this inequality iteratively and using the semi-group property of  $g_\gamma$ , we find

$$|x - y|(t) \leq L^n g_{n\alpha} * |x - y|.$$

Fixing  $T > 0$ , the right hand side goes to zero uniformly on  $[0, T]$ . Then, we find that  $x = y$  on  $[0, T]$  for this sample path. Since both solutions are continuous almost surely, then  $x = y$  on  $[0, T]$  almost surely. By the arbitrariness of  $T$ ,  $x = y$  almost surely. The uniqueness then is shown. This then completes the proof of the theorem.  $\square$

## References

1. Biagini, F., Hu, Y., Øksendal, B., Zhang, T.: Stochastic Calculus for Fractional Brownian Motion and Applications. Springer, New York (2008)
2. Callen, H.B., Welton, T.A.: Irreversibility and generalized noise. Phys. Rev. **83**(1), 34 (1951)
3. Decreusefond, L., Ustünel, A.S.: Stochastic analysis of the fractional Brownian motion. Potential Anal. **10**(2), 177–214 (1999)
4. Deng, W., Barkai, E.: Ergodic properties of fractional Brownian-Langevin motion. Phys. Rev. E **79**(1), 011112 (2009)
5. Diethelm, K.: The Analysis of Fractional Differential Equations: An Application-oriented Exposition Using Differential Operators of Caputo Type. Springer, Berlin (2010)
6. Diethelm, K., Ford, N.: Analysis of fractional differential equations. J. Math. Anal. Appl. **265**(2), 229–248 (2002)
7. Duncan, T.E., Hu, Y., Pasik-Duncan, B.: Stochastic calculus for fractional Brownian motion I. Theory. SIAM J. Control Optim. **38**(2), 582–612 (2000)
8. Felderhof, B.: On the derivation of the fluctuation-dissipation theorem. J. Phys. A **11**(5), 921 (1978)
9. Ford, G., Kac, M., Mazur, P.: Statistical mechanics of assemblies of coupled oscillators. J. Math. Phys. **6**(4), 504–515 (1965)
10. Freidlin, M.: Some remarks on the Smoluchowski-Kramers approximation. J. Stat. Phys. **117**(3–4), 617–634 (2004)
11. Givon, D., Kupferman, R., Stuart, A.: Extracting macroscopic dynamics: model problems and algorithms. Nonlinearity **17**(6), R55 (2004)
12. Gorenflo, R., Mainardi, F.: Fractional Calculus: Integral and Differential Equations of Fractional Order. Springer, Wien (1997)
13. Hairer, M.: Ergodicity of stochastic differential equations driven by fractional Brownian motion. Ann. Probab. **33**, 703–758 (2005)

14. Hairer, M., Majda, A.J.: A simple framework to justify linear response theory. *Nonlinearity* **23**(4), 909 (2010)
15. Hairer, M., Pillai, N.S.: Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré B* **47**, 601–628 (2011)
16. Jakšić, V., Pillet, C.: Ergodic properties of the non-Markovian Langevin equation. *Lett. Math. Phys.* **41**(1), 49–57 (1997)
17. Kilbas, A.A., Srivastava, H.M., Trujillo, J.J.: *Theory and Applications of Fractional Differential Equations*, vol. 204. Elsevier, Amsterdam (2006)
18. Kou, S., Xie, X.S.: Generalized langevin equation with fractional gaussian noise: subdiffusion within a single protein molecule. *Phys. Rev. Lett.* **93**(18), 180603 (2004)
19. Kou, S.C.: Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *Ann. Appl. Stat.* **2**, 501–535 (2008)
20. Kubo, R.: The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **29**(1), 255 (1966)
21. Li, L., Liu, J.G.: A generalized definition of Caputo derivatives and its application to fractional ODEs. arXiv preprint [arXiv:1612.05103v2](https://arxiv.org/abs/1612.05103v2) (2017)
22. Magdziarz, M., Weron, A., Burnecki, K., Klafter, J.: Fractional brownian motion versus the continuous-time random walk: a simple test for subdiffusive dynamics. *Phys. Rev. Lett.* **103**(18), 180602 (2009)
23. Mainardi, F., Paradisi, P., Gorenflo, R.: Probability distributions generated by fractional diffusion equations. arXiv preprint [arXiv:0704.0320](https://arxiv.org/abs/0704.0320) (2007)
24. Mandelbrot, B.B., Van Ness, J.W.: Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**(4), 422–437 (1968)
25. Marconi, U., Puglisi, A., Rondoni, L., Vulpiani, A.: Fluctuation-dissipation: response theory in statistical physics. *Phys. Rep.* **461**(4), 111–195 (2008)
26. Marty, R., Sølna, K.: A general framework for waves in random media with long-range correlations. *Ann. Appl. Probab.* **2**, 115–139 (2011)
27. Mémin, J., Mishura, Y., Valkeila, E.: Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion. *Stat. Probab. Lett.* **51**(2), 197–206 (2001)
28. Mikosch, T., Norvaiša, R.: Stochastic integral equations without probability. *Bernoulli* **6**, 401–434 (2000)
29. Min, W., Luo, G., Cherayil, B.J., Kou, S., Xie, X.S.: Observation of a power-law memory kernel for fluctuations within a single protein molecule. *Phys. Rev. Lett.* **94**(19), 198302 (2005)
30. Mori, H.: A continued-fraction representation of the time-correlation functions. *Prog. Theor. Phys.* **34**(3), 399–416 (1965)
31. Nualart, D.: Fractional Brownian motion: stochastic calculus and applications. *Int. Congr. Math.* **3**, 1541–1562 (2006)
32. Nualart, D., Ouknine, Y.: Regularization of differential equations by fractional noise. *Stoch. Proc. Appl.* **102**(1), 103–116 (2002)
33. Nyquist, H.: Thermal agitation of electric charge in conductors. *Phys. Rev.* **32**(1), 110 (1928)
34. Ottobre, M., Pavliotis, G.: Asymptotic analysis for the generalized Langevin equation. *Nonlinearity* **24**(5), 1629 (2011)
35. Pavliotis, G.A.: *Stochastic Processes and Applications, Diffusion Processes, the Fokker-Planck and Langevin Equations*. Springer, New York (2014)
36. Pipiras, V., Taqqu, M.: Integration questions related to fractional Brownian motion. *Probab. Theory Relat. Fields* **118**(2), 251–291 (2000)
37. Pipiras, V., Taqqu, M.: Are classes of deterministic integrands for fractional Brownian motion on an interval complete? *Bernoulli* **7**(6), 873–897 (2001)
38. Rey-Bellet, L., Thomas, L.E.: Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. *Commun. Math. Phys.* **225**(2), 305–329 (2002)
39. Sakthivel, R., Revathi, P., Ren, Y.: Existence of solutions for nonlinear fractional stochastic differential equations. *Nonlinear Anal. Theor.* **81**, 70–86 (2013)
40. Widder, D.: *The Laplace Transform*. Princeton University Press, Princeton (1941)
41. Zähle, M.: Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Relat. Fields* **111**(3), 333–374 (1998)
42. Zwanzig, R.: Nonlinear generalized Langevin equations. *J. Stat. Phys.* **9**(3), 215–220 (1973)