Fractional stochastic differential equations satisfying fluctuation-dissipation theorem

Lei Li*[1], Jian-Guo Liu†[2], and Jianfeng Lu‡[3]

[1]Department of Mathematics, Duke University, Durham, NC 27708, USA.
[2] Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA.
[3]Department of Mathematics, Department of Physics, and Department of Chemistry, Duke University, Box 90320, Durham NC 27708, USA.

Abstract

We consider in this work stochastic differential equation (SDE) model for particles in contact with a heat bath when the memory effects are non-negligible. 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 based on this we consider fractional stochastic differential equations (FSDEs), which should be understood in an integral form. We establish the existence of strong solutions for such equations. In the linear forcing regime, we compute the solutions explicitly and analyze the asymptotic behavior, through which we verify that satisfying fluctuation-dissipation indeed leads to the correct physical behavior. We further discuss possible extensions to nonlinear forcing regime, while leave the rigorous analysis for future works.
Keywords Fractional SDE, Fluctuation-dissipation-theorem, Caputo derivative, Fractional Brownian motion, Generalized Langevin equation

1 Introduction

1.1 Physical Background

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

*leili@math.duke.edu
†jliu@phy.duke.edu
‡jianfeng@math.duke.edu

velocity of the particle

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

where dot denotes derivative on time, $-\gamma u$ counts for friction and η 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

$$mdu = -\gamma u dt + \sqrt{2D_x}dW,$$

where W is a standard Brownian motion and D_x is some constant to be determined. If we add the equation for position and consider external force, one has the Langevin equation:

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

Since the friction $-\gamma u$ and random force η both come from interactions between the particle and the environment, they should be related. The 'fluctuation-dissipation theorem' [1] provides a precise connection between them, such that the correlation 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 and then justified by Callen and Welton [1, 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 [3]. 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:

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

In [4, 5], 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:

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

[1]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.

where $R(t)$ is some random force. Note that the friction γ becomes a kernel now. For the particle to achieve equilibrium at the prescribed temperature, the force $R(t)$ and the friction kernel γ must be related. Without the external force (i.e. $\nabla V = 0$), Kubo assumed that $\mathbb{E}(u(t_0)R(t)) = 0, t > t_0$ and that u is a stationary process. He derived formally (though he used the existence of the one-sided Fourier transform of γ, 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

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

There are other formal derivations as well (e.g. [6]). These derivations are not fully convincing though on the mathematical rigorous level. In [5], Kubo assumed the relation $\mathbb{E}(u(t_0)R(t)) = 0, t > t_0$ arguing using causality. The issue is though $R(t)$ does not affect $u(t_0)$, $u(t_0)$ can affect $R(t)$. In [6], Felderholf 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 [7, 8, 9, 10]). 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. From this model, 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', all in the context of generalized Langevin equations. In [11], the author tried to rephrase the 'fluctuation-dissipation theorems' and the related linear response theory in mathematical language. Hairer and Majda in [12] developed a framework to justify the use the linear response theory through the 'fluctuation-dissipation theorem' for studying climate models.

In the following discussion, we will simply set $kT = 1$ for convenience, and the variables k and T might be used to denote other quantities.

1.2 The motivation of FSDE

Motivated by the discussions in [10, 13], we consider the random force given by the (distributional) derivative of fractional Brownian motion $R \sim \dot{B}_H$. To understand this, let us first review the basics of the fractional Brownian motion B_H. (See [14, 15] for more detailed discussions.)

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

$$\mathbb{E}(B_t^H B_s^H) = R_H(s,t) = \frac{1}{2}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right), \tag{1.6}$$

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: $B_H(t) \overset{\mathrm{d}}{=} a^{-H} B_H(at)$ where '$\overset{\mathrm{d}}{=}$' 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 [14], the following representation is given

$$\begin{aligned} B_H(t) = C_1(H) \left(\int_0^t (t-s)^{H-\frac{1}{2}} dW(s) + \int_{-\infty}^0 ((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}) dW(s) \right) \\ = C_1(H) \int_{-\infty}^0 (-r)^{H-\frac{1}{2}} (dW(r+t) - dW(r)), \end{aligned} \tag{1.7}$$

where W is a normal Brownian motion and $C_1(H)$ is a constant to make (1.6) valid. This is also used in [16]. In [17, 18], one uses

$$B_H(t) = C_2(H) \int_0^t (t-s)^{H-\frac{1}{2}} F\left(H - \frac{1}{2}, \frac{1}{2} - H, H + \frac{1}{2}, 1 - \frac{t}{s}\right) dW(s), \tag{1.8}$$

where F is the Gauss hypergeometric function. Another representation in [19] 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

$$\begin{aligned} \lim_{h \to 0^+, h_1 \to 0} \mathbb{E}\left(\frac{B_H(h)}{h} \frac{B_H(t+h_1) - B_H(t)}{h_1} \right) \\ = \lim_{h \to 0^+, h_1 \to 0} \frac{1}{2hh_1} \left((t+h_1)^{2H} - (t+h_1-h)^{2H} - t^{2H} + (t-h)^{2H} \right) \\ = H(2H-1) t^{2H-2}. \end{aligned} \tag{1.9}$$

If we pick the initial time in (1.4) as $t_0 = 0$ and choose the random noise as

$$R_H(t) = \frac{1}{\sqrt{H(2H-1)\Gamma(2H-1)}}\dot{B}_H(t), \tag{1.10}$$

we then have the model

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

following the 'fluctuation-dissipation theorem'.

We will assume throughout the paper that

$$H \in \left(\frac{1}{2}, 1\right), \tag{1.11}$$

as they are the physically most realistic regimes [10] and consequently $2-2H \in (0,1)$.

