ESAIM: M2AN 54 (2020) 431–463
https://doi.org/10.1051/m2an/2019067

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

NUMERICAL APPROXIMATION AND FAST EVALUATION OF THE OVERDAMPED GENERALIZED LANGEVIN EQUATION WITH FRACTIONAL NOISE

Di Fang[1,*] and Lei Li[2]

Abstract. The generalized Langevin equation (GLE) is a stochastic integro-differential equation that has been used to describe the movement of microparticles with sub-diffusion phenomenon. It has been proved that with fractional Gaussian noise (fGn) mostly considered by biologists, the overdamped Generalized Langevin equation satisfying fluctuation dissipation theorem can be written as a fractional stochastic differential equation (FSDE). In this work, we present both a direct and a fast algorithm respectively for this FSDE model in order to numerically study ergodicity. The strong orders of convergence are proven for both schemes, where the role of the memory effects can be clearly observed. We verify the convergence theorems using linear forces, and then verify the convergence to Gibbs measure algebraically for the double well potentials in both 1D and 2D setups. Our work is new in numerical analysis of FSDEs and provides a useful tool for studying ergodicity. The idea can also be used for other stochastic models involving memory.

Mathematics Subject Classification. 65C20, 60H35.

Received January 10, 2019. Accepted August 31, 2019.

1. Introduction

Diffusion in statistical mechanics is a class of ubiquitous phenomena that appears commonly in nature and has been extensively studied from both physical and mathematical points of view. While the normal diffusion, typically formalized by random walk, is well understood through the rather classical Brownian Motion theory and the Langevin equation (LE), the so-called *anomalous diffusion* processes, however, remain far from fully explored. Among them, subdiffusion, in which the mean square displacement $\langle \Delta x(t)^2 \rangle$ scales as t^β with $0 < \beta < 1$, has been found in many different physical contexts such as cytoplasmic macromolecules in living cells [54, 58], the movement of lipids and single-molecule on membranes [49, 50, 57], the solute transport in porous media [10], the translocation of polymer solutions [20, 40, 41], and the conformational dynamics and

Keywords and phrases. Generalized Langevin equation, fractional Brownian motion, fractional stochastic differential equations, fast algorithm, strong convergence.

[1] Department of Mathematics, University of California, Berkeley, Berkeley, CA 94720, USA.

[2] School of Mathematical Sciences, Institute of Natural Sciences, MOE-LSC, Shanghai Jiao Tong University, 200240 Shanghai, P.R. China.

*Corresponding author: difang@berkeley.edu, di@math.wisc.edu

fluctuations of protein molecules [35,60]. To better describe subdiffusive phenomena, the Generalized Langevin equation (GLE) with fractional Gaussian noise (fGn) is therefore introduced [23] from a model point of view. This particular GLE is given by

$$m\ddot{x}(t) = -\nabla V(x) - \int_{t_0}^{t} \gamma(t-s)\dot{x}(s)\,\mathrm{d}s + \eta(t), \tag{1.1}$$

where a particle with mass m and position $x(t)$ is considered. Besides the external force $-\nabla V(x) = b(x)$, the particle is also driven by the friction (dissipation) and a random force (fluctuation), in which the friction term depending on the history of the particle velocity – instead of the instantaneous velocity – has a memory effect with the memory kernel $\gamma(t)$, and the random fluctuation force $\eta(t) = \sigma\dot{B}_H$ is the fractional Gaussian noise with the Hurst index H and some fluctuation constant σ. Here, $\dot{B}_H$ is understood as the distributional time derivative of the fractional Brownian motion B_H. See Section 2.2 for the brief introduction and we refer the readers to [11, 52] for more details. This equation is sometimes also referred to as the fractional Langevin equation (FLE) in physical literature [4,5,22,53].

The intuitive picture of the GLE (or LE) description, as is provided by Kubo [25], is to consider for example the colloidal particles floating in a liquid medium. The random impacts of surrounding particles are responsible for two effects – the random force and the systematic friction, and hence the two parts must be related. To be specific, the energy restored by fluctuation must be balanced with the energy loss by dissipation so that the particle achieves the equilibrium with the correct temperature. This internal relationship linking both parts of the microscopic forces is described in general as the fluctuation-dissipation theorem (FDT) [6, 25, 38]. To be specific, $\eta(t)$ and the kernel $\gamma(t)$ satisfy

$$\mathbb{E}(\eta(t)\eta(t+\tau)) = kT\gamma(|\tau|),\ \forall\tau \in \mathbb{R}, \tag{1.2}$$

where $\mathbb{E}$ means "ensemble average" in physics or expectation in mathematics.