In the cases that inertia can be neglected, it is natural to consider the over-damped equation with fractional noise:

$$\frac{1}{\Gamma(2H-1)}\int_0^t (t-s)^{2-2H}u(s)ds = -\nabla V(x) + R_H(t). \tag{1.12}$$

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

$$D_c^\alpha v = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{\dot{v}(s)}{(t-s)^\alpha}ds. \tag{1.13}$$

In [20] by two of the authors, a definition of the Caputo derivative that relies on a convolution group was proposed so that it can be defined for a large class of locally integrable functions, which agrees with (1.13) when the function is continuous on $[0,t]$ and has a weak derivative on $(0,t]$. Note that $u(s) = \dot{x}(s)$, the left hand side of Equation (1.12) formally becomes the Caputo derivative of x with $\alpha = 2-2H$ and the equation becomes a fractional SDE:

$$D_c^{2-2H}x = -\nabla V(x) + R_H(t). \tag{1.14}$$

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 α:

$$D_c^\alpha x = -V'(x) + C_H\dot{B}_H, \tag{1.15}$$

where

$$C_H = \frac{1}{\sqrt{H(2H-1)\Gamma(1-\alpha)}} \tag{1.16}$$

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 kernel associated with the Caputo derivative as

$$\gamma(t) = \frac{1}{\Gamma(1-\alpha)} t^{-\alpha}. \tag{1.17}$$

By [20], we may de-convolve and change the Caputo derivative to integral form as

$$x(t) = x(0) - \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} V'(x(s))ds + \frac{C_H}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} dB_H, \tag{1.18}$$

where we formally used $R_H ds = C_H dB_H$. This integral will then be understood as the rigorous definition of the FSDE (1.15). The last term in (1.18) is an integral with respect to fractional Brownian motion, which we will make the meaning precise later.

While FSDEs have been discussed in some previous works already, the above equation (1.18) motivated by the 'fluctuation-dissipation theorem' seems to be new. In [18, 16, 21], they discussed the FSDEs driven by fractional Brownian motions but there is no memory effect in the dissipating term. In [22], 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.

In this work, we will study FSDE (1.18) and try to understand the role of the 'fluctuation-dissipation theorem'. For convenience, we denote

$$G(t) = \frac{C_H}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} dB_H(s) = \int_0^\infty f_t(s) dB_H(s), \tag{1.19}$$

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 the next section.

The rest of the paper is organized as follows. In Section 2, we study the process $G(t)$ (1.19) for FSDE (1.18) in detail, including the meaning of the stochastic integral and the regularity property of $G(t)$. Based on the study of $G(t)$, in Section 3, we show the existence and uniqueness of strong solutions for FSDE (1.18) on the interval $[0, \infty)$ provided $V'(\cdot)$ is Lipschtiz continuous. In Section 4, we focus on the asymptotic behavior of the strong solutions of (1.18), and argue that satisfying the 'fluctuation-dissipation theorem' leads the correct physical behavior. 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. We discuss that when the 'fluctuation-dissipation theorem' is satisfied, there is balance between the dissipation and fluctuation effect from the random forcing and that the Gibbs measure is the final equilibrium distribution. In the nonlinear regime, we argue formally that when the 'fluctuation-dissipation theorem' is satisfied, the FSDE can be reduced from some Markovian processes in infinite dimensions. These Markovian approaches might be useful for studying our FSDE. We can see formally from the second approach that the final equilibrium distribution should be the Gibbs measure as well. The rigorous study of the nonlinear regimes is left for future works.

2 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.

2.1 Stochastical integrals driven by fractional Brownian motions

The stochastic integrals with respect to fractional Brownian motions have been thoroughly discussed in literature [23, 24, 17, 25]. In [23, 24], the stochastic integrals are defined pathwise using the Riemann-Stieltjes integrals by making use of certain properties of the paths. In [17, 25], 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 [15, 26]. In the case that the integrand is deterministic, those two definitions agree. By (1.18), 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 $\mathscr{E}$ be the set of all step functions on $[0, T]$, i.e. $\varphi \in \mathscr{E}$ is given by

$$\varphi = \sum_{j=1}^{m} a_j 1_{(t_{j-1}, t_j]}(t), \tag{2.1}$$

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

$$B^H(\varphi) = \int_0^T \varphi \, dB_H(t) = \sum_{j=1}^{m} a_j \Big(B_H(t_j) - B_H(t_{j-1}) \Big). \tag{2.2}$$

Consider the inner product

$$\langle \varphi_1, \varphi_2 \rangle_{\mathscr{H}} = \mathbb{E}(B^H(\varphi_1) B^H(\varphi_2)). \tag{2.3}$$

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

$$\langle \varphi_1, \varphi_2 \rangle_{\mathscr{H}} = H(2H-1)\int_0^T \int_0^T |r-u|^{2H-2}\varphi_1(r)\varphi_2(u)dudr$$
$$= \frac{\pi\kappa(2\kappa+1)}{\Gamma(1-2\kappa)\sin(\pi\kappa)}\int_0^T s^{-2\kappa}(I^\kappa u^\kappa f)(s)(I^\kappa u^\kappa g)(s)ds, \quad (2.4)$$

where $\kappa = H - \frac{1}{2}$ and I^κ is the right Riemann-Liouville fractional calculus, given by ([19]):

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

This then motivates the definition of

$$\mathscr{H}_0 = \left\{\varphi \in L^1_{loc}[0,T] : \int_0^T \int_0^T |r-u|^{2H-2}|\varphi(r)||\varphi(u)|drdu < \infty\right\} \quad (2.5)$$

and

$$\Lambda = \left\{f \in L^1_{loc}[0,T] : \int_0^T s^{-2\kappa}(I^\kappa u^\kappa f)^2(s)ds < \infty\right\}. \quad (2.6)$$

Clearly, $\mathscr{H}_0 \subset \Lambda$. The integral $B^H(\varphi)$ can then be defined for $\varphi \in \Lambda$ by approximating them with step functions. In [27, 19], it is shown that both inner product spaces $\mathscr{H}_0$ and Λ are not complete and therefore not Hilbert spaces. However, the space $B^H(\mathscr{E})$ 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 $\mathscr{H}$ be the space of the closure of $\mathscr{E}$ under the inner product (2.3) and thus $\mathscr{H}$ contains some distributions. $\forall \varphi_1, \varphi_2 \in \mathscr{H}_0 \subset \mathscr{H}$,

$$\langle \varphi_1, \varphi_2 \rangle_{\mathscr{H}} = \mathbb{E}(B^H(\varphi_1)B^H(\varphi_2)) = H(2H-1)\int_0^T \int_0^T |r-u|^{2H-2}\varphi_1(r)\varphi_2(u)dudr. \quad (2.7)$$

The following lemma provides a convenient way to check that some deterministic processes can be integrated by fractional Brownian motion ([28, 15]):

Lemma 1. *If $H > 1/2$ and $\varphi \in L^{1/H}([0,T])$, then*

$$\|\varphi\|_{\mathscr{H}_0} \le b_H \|\varphi\|_{L^{1/H}[0,T]}. \quad (2.8)$$

where

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

2.2 Some basic properties of G

We can easily verify that $f_t \in L^{1/H}[0,T]$ whenever $t \le T$, and hence the integral on $[0,T]$ is well defined. Further, for any $T_1 > t, T_2 > t$, the integral 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*

$$\phi(t_1,t_2) = \mathbb{E}(G(t_1)G(t_2)) = \frac{B(2H-1,\alpha)}{B(\alpha,1-\alpha)\Gamma(\alpha)}\times \int_0^{\min(t_1,t_2)} dr\Big((t_1-r)^{\alpha-1}(t_2-r)^{2H-2+\alpha} + (t_2-r)^{\alpha-1}(t_1-r)^{2H-2+\alpha}\Big). \quad (2.9)$$

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

$$\beta_H = \frac{\sqrt{2}}{\sqrt{\Gamma(3-2H)}}. \quad (2.10)$$

In other words, $G(t)$ is a fractional Brownian motion with Hurst parameter $1-H$ up to a constant β_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 $t_2 \ge t_1 \ge 0$. The covariance can be computed using the isometry (2.7)

$$\mathbb{E}(G(t_1)G(t_2)) = \langle f_{t_1}, f_{t_2}\rangle_{\mathscr{H}} = \frac{1}{\Gamma(\alpha)B(\alpha,1-\alpha)}\int_0^{t_1}\int_0^{t_2} |r-u|^{2H-2}(t_1-r)^{\alpha-1}(t_2-u)^{\alpha-1}dudr.$$

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

$$I_1 = \frac{1}{\Gamma(\alpha)B(\alpha,1-\alpha)}\iint_{u\ge r}\dots dudr,\ \ I_2 = \frac{1}{\Gamma(\alpha)B(\alpha,1-\alpha)}\iint_{r\ge u}\dots dudr.$$

By explicit computation,

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

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

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

If $\alpha = \alpha^* = 2-2H$, the integrals can be evaluated exactly and we have

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

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 $\mathrm{Var}(x) \propto t^{2-2H}$. In other words, we have subdiffusion.*

This agrees with the Langevin model in [10, Theorem 2.2], though the author was discussing the case with mass.

Lemma 3. *Suppose $0 < \beta \le 1$ and $a \ge b \ge 0$. Then,*

$$a^\beta - b^\beta \le (a-b)^\beta. \tag{2.11}$$

Proof. This claim follows trivially from $\int_0^b (a-r)^{\beta-1}dr \le \int_0^b (b-r)^{\beta-1}dr$. □

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

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 $t_2 \ge t_1$. Applying $a+b \ge 2\sqrt{ab}$ whenever $a \ge 0, b \ge 0$, we have

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

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

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

Hence,

$$\mathbb{E}|G(t_2)-G(t_1)|^2 \leq C_1\Big(t_1^{2H+2\alpha-2} - t_2^{2H+2\alpha-2} + 2(t_2-t_1)^{2H+2\alpha-2}\Big) \leq 2C_1(t_2-t_1)^{2H+2\alpha-2},$$

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

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

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

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

Note that $2H+2\alpha-4 < 0$. Taking $\epsilon \to 0$ 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. $\square$

Lemma 4. *Let $\{g_\beta\}$ be the convolution group in [20]. In particular, for $\beta > -1$*

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

D here means the distributional derivative. 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 $\eta_\epsilon = \frac{1}{\epsilon}\eta(\frac{t}{\epsilon})$ with $\eta \in C_c^\infty(-\infty, 0)$, $0 \le \eta \le 1$ and $\int \eta dt = 1$. 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 [20]. Taking $\epsilon \to 0$ and using the Hölder continuity of B_H, we arrive at the conclusion. □

3 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 [18]. However, our FSDEs are different from those studied as we have both the Caputo derivatives and fractional Brownian motions. 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* (1.15) *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 $\mathcal{G}_t = \cap_{s>t}(\sigma(B_H(\tau), 0 \le \tau \le s) \cup \sigma(x_0))$, $\forall t \in [0, T)$, satisfying*

(1) $P(x(0) = x_0) = 1$.

(2) With probability one, we have $\forall t \in [0, T)$, Equation (1.18) *holds.*

We now prove that the strong solution exists and is unique given the initial data.

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* (1.15) *for a given fractional Brownian motion and initial distribution in the sense of Definition 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 $g_\gamma = \frac{1}{\Gamma(\gamma)} t^{\gamma-1}$, we find that $\{g_\gamma\}_{\gamma>0}$ forms a convolution semigroup. 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

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

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, $\forall t > 0$,

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

Applying this inequality iteratively and using the semi-group property of g_γ, 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. $\square$

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

4 Asymptotic analysis

4.1 Linear force case

Consider that $V'(x) = kx$ for some $k > 0$. By Theorem 1, the solution exists and is unique. For each continuous sample path $G(t)$, the equation can be solved exactly. To see this, let $T > 0$. We set $\widetilde{G}(t) = G(T)$ when $t > T$. Consider the equation

$$x(t) = x(0) - \frac{k}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} x(s)ds + \widetilde{G}(t). \tag{4.1}$$

As all functions are continuous, we can then take the Laplace transform (denoted by $\mathcal{L}$) on both sides. Since $\mathcal{L}(t^{\alpha-1}) = \Gamma(\alpha)s^{-\alpha}$, we find

$$\mathcal{L}(x) = \frac{x_0 s^{\alpha-1}}{s^\alpha + k} + \mathcal{L}(\widetilde{G})\Big(1 - \frac{k}{s^\gamma + k}\Big). \tag{4.2}$$

Denote

$$e_{\alpha,k}(t) = E_\alpha(-kt^\alpha), \tag{4.3}$$

where

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\alpha+1)} \tag{4.4}$$

is the Mittag-Leffler function. We have by the Laplace transform of $e_{\alpha,k}$ that

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

Recall again that the dot means derivative on time. This is valid for $t \le T$. Since T is arbitrary, then, we have for any $t \ge 0$:

$$x(t) = x_0 e_{\alpha,k}(t) + \left(G(t) + \int_0^t G(t-s)\dot{e}_{\alpha,k}(s)ds\right) =: X_1 + X_2. \tag{4.5}$$

Since $G(t)$ is a Gaussian process, $X_2 = G(t) + \int_0^t G(t-s)\dot{e}_{\alpha,k}(s)ds$ is Gaussian. The mean of X_2 is clearly zero. We can investigate the variance to see its asymptotic behavior.