We point out that the GLE is, of course, by no means introduced merely for the sake of subdiffusion. Proposed by Mori [42] and Kubo [25] in the sixties, the GLE is a well-appreciated object enjoying a long history. From a model viewpoint, GLE appears naturally according to the fluctuation-dissipation theorem as a generalization of the Langevin equation when the random force in considerations is no longer memoryless; In terms of derivation, it can be derived from the Mori–Zwanzig formulation [43,63] as a powerful tool for dimension reduction in many different forms depending on applications, such as molecular dynamics [7, 16, 30, 31] and recently uncertainty quantification [56]. An ergodicity condition of the equation has been recently studied in [26]. For different random forces, their corresponding memory kernels are different as a consequence of the fluctuation dissipation theorem (1.2). In the subdiffusion model with fractional Gaussian noise, the reasonable memory kernel turns out to behave as a power law $\gamma(t) \propto t^{-\alpha}$ with certain constant α and the friction term becomes the fractional derivative, as is shown both formally and rigorously that in the contexts of the absence of external force and with quadratic potential in the *overdamped* regime ($m \ll 1$) [12, 23, 25, 32], respectively. It worths pointing out that due to the complicated memory effects, a rigorous proof of the convergence of the law to Gibbs measure in the overdamped regime is by no means an easy task. In above cases (*i.e.* absence of external force or with quadratic potential), the solution can be formally written down explicitly, which makes the proofs feasible. For general potentials, however, a validation analysis of the ergodicity in equation (1.1) remains unclear. Therefore, trying to understand this problem numerically serves as one main motivation of this work.

Following the analytical work, we also restrict ourselves to the overdamped GLE ($m \to 0$) as in [32]. The corresponding fractional stochastic differential equation (FSDE) reads [32]

$$D_c^{\alpha} x = -\nabla V(x) + \sigma\dot{B}_H, \tag{1.3}$$

which is also known as the overdamped FLE with fractional noise, where

$$\alpha = 2 - 2H, \quad \sigma = \frac{\sqrt{2}}{\sqrt{\Gamma(2H+1)}},$$

as a result of FDT (1.2) (see Sect. 2.2 for more details). The fractional derivative in the Caputo sense [14,21] is given by

$$D_c^\alpha x(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\dot{x}(s)}{(t-s)^\alpha} \, \mathrm{d}s. \tag{1.4}$$

Though motivated by the understanding of overdamped GLE in this FSDE model, we point out that the numerical analysis of FSDE is by itself an interesting mathematical problem. Although there has been some numerical simulations of the FLE [5, 15], the numerical analysis remain untouched. We remark that there have been many recent developments in the numerical schemes of the time fractional differential equations, for example, the L^1 schemes [29,33,44], spectral methods [37,51] and higher-order algorithms [8,9,61].

The main difficulties of our problem are of two folds: (a) In terms of numerical analysis, different from the usual SDE where the correlation between increments of the standard Wiener process is simply absent, in FSDE the increments of fBm depends on the history, resulting in the analysis of strong convergence much harder comparing to usual case. (b) From a computational viewpoint, due to the memory effect, a straightforward discretization will be very memory-consuming, as one needs access to all history values at each time step. This becomes particularly troublesome when computing a number of sample paths. We summarize here the contributions of this work:

(1) We propose a fast algorithm that reduces the computational cost of a single path from $O(N^2)$ to $O(NM)$, where N is the number of steps in time discretization, and $M \ll N$ is the number of modes used in the approximation by exponential sums. This makes the numerical study of multiple sample paths and ergodicity feasible.
(2) We provide rigorous strong convergence theorems of both direct and fast solvers to the FSDE (Thms. 3.2 and 4.4). The work is new in numerical analysis of FSDEs, and the idea can be used for other stochastic models involving memory.
(3) Numerical studies of ergodicity are conducted with harmonic and double-well potentials in 1D and 2D cases, which shows the versatility of our algorithm for general potentials.