We first of all introduce a lemma regarding the behavior of $e_{\alpha,k}$:

Lemma 5. *$e_{\alpha,k}$ solves the equation $D_c^\alpha e_{\alpha,k} = -ke_{\alpha,k}$, $e_{\alpha,k}(0) = 1$. It is continuous on $[0,\infty)$ and smooth on $(0,\infty)$. $e_{\alpha,k} = O(t^{-\alpha})$ as $t \to \infty$. $\dot{e}_{\alpha,k}(t) < 0$ and $\dot{e}_{\alpha,k}(t) \sim Ct^{\alpha-1}$ when $t \to 0+$. There exist $C_1 > 0, C_2 > 0$ such that for $t \ge 1$,*

$$C_1 t^{-\alpha-1} \le |\dot{e}_{\alpha,k}(t)| \le C_2 t^{-\alpha-1}. \tag{4.6}$$

Denoting the Heaviside step function as $u(t)$ and $g_\alpha = \frac{u(t)}{\Gamma(\alpha)} t^{\alpha-1}$, we have

$$u(t)\dot{e}_{\alpha,k} = -kg_\alpha - kg_\alpha * (u(t)\dot{e}_{\alpha,k}). \tag{4.7}$$

Proof. The fact that $e_{\alpha,k}$ is the solution to the IVP is well-known. One can refer to [29, 20]. Using the group technique and the inverse formula introduced in [20], we find that

$$u(t)e_{\alpha,k} = u(t)\left(1 + g_\alpha * (-ku(t)e_{\alpha,k})\right).$$

Taking the distributional derivative on both sides, we find that

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

Since all distributions are locally integrable, they can be understood in the Lebesgue sense and we have the equality.

By the series expansion of Mittag-Leffler functions (Eq. (4.4)), 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 $t \to \infty$, is obtained by Tauberian analysis ([30]) using the Laplace transforms of $\dot{e}_{\alpha,k}(t)$ and $\ddot{e}_{\alpha,k}(s)$ (Note $\mathcal{L}(\dot{e}_{\alpha,k}) = -\frac{k}{s^\alpha + k}$), or the asymptotic behavior of Mittag-Leffler function directly. $\square$

Computing the variance of X_2 directly yields an integral that is hard to evaluate. To compute the variance, we should find an alternative form of X_2. As in [10], one can compute formally that,

$$\mathcal{L}(G) = \frac{C_H}{\Gamma(\alpha)} \int_0^\infty e^{-st} \int_0^t (t-\tau)^{\gamma-1} dB_H(\tau) = C_H s^{-\gamma} \int_0^\infty e^{-s\tau} dB_H(\tau).$$

Hence, $\mathcal{L}(G)(1 - \frac{\lambda}{s^\gamma + \lambda}) = C_H \int_0^\infty \frac{e^{-s\tau}}{s^\gamma + \lambda} dB_H$. This then motivates the following alternative form of X_2 which we will prove in another way:

Lemma 6. *The process X_2 can be written as*

$$X_2 = -\frac{C_H}{k} \int_0^t \dot{e}_{\alpha,k}(t-\tau) dB_H(\tau). \tag{4.8}$$

Proof. 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 4, we may mollify the random path. Then, we can change the order of integration. Taking the mollifying parameter to zero, we get

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

By the identity for $\dot{e}_{\alpha,k}$ (Eq. (4.7)), we have

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

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. □

To be notational convenient, let us denote

$$r(t)=-\dot{e}_{\alpha,k}\geq 0. \tag{4.9}$$

By the isometry, we can compute that

$$\begin{aligned}\sigma(t)=\mathrm{Var}(X_2(t))&=H(2H-1)\frac{C_H^2}{k^2}\int_0^t\int_0^t r(t-u)r(t-v)|u-v|^{2H-2}dvdu\\ &=\frac{1}{k^2\Gamma(1-\alpha)}\int_0^t\int_0^t r(u)r(v)|v-u|^{2H-2}dvdu. \qquad (4.10)\end{aligned}$$

Lemma 7. *Let $\alpha\in(1-H,1)$. $\sigma=\lim_{t\to\infty}\sigma(t)$ exists and there exist $C_1>0, C_2>0$ such that*

$$C_1t^{2H-2-\alpha}<\sigma-\sigma(t)<C_2t^{2H-2-\alpha}. \tag{4.11}$$

Proof. Note that r is positive. By the formula of r, we find

$$\int_0^\infty rdt=1.$$

By Lemma 5, there exist $C_1>0, C_2>0$ such that for $t\geq 1$

$$C_1t^{-\alpha-1}\leq r\leq C_2t^{-\alpha-1}.$$

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

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

$$k^2\Gamma(1-\alpha)(\sigma-\sigma(t))=2\int_t^\infty dur(u)\int_0^u r(v)(u-v)^{2H-2}dv.$$

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

$$J(u) = \int_0^u r(v)(u-v)^{2H-2}dv \leq \int_0^1 r(v)(u-v)^{2H-2}dv + \int_1^u Cv^{-\alpha-1}|u-v|^{2H-2}dv.$$

The first term is controlled by $(u-1)^{2H-2}\int_0^1 r(v)dv$. The second term

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

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

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

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

Theorem 2. *Let $V = \frac{1}{2}kx^2$. As $t \to \infty$, $x(t)$ in (4.5) converges in distribution to a normal distribution, i.e., $x(t)$ tends to a stationary Gaussian process: $x_\infty(t)$. The covariance $\theta(\tau) = \mathbb{E}(x_\infty(t)x_\infty(t+\tau))$ of this stationary process satisfies*

$$\mathcal{F}(\theta(\tau)) = \frac{2\Gamma(2H+1)\sin(H\pi)}{\Gamma(1-\alpha)}\frac{|\omega|^{1-2H}}{|(i\omega)^\alpha + k|^2}, \tag{4.12}$$

where $\mathcal{F}(\cdot)$ is the Fourier transform operator for tempered distributions. If $\alpha = \alpha^*$, the covariance is given exactly by

$$\theta(\tau) = \frac{1}{k}e_{\alpha,k}(\tau). \tag{4.13}$$

In particular, if $\alpha = \alpha^*$, $x_\infty(t)$ satisfies the Gibbs measure

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

Proof. By inspection of the solution (4.5), it is clear that $X_1 \to 0$ almost surely and in L^2 as $t \to \infty$. We only have to focus on X_2.

Since X_2 is a Gaussian process with mean zero, we only have to show that $\mathrm{Var}(X_2)$ converges. We compute the covariance

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

Hence, as $t \to \infty$,

$$\sigma(\tau;t) \to \theta(\tau) = \frac{1}{k^2\Gamma(1-\alpha)}\int_0^\infty\int_0^\infty r(u)r(v)|v-\tau-u|^{2H-2}dvdu.$$

This is valid for $-\infty < \tau < \infty$ by the decay rate of r. Hence, X_2 converges in distribution to a normal distribution.