The rest of the paper is organized as follows. To start with, we give a brief review of fBm and notations in Section 2. To help the understanding of the model, some basic estimates of FSDE are also provided in preparation of later numerical analysis. In Section 3 we propose a direct numerical scheme for the FSDE with general parameters

$$\alpha \in (0,1), \ H \in (1/2,1), \quad \sigma > 0, \tag{1.5}$$

based on the integral formulation of the FSDE and prove its convergence result in the strong sense. The optimal strong order for the overdamped GLE in the case $b(\cdot)$ is linear has been obtained by careful estimates of correlation for increments of fBm. However, due to the memory effects, the computational cost is rather formidable. In particular, a speeding up of the solver becomes crucial in the considerations of the ergodicity, and moreover the stochastic nature of our problem where multiple sample paths need to be computed. To tackle the task, a fast numerical algorithm is then introduced in Section 4. The idea is to make use of the approximation by exponential sums (also known as the sum-of-exponentials approximation), introduced by Beylkin and Monzón in [2,3] to re-represent the algebraic memory kernel, which can be understood intuitively here as Markovian approximations of the non-Markovian process. We point out that such approximation has been applied in various situations for improving the computational efficiency of convolution integrals, for instance [1,17–19,27,59,62]. The convergence of the fast algorithm is proved by establishing a stability lemma (Lem. 4.3) based on comparison principles. The sampling strategy of fBm and the complexity of both algorithms are summarized in Section 5. Finally Section 6 is devoted to the numerical results: we carry out numerical tests of linear forces to verify the convergence and efficiency of the algorithms – whose strong orders agree with the theorems proven in previous sections – and provide extensive numerical studies for the ergodicity of the double-well potentials in both 1D

and 2D cases, which illustrates the potential of our new algorithm in studying the ergodicity of the system in general.

2. FSDE and fractional Brownian motion

2.1. Preliminaries and notations

In this paper, we will fix the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. x_0 is a random variable defined on this space, while B_H is a fractional Brownian motion defined on this space (see Sect. 2.2 for brief introduction). We will use the filtration $(\mathcal{G}_t)$ with

$$\mathcal{G}_t = \cap_{s>t}\sigma\Big(B_H(\tau), 0 \leqslant \tau \leqslant s, x_0\Big), \quad \forall t \in [0, T).$$

The notation $\mathbb{E}$ represents the expectation (integral) under probability measure $\mathbb{P}$.

The FSDE model (1.3) with general parameters given in (1.5) is rigorously defined through the following integral formulation

$$x(t) = x_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1} b(x(s))\,\mathrm{d}s + \frac{\sigma}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}B_H, \tag{2.1}$$

where $b(x) = -\nabla V$. In the discussion below, we will consider a general force field $b(x)$ that is not necessarily conservative. As is shown in Theorem 1 of [32], if b is Lipschitz, the FSDE has a unique strong solution $x(t)$, which is a stochastic process adapted to the filtration $(\mathcal{G}_t)$.

For the convenience of discussion, we introduce the norm of a random variable $v \in L^2(\Omega; \mathbb{P})$

$$\|v\| = \sqrt{\mathbb{E}|v|^2}, \tag{2.2}$$

together with the associated inner product

$$\langle u, v\rangle = \mathbb{E}uv. \tag{2.3}$$

Occasionally, we will drop the measure $\mathbb{P}$ and use $L^2(\Omega)$ to mean the space of square integrable random variables. Moreover, we note that the solution to the fractional ODE [28]

$$D_c^\alpha u = \lambda u, \quad u(0) = A, \tag{2.4}$$

where λ and A are constants, is given by

$$u(t) = AE_\alpha(\lambda t^\alpha), \tag{2.5}$$

where $E_\alpha(\cdot)$ is the Mittag-Leffler function defined by

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1+n\alpha)}. \tag{2.6}$$

In the following subsections, we shall briefly revisit the basics of fractional Brownian motion, and then prove some basic estimates for the FSDE that prepares us for the numerical analysis in later sections.

2.2. Fractional Brownian motion

The fractional Brownian motion B_H (see [36, 45] for more detailed discussions) with Hurst parameter $H \in (0, 1)$ is a Gaussian process defined on some probability space $(\Omega, \mathcal{F}, P)$ such that $B_H(0) = 0$, with mean zero and covariance

$$\mathbb{E}(B_t^H B_s^H) = R_H(s,t) = \frac{1}{2}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right). \tag{2.7}$$

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). With this definition and the fact 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 $\dot{B}_H(t)$, one has for $\tau > 0$

$$\begin{aligned}\mathbb{E}(\dot{B}_H(t)\dot{B}_H(\tau+t)) &= \lim_{h\to 0, h_1\to 0} \mathbb{E}\left(\frac{B_H(t+h_1)-B_H(t)}{h_1}\frac{B_H(t+\tau+h)-B_h(t+\tau)}{h}\right)\\ &= \lim_{h\to 0, h_1\to 0} \frac{1}{2hh_1}\left((\tau+h_1)^{2H} - (\tau+h_1-h)^{2H} - \tau^{2H} + (\tau-h)^{2H}\right)\\ &= H(2H-1)\tau^{2H-2}. \end{aligned} \tag{2.8}$$