It is not hard to show that $\theta(\tau)$ is bounded, and therefore it is a tempered distribution. The Fourier transform exists. The following formal computation can be justified by considering $\theta(\tau)e^{-\epsilon\tau^2}$ and then taking the limit $\epsilon \to 0$ under the topology of the tempered distribution.

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

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

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

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

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

Hence, we find in this case

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

If follows that the final equilibrium is a normal distribution with variance $1/k$ and the last claim follows. □

Remark 1. *It is clear that X_2 never converges in L^p or almost surely, as the random force is being present for all the time.*

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 $t^{2H-2-\alpha}$. Hence, 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 α 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 α 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.

4.2 The general case

We have seen that for linear regimes, when $\alpha = \alpha^*$ is considered, the distribution converges to the Gibbs measure with algebraic rate. For general regimes, even for this critical case $\alpha = \alpha^*$, proving that the distribution converges to a stationary process algebraically seems hard. In the following, we propose two possible Markovian embedding approaches that may be helpful for studying the asymptotic behavior. We believe that for general $V(\cdot)$, algebraic convergence to the Gibbs measure is still true if the 'fluctuation-dissipation theorem' is satisfied, i.e., $\alpha = \alpha^*$.

4.2.1 A first Markovian approach

If the kernel $\gamma(t)$ is the sum of finitely many exponentials, then, the GLE has a Markovian representation (see [31] for the details). In our FSDE, the kernel $\gamma(t) = \frac{1}{\Gamma(1-\alpha)}t^{-\alpha}$ is completely monotone. By the famous Bernstein theorem [30], any completely function on $(0,\infty)$ is the Laplace transform of a Radon measure on $[0,\infty)$. In other words, the kernel $\gamma(\cdot)$ can be written as superpositions of infinitely many exponentials. Based on this fact, we will generalize the representation in [31] formally to our case here, though the total mass for the measure is not finite.

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

$$D_c^\alpha x = \gamma(t) * (u(t)\dot{x}) = x, \quad x(0) = x_0. \tag{4.14}$$

where $x(\cdot)$ is some unknown continuous function and $\dot{x}$ is understood as the distributional derivative. It is well-known that the solution of this equation is $x = x_0 E_\alpha(t^\alpha)$, which is continuous on $[0,\infty)$ and smooth on $(0,\infty)$, and further $\dot{x} \ge 0$ [20].

The kernel $\gamma(t)$ is completely monotone and $\gamma = \int_{[0,\infty)} e^{-\lambda t}\mu(d\lambda)$. It turns out the Radon measure μ is absolutely continuous with respect to the Lebesgue measure:

$$\gamma(t) = \int_0^\infty e^{-\lambda t}\mu(d\lambda) = \int_0^\infty e^{-\lambda t}\rho(\lambda)d\lambda. \tag{4.15}$$

By explicit computation, we find that

$$\rho(\lambda) = \frac{1}{B(\alpha, 1-\alpha)}\lambda^{\alpha-1}. \tag{4.16}$$

where $B(\cdot,\cdot)$ is the Beta function.

Consider the following system

$$\begin{cases} x = \xi, \quad t > 0, & x(0+) = x_0, \\ \dot{\xi}_\lambda = -\lambda\xi_\lambda + \sqrt{\rho}\dot{x}, & \xi_\lambda(0) = 0, \\ \xi = \lim_{\epsilon\to 0}\int_0^\infty e^{-\lambda\epsilon}\sqrt{\rho}\xi_\lambda d\lambda. \end{cases} \tag{4.17}$$

From the second equation, one obtains that

$$\xi_\lambda = \int_0^t \sqrt{\rho}e^{-\lambda(t-s)}\dot{x}(s)ds, \tag{4.18}$$

which implies that ξ_λ is Lebesgue-measurable in λ and ξ in the third equation is well-defined. Provided the properties of x in advance, $\int \rho e^{-\lambda(t-s+\epsilon)}|\dot{x}|ds$ is convergent for $t > 0$. Switching the order of integration and applying monotone convergence theorem ($\dot{x}$ is positive),

$$\xi = \lim_{\epsilon\to 0}\int_0^\infty \int_0^t \rho e^{-\lambda(t-s+\epsilon)}\dot{x}dsd\alpha = \lim_{\epsilon\to 0}\frac{1}{\Gamma(1-\alpha)}\int_0^t (t-s+\epsilon)^{-\alpha}\dot{x}(s)ds = D_c^\alpha x. \tag{4.19}$$

The first equation tells us that

$$x = D_c^\alpha x,\ \ t > 0, \tag{4.20}$$

which recovers the equation. This system then decouples the memory to a system of uncountable Markovian functions.

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

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

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

Remark 2. *Note that the new system is Markovian in the sense that if all the information of $\xi_\lambda(t_1)$ and also $\xi(t_1+) = \xi(t_1)$ are given at some $t_1 > 0$, then we can continue the system without knowing the history. For example, we can solve that*

$$\xi_\lambda(t) = \xi_\lambda(t_1)e^{-\lambda(t-t_1)} + \int_{t_1}^t e^{-\lambda(t-s)}\sqrt{\rho}\dot{\xi}(s)ds.$$

By the same approach, we find for $t > t_1$ that

$$x(t) = \lim_{\epsilon\to 0}\int_0^\infty e^{-\lambda\epsilon}e^{-\lambda(t-t_1)}\sqrt{\rho}\xi_\lambda(t_1)d\lambda + D_{c,t_1}^\alpha x. \tag{4.21}$$

where the symbol $D_{c,t_1}^\alpha x$ means the Caputo derivative starting from t_1. We need to know all the $\xi_\lambda(t_1)$ values to continue. Note that we must also specify $\xi(t_1+) = \xi(t_1)$ to continue the process. This is the same subtlety as what we discussed above for $t = 0$. The reason is that we do not require Equation (4.21) to be valid at $t = t_1$. If we specify it to be valid at $t = t_1$, then $\lim_{t\to t_1} D_{c,t_1}^\alpha x = 0$, which is equivalent to the continuity condition $\xi(t_1+) = \xi(t_1)$.

Back to our FSDE (1.15), the computation for the deterministic case then leads us to consider:

$$\begin{cases} \xi = -V'(x), & t > 0 \\ \xi = \lim_{\epsilon\to 0^+}\int_0^\infty \xi_\lambda e^{-\epsilon\lambda}\rho^{1/2}d\lambda, & t > 0 \\ \dot{\xi}_\lambda = -\lambda\xi_\lambda + \sqrt{\rho}\dot{x}(t) + \sqrt{2\lambda}\dot{W}_\lambda(t). & \end{cases} \tag{4.22}$$

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

Clearly, as long as we have 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 ξ_λ formally, we have

$$\xi(t)=\lim_{\epsilon\to0^+}\int_{[0,\infty)}\xi_\lambda(0)\sqrt{\rho}e^{-\lambda(t+\epsilon)}d\lambda+\lim_{\epsilon\to0^+}\int_{[0,\infty)}\int_0^t\rho(\lambda)e^{-\lambda(t-s+\epsilon)}\dot{x}(s)dsd\lambda$$
$$+\lim_{\epsilon\to0}\int_{[0,\infty)}\int_0^t\sqrt{2\lambda\rho}e^{-\lambda(t-s+\epsilon)}dW_\lambda(s)d\lambda. \quad (4.23)$$

Denote the random noise as

$$R(t)=\lim_{\epsilon\to0}\int_{[0,\infty)}\xi_\alpha(0)\sqrt{\rho}e^{-\lambda(t+\epsilon)}d\lambda+\int_{[0,\infty)}\int_0^t\sqrt{2\lambda\rho}e^{\lambda(t+\epsilon-s)}dW_\lambda(s)d\lambda. \quad (4.24)$$

In the case $t>0,\tau\ge0$, we have

$$\mathbb{E}(R(t)R(t+\tau))=\int_{[0,\infty)}\rho e^{-\lambda(2t+\tau)}\operatorname{Var}(\xi_0)d\lambda+\int_{[0,\infty)}\int_0^t 2\lambda\rho(\alpha)e^{-\lambda(2t+\tau-2s)}dsd\lambda$$
$$=\gamma(2(\tau+2t))+\gamma(\tau)-\gamma(2(\tau+2t))=\gamma(\tau). \quad (4.25)$$

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.

We now check the other term. Since $\rho e^{-\epsilon\lambda}\in L^1[0,\infty)$, we may change the order of integration and have

$$\lim_{\epsilon\to0}\int_0^t\gamma(t-s+\epsilon)dx(s)=D_c^\alpha x, t>0. \quad (4.26)$$

Hence,

$$\xi=D^\gamma x+R(t), t>0. \quad (4.27)$$

This then formally verifies that FSDE (1.15) can be obtained from the Markovian DAE system.

The same subtlety appears here. As $t\to0$, the integral $\int_0^t\gamma(t-s)dx(s)$ may not vanish. This means the limit $\lim_{t\to0}$ and the limit $\lim_{\epsilon\to0}$ can not be switched. Hence,

$\xi(0) \neq D^\alpha x|_{t=0} + R(0)$. Formally, that $\lim_{t\to 0+} \xi(t)$ has a nonzero limit which has nothing to do with $\{\xi_\lambda(0)\}$ allows us to specify the initial condition x_0.

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.

4.2.2 A second Markovian approach

We find that the formulation proposed in [32, 33] for the generalized Langevin equation may be another promising direction to study the asymptotic behavior of our FSDE. (This formulation is the continuous version of the Kac-Zwanzig model mentioned in [8, 9, 10].) Formally, if one takes the $m \to 0$ limit for the special kernel, our FSDE can be obtained. This limit for the classical Langevin equation (Eq. (1.1)) is called the Smoluchowski-Kramers approximation [34] and the limit for generalized Langevin equation has not been studied yet to our best knowledge. We will summarize the formulation here with some modifications that are better suited to our case and then give a brief discussion.

Assume that the heat bath is modeled by infinitely many free phonons and the corresponding scalar field φ is given by the massless Klein-Gordon equation, which is the standard wave equation,

$$(-\partial_t^2 + \Delta)\varphi = 0. \tag{4.28}$$

The Lagrangian density of this equation reads

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi, \tag{4.29}$$

where μ goes over the time-spatial coordinate in relativity. This then motivates the Hamiltonian of the heat bath

$$\mathcal{H}_h = \frac{1}{2}\int_{\mathbb{R}^n}(|\nabla\varphi|^2 + |\pi|^2)dx, \tag{4.30}$$

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

$$\langle f, g\rangle = \int_{\mathbb{R}^n}(\nabla f_1 \cdot \nabla g_1 + f_2 g_2)dx, \forall f = (f_1, f_2) \in \mathcal{V}, g = (g_1, g_2) \in \mathcal{V}. \tag{4.31}$$

Note that Gaussian measures can be constructed over this Hilbert space. $\forall f, g \in \mathcal{V}$ and ξ 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,

$$\mathbb{E}(\langle f, \xi - \phi_0\rangle\langle \xi - \phi_0, g\rangle) = \beta^{-1}\langle f, g\rangle. \tag{4.32}$$

Formally, one can understand the Gaussian measure as centered at $\phi_0 = (\varphi_0, \pi_0)$ and

$$\mu_{\phi_0}^{\beta}(d\phi) = C_\beta \exp\left(-\frac{\beta}{2}\langle \phi - \phi_0, \phi - \phi_0 \rangle\right) d\phi,$$

where C_β is a normalization constant and $d\phi$ is like a 'Lebesgue measure' in the Hilbert space. To be convenient, later we use μ_{ϕ_0} to mean $\mu_{\phi_0}^1$.

The coupling between the particle we consider with 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 ρ is a radially symmetric function which can be understood as the coupling strength. In literature [32, 33], ρ is assumed to be in L^2, so that the coupling strength is finite. The cases where ρ is not square integrable must be taken as the ideal limit under certain regimes. If $\rho \in L^2$, the interaction Hamiltonian then can be approximated by the dipole expansion:

$$\mathcal{H}_I = q \cdot \int_{\mathbb{R}^n} \nabla\varphi\rho dx + \frac{q^2}{2}\int_{\mathbb{R}^n} |\rho^2| dx. \tag{4.33}$$