For the complete auto-correlation function of $\dot{B}_H$, we assume there is no extra singularity like Dirac delta function or its derivatives at $\tau = 0$ and check for $s < t$:

$$\begin{aligned}\mathbb{E}(B_t^H B_s^H) &= \int_0^t \int_0^s \mathbb{E}\left(\dot{B}_H(z)\dot{B}_H(w)\right)\mathrm{d}z\,\mathrm{d}w\\ &= \int_0^s \int_0^t H(2H-1)|z-w|^{2H-2}\,\mathrm{d}z\,\mathrm{d}w = \frac{1}{2}(s^{2H}+t^{2H}-(t-s)^{2H}).\end{aligned} \tag{2.9}$$

Now that (2.9) agrees with (2.7). Hence, above assumption is reasonable. The auto-correlation function of $\dot{B}_H$ is therefore $H(2H-1)|\tau|^{2H-2}$, and this explains why the fractional noise leads to power law kernel by the (1.2) and why we have the FSDE in the over-damped limit as mentioned in the introduction. Note that When $H < 1/2$, the auto-correlation function is no longer integrable and could contain singular distributions supported at $\tau = 0$. In light of this, the FSDE in the overdamped regime might contain a mixture of fractional and integer derivatives, which is another interesting equation type but beyond the scope of this paper.

For the convenience, we denote

$$G(t) = \frac{\sigma}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}B_H. \tag{2.10}$$

The process $G(t)$ is clearly a Gaussian process because $B_H(t)$ is a Gaussian process. In [32], it has been shown that

Lemma 2.1 ([32]). *The increments of G satisfies*

$$\mathbb{E}|G(t_2)-G(t_1)|^2 \leqslant C|t_2-t_1|^{2H+2\alpha-2}. \tag{2.11}$$

Consequently, $G(t)$ is $H+\alpha-1-\epsilon$ Hölder continuous for any $\epsilon > 0$. Moreover, if $\alpha = 2-2H$ and $\sigma = \frac{\sqrt{2}}{\sqrt{\Gamma(2H+1)}}$, then

$$G(t) \overset{d}{=} \beta_H B_{1-H} \tag{2.12}$$

is a fractional Brownian motion up to a factor. Here, β_H is a constant given by

$$\beta_H = \frac{\sqrt{2}}{\sqrt{\Gamma(3-2H)}}. \tag{2.13}$$

2.3. Some estimates of the FSDE

In this section, we prove some basic estimates of the FSDE, which helps the understanding of the equation and prepares us for the numerical analysis in the later sections.

For the FSDE (2.1), assume that

$$|b(x) - b(y)| \leqslant L|x - y|. \tag{2.14}$$

In [32], it has been shown that (2.1) has a unique continuous strong solution. Moreover, we have the following moment control:

Lemma 2.2. *If* $\mathbb{E}(|x_0|^2 + |b(x_0)|^2) < \infty$, *then*

$$\sup_{0\leqslant t\leqslant T} \|x(t)\|^2 \leqslant C(T). \tag{2.15}$$

Proof. Using (2.14), the strong solution $x(t)$ satisfies

$$|x(t) - x_0| \leqslant \frac{t^\alpha}{\Gamma(\alpha+1)}|b(x_0)| + \frac{L}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}|x - x_0|(s)\,\mathrm{d}s + |G(t)|.$$

Taking square and using the elementary inequality $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$, we have

$$|x(t) - x_0|^2 \leqslant 3\left[\frac{t^{2\alpha}}{\Gamma(\alpha+1)^2}|b(x_0)|^2 + \frac{L^2}{\Gamma(\alpha)^2}\left(\int_0^t (t-s)^{\alpha-1}|x - x_0|(s)\,\mathrm{d}s\right)^2 + |G(t)|^2\right]. \tag{2.16}$$

For the second term, Hölder inequality yields

$$\begin{aligned}\left(\int_0^t (t-s)^{\alpha-1}|x - x_0|(s)\,\mathrm{d}s\right)^2 &\leqslant \int_0^t (t-s)^{\alpha-1}\,\mathrm{d}s \int_0^t (t-s)^{\alpha-1}|x - x_0|^2(s)\,\mathrm{d}s \\ &= \frac{t^\alpha}{\alpha}\int_0^t (t-s)^{\alpha-1}|x - x_0|^2(s)\,\mathrm{d}s.\end{aligned} \tag{2.17}$$