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 is then given by

$$\begin{aligned}\mathcal{H} = \frac{1}{2m}p^2 + V(q) + \frac{1}{2}\int_{\mathbb{R}^n}(|\pi|^2 + |\nabla\varphi|^2)dx + q \cdot \int_{\mathbb{R}^n} \nabla\varphi\rho dx + \frac{q^2}{2}\int_{\mathbb{R}^n}|\rho^2|dx \\ = \frac{1}{2m}p^2 + V(q) + \frac{1}{2}\int_{\mathbb{R}^n} |\nabla\varphi + q\rho|^2 + |\pi^2| dx. \end{aligned}\tag{4.34}$$

where $\lim_{|q|\to\infty} V(q) = \infty$ and $\exp(-\beta V(\cdot)) \in L^1(\mathbb{R}^n)$ for any $\beta > 0$. At this point, it is convenient to find $\alpha_j \in H^1$ so that $\partial_i \alpha_j = \rho\delta_{ij}$. This can be constructed easily in Fourier space

$$\hat{\alpha}_j(k) = \frac{n\hat{\rho}(k)k^j}{i|k|^2}. \tag{4.35}$$

Note that $\hat{\rho}$ is also radially symmetric so that $\int nk_1^2 f(|k|)dk = \int |k|^2 f(|k|)dk$. Denote $\alpha = (\alpha_j)$. Then,

$$\nabla\varphi + q\rho = \nabla(\varphi + q\cdot\alpha). \tag{4.36}$$

The system of equations then are given by

$$\dot{q} = \frac{\partial\mathcal{H}}{\partial p} = p/m, \quad \dot{p} = -\frac{\partial\mathcal{H}}{\partial q} = -V'(q) - \int_{\mathbb{R}^n} \nabla(\varphi + q\cdot\alpha)\rho dx, \tag{4.37}$$

$$\dot{\varphi} = \frac{\delta\mathcal{H}}{\delta\pi} = \pi, \quad \dot{\pi} = -\frac{\delta\mathcal{H}}{\delta\varphi} = \Delta(\varphi + q\cdot\alpha). \tag{4.38}$$

Introducing $u = p/m$, the Hamiltonian system admits an invariant measure proportional to $\exp(-\mathcal{H})$, which is the Gibbs measure

$$d\mu = Z^{-1}\left(\exp\Big(-\frac{mu^2}{2} - V(q)\Big)dqdu\right) \times d\mu_{-q\cdot\alpha}, \tag{4.39}$$

where Z is a normalization constant. Conditioning on q, the measure for (φ, π) is a Gaussian measure $\mu_{-q\cdot\alpha}$ in $\mathcal{V}$. If we average out the heat bath variables, the marginal distribution of for the particle is exactly,

$$\mu_A(dqdu) = Z_A^{-1}\exp\left(-\frac{mu^2}{2} - V(q)\right)dqdu, \tag{4.40}$$

with Z_A being a normalization constant.

From here on, we consider only $n = 1$ (general dimension is similar but the notations are messier) and assume that the initial distribution is given by

$$d\mu|_{t=0} = \mu_A^0(q,u) \times \mu_{-q\cdot\alpha}. \tag{4.41}$$

This means we allow the measure the the particle to be arbitrary but the heat bath is Gaussian conditioning on the particle position, i.e. $(\varphi + q(0)\alpha, \pi)$ is mean zero Gaussian conditioning on $q(0)$.

Let $\phi = \begin{pmatrix} \varphi + q(t)\alpha \\ \pi \end{pmatrix}$. The second group of equation can be written in the vector form

$$\partial_t\phi = \mathcal{A}\phi + \begin{pmatrix} \dot{q}\alpha \\ 0 \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} 0 & 1 \\ \partial_x^2 & 0 \end{pmatrix}. \tag{4.42}$$

Then,

$$\phi = e^{\mathcal{A}t}\phi_0 + \int_0^t e^{\mathcal{A}(t-s)}\begin{pmatrix} \dot{q}\alpha \\ 0 \end{pmatrix} ds. \tag{4.43}$$

Define $\xi = (\alpha, 0)$ (α is a scalar since $n = 1$). Then

$$\int_{\mathbb{R}} \partial_x(\varphi + q\alpha)\rho dx = \langle \phi, \xi\rangle, \tag{4.44}$$

where $\langle\cdot,\cdot\rangle$ is the inner product (4.31) for $n = 1$. By taking the Fourier transform, it can be shown then that

$$-\left\langle \int_0^t e^{\mathcal{A}(t-s)}\begin{pmatrix} \dot{q}\alpha \\ 0 \end{pmatrix} ds, \xi \right\rangle = -\int_0^t \gamma(t-s)\dot{q}(s)ds, \tag{4.45}$$

where

$$\gamma(t) = \int_{\mathbb{R}} |\hat{\rho}|^2 e^{ikt} dk. \tag{4.46}$$

Since $\rho \in L^2(\mathbb{R})$, $\gamma(t)$ is an even function and continuous and $\gamma(0) = \int |\rho|^2 dx$.

Hence, we obtain

$$\dot{q} = u, \quad m\dot{u} = -V'(q) - \int_0^t \gamma(t-s)\dot{q}(s) + R(t), \tag{4.47}$$

where

$$R(t) = -\left\langle \partial_x \phi_0, e^{-\mathcal{A}t}\xi \right\rangle. \tag{4.48}$$

Recall $\phi_0 = (\varphi + q(0)\alpha, \pi)$ satisfies the mean zero Gaussian distribution conditioning on $q(0)$, $R(t)$ is also a mean zero Gaussian process. It can then be verified easily by Equation (4.32) that

$$\mathbb{E}(R(t)R(s)) = \gamma(|t-s|), \tag{4.49}$$

regardless of the value $q(0)$. It follows that $R(t)$ is a stationary process. Note that $R(t)$ and $q(0)$ are not independent! Different from [32] where the GLE is in a different form and the convolution is between a kernel and q instead of $\dot{q}$, we impose the initial condition (4.41) so that the random force $R(t)$ satisfies the usual 'fluctuation-dissipation theorem' (1.5).

By the results in [32, 33], we can summarize the following claim for $n = 1$:

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

$$\mathbb{E}(R(t)R(s)) = \gamma(|t-s|). \tag{4.50}$$

If γ 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

$$\dot{q} = u, \quad m\dot{u} = -V'(q) - \int_0^t \gamma(t-s)\dot{q}(s)ds + R(t) \tag{4.51}$$

admits the Gibbs measure (4.40) as the invariant measure.

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

The first claim follows from the discussion above. One can construct a heat bath so that the coupling is given by (4.41) and (4.48).

For the second claim, one must show that the Hamiltonian system is ergodic. According to [32], the special coupling (4.41) and (4.48) guarantees that the joint distribution converges weakly to the joint Gibbs measure (4.39). For an arbitrary coupling between q_0 and $R(t)$, we must check what happens for a given q_0. Conditioning on q_0, the random noise according to (4.48) has the same distribution with $R(t)$, and thus the conditional distribution for $\mu_A^t(\cdot|q_0) = P((q(t), u(t)) \in \cdot \mid q_0)$ is the same for any coupling. $\mu_A^t(\cdot|q_0)$ converges weakly to the Gibbs measure μ_A. If μ_A^0 is absolutely continuous with respect to the Gibbs measure, one can then have uniform estimates on the joint distribution.

Note that our final goal would be to consider the case $\gamma(t) \propto |t|^{-\alpha}$ and $m \to 0$. For kernel $|t|^{-\alpha}$, $\rho \notin L^2(\mathbb{R})$. One can therefore mollify γ by

$$\gamma_\epsilon(t) = \eta_\epsilon * \gamma(t), \tag{4.52}$$

so that the smoothness gives decay in Fourier side and we have $\rho_\epsilon \in L^2(\mathbb{R})$.

Remark 3. *In this sense, what matters is the smoothness of $\gamma(\cdot)$ instead of its tail behavior as $t \to \infty$. Even if the kernel γ is not integrable but as long as it is smooth, the above construction works.*

Formally, if final equilibrium is preserved with $\epsilon \to 0$ limit, then the Gibbs measure is the equilibrium measure for the GLE with kernel $|t|^{-\alpha}$. Then, formally, the $m \to 0$ limit yields that the Gibbs measure proportional to $\exp(-V(q))$ is the final equilibrium measure of our FSDE (1.18). Regarding the limit $m \to 0$, one should be careful. The limit equation for a general kernel γ may not be a good initial value problem. The initial value problem

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

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 $m \to 0$ limit can pass to the final equilibrium measures.

Acknowledgements

The work of J.-G. Liu is partially supported by KI-Net NSF RNMS grant No. 1107291 and NSF DMS grant No. 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.

References

[1] Harry Nyquist. Thermal agitation of electric charge in conductors. *Phys. Rev.*, 32(1):110, 1928.

[2] Herbert B Callen and Theodore A Welton. Irreversibility and generalized noise. *Phys. Rev.*, 83(1):34, 1951.

[3] U. Marconi, A. Puglisi, L. Rondoni, and A. Vulpiani. Fluctuation–dissipation: response theory in statistical physics. *Phys. Rep.*, 461(4):111–195, 2008.

[4] Hazime Mori. A continued-fraction representation of the time-correlation functions. *Prog. Theor. Phys.*, 34(3):399–416, 1965.

[5] R. Kubo. The fluctuation-dissipation theorem. *Rep. Prog. Phys.*, 29(1):255, 1966.

[6] B.U. Felderhof. On the derivation of the fluctuation-dissipation theorem. *J. Phys. A-Math. Gen.*, 11(5):921, 1978.

[7] G.W. Ford, M. Kac, and P. Mazur. Statistical mechanics of assemblies of coupled oscillators. *J. Math. Phys.*, 6(4):504–515, 1965.

[8] R. Zwanzig. Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9(3):215–220, 1973.

[9] D. Givon, R. Kupferman, and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55, 2004.

[10] Samuel C Kou. Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *Ann. Appl. Stat.*, pages 501–535, 2008.

[11] Grigorios A Pavliotis. *Stochastic processes and applications, Diffusion Processes, the Fokker-Planck and Langevin equations*. Springer, 2014.

[12] Martin Hairer and Andrew J Majda. A simple framework to justify linear response theory. *Nonlinearity*, 23(4):909, 2010.

[13] Weihua Deng and Eli Barkai. Ergodic properties of fractional Brownian-Langevin motion. *Phys. Rev. E*, 79(1):011112, 2009.

[14] Benoit B Mandelbrot and John W Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10(4):422–437, 1968.

[15] D. Nualart. Fractional Brownian motion: stochastic calculus and applications. In *International Congress of Mathematicians*, volume 3, pages 1541–1562, 2006.

[16] M. Hairer. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.*, pages 703–758, 2005.

[17] L. Decreusefond and Ustunel A.S. Stochastic analysis of the fractional Brownian motion. *Potential Anal.*, 10(2):177–214, 1999.

[18] D. Nualart and Y. Ouknine. Regularization of differential equations by fractional noise. *Stoch. Proc. Appl.*, 102(1):103–116, 2002.

[19] V. Pipiras and M.S. Taqqu. Are classes of deterministic integrands for fractional Brownian motion on an interval complete? *Bernoulli*, 7(6):873–897, 2001.

[20] L. Li and J.-G. Liu. On convolution groups of completely monotone sequences/functions and fractional calculus. preprint.

[21] Martin Hairer and Natesh S Pillai. Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. In *Annales de l'institut Henri Poincaré (B)*, volume 47, pages 601–628, 2011.

[22] R Sakthivel, P Revathi, and Y. Ren. Existence of solutions for nonlinear fractional stochastic differential equations. *Nonlinear Anal.-Theor.*, 81:70–86, 2013.

[23] Martine Zähle. Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Relat. Fields*, 111(3):333–374, 1998.

[24] T. Mikosch and R. Norvaiša. Stochastic integral equations without probability. *Bernoulli*, pages 401–434, 2000.

[25] Tyrone E Duncan, Yaozhong Hu, and Bozenna Pasik-Duncan. Stochastic calculus for fractional Brownian motion I. Theory. *SIAM J. Control Optim.*, 38(2):582–612, 2000.

[26] Francesca Biagini, Yaozhong Hu, Bernt Øksendal, and Tusheng Zhang. *Stochastic calculus for fractional Brownian motion and applications*. Springer Science & Business Media, 2008.

[27] V. Pipiras and M.S. Taqqu. Integration questions related to fractional Brownian motion. *Probab. Theory Relat. Fields*, 118(2):251–291, 2000.

[28] Jean Mémin, Yulia Mishura, and Esko Valkeila. Inequalities for the moments of Wiener integrals with respect to a fractional Brownian motion. *Stat. Probabil. Lett.*, 51(2):197–206, 2001.

[29] F. Mainardi, P. Paradisi, and R. Gorenflo. Probability distributions generated by fractional diffusion equations. *arXiv preprint arXiv:0704.0320*, 2007.

[30] D.V. Widder. *The Laplace Transform*. Princeton University Press, 1941.

[31] M. Ottobre and G.A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629, 2011.

[32] V. Jakšić and C.A. Pillet. Ergodic properties of the non-Markovian Langevin equation. *Lett. Math. Phys.*, 41(1):49–57, 1997.

[33] L. Rey-Bellet and L. E. Thomas. Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. *Commun. Math. Phys.*, 225(2):305–329, 2002.

[34] M. Freidlin. Some remarks on the Smoluchowski–Kramers approximation. *J. Stat. Phys.*, 117(3-4):617–634, 2004.