Further, by the result in Proposition 1 of [32]

$$\mathbb{E}|G(t)|^2 \leqslant C(T). \tag{2.18}$$

Combining (2.16)–(2.18), we have

$$\mathbb{E}|x(t) - x_0|^2 \leqslant C_1(T) + C_2(T)\frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\mathbb{E}|x - x_0|^2(s)\,\mathrm{d}s. \tag{2.19}$$

Using the Grönwall inequality in Proposition 5 of [13], we have

$$\mathbb{E}|x(t) - x_0|^2 \leqslant C(T), \forall t \in [0, T]. \tag{2.20}$$

The claim therefore follows. □

Now we are ready to estimate the increments of the solution.

Lemma 2.3. *There exists a constant* $C(T)$ *such that for all* $\delta \in (0,1)$, *we have*

$$\|x(t+\delta) - x(t)\|^2 \leqslant C(T)\delta^{2H+2\alpha-2} \tag{2.21}$$

for all $t \leqslant T, t+\delta \leqslant T$.

Proof. By (2.1), we have

$$x(t+\delta)-x(t)=\frac{1}{\Gamma(\alpha)}\int_0^t\left[(t+\delta-s)^{\alpha-1}-(t-s)^{\alpha-1}\right]b(x(s))\,\mathrm{d}s + \frac{1}{\Gamma(\alpha)}\int_t^{t+\delta}(t+\delta-s)^{\alpha-1}b(x(s))\,\mathrm{d}s+(G(t+\delta)-G(t))=:I_1+I_2+I_3.$$

Using again $(a+b+c)^2\leqslant 3(a^2+b^2+c^2)$, we have

$$\|x(t+\delta)-x(t)\|^2\leqslant 3(\|I_1\|^2+\|I_2\|^2+\|I_3\|^2).$$

Using the Hölder inequality,

$$\mathbb{E}\left(\int_0^t[(t+\delta-s)^{\alpha-1}-(t-s)^{\alpha-1}]b(x(s))\,\mathrm{d}s\right)^2\leqslant\int_0^t\left[(t-s)^{\alpha-1}-(t+\delta-s)^{\alpha-1}\right]\mathrm{d}s \times\mathbb{E}\int_0^t\left[(t-s)^{\alpha-1}-(t+\delta-s)^{\alpha-1}\right]|b(x(s))|^2\,\mathrm{d}s.$$

Since

$$0\leqslant\int_0^t(t-s)^{\alpha-1}-(t+\delta-s)^{\alpha-1}\,\mathrm{d}s=\frac{1}{\alpha}(t^\alpha+\delta^\alpha-(t+\delta)^\alpha)\leqslant\frac{1}{\alpha}\delta^\alpha,$$

we have

$$\|I_1\|^2\leqslant C(\alpha)\delta^{2\alpha}\sup_{0\leqslant s\leqslant T}\mathbb{E}|b(x(s))|^2.$$

Similarly, one can apply Hölder inequality for $\mathbb{E}(\int_t^{t+\delta}(t+\delta-s)^{\alpha-1}b(x(s))\,\mathrm{d}s)^2$. Finally, we have

$$\|x(t+\delta)-x(t)\|^2\leqslant C\delta^{2\alpha}\sup_{0\leqslant s\leqslant T}\mathbb{E}|b(x(s))|^2+\mathbb{E}|G(t+\delta)-G(t)|^2\leqslant C(T)\left(\delta^{2\alpha}+\delta^{2H+2\alpha-2}\right),\tag{2.22}$$

where we used Lemmas 2.1 and 2.2. Apparently $2\alpha>2H+2\alpha-2$, and the claim follows. □

Remark 2.4. When FDT is satisfied, $\alpha=2-2H$, the order in the right hand side of equation (2.22) becomes

$$\|x(t+\delta)-x(t)\|^2\leqslant C\delta^{2-2H}.$$

If $H=\frac{1}{2}$ and $B_H=W$ is the standard Wiener process, this is a well-known result for diffusion processes.

3. Direct discretization

3.1. A direct scheme

For the fixed terminal time T, we define $t_j=jk$, where $k=T/N$ denotes the time step and N is a positive integer. We approximate $b(x(t))$ with a function $\tilde{b}(t)$ such that

$$\tilde{b}(t)=b(x_{j-1}),\ t\in[t_{j-1},t_j).\tag{3.1}$$

This then gives a numerical scheme for the FSDE (2.1):

$$\begin{aligned}x_n&=x_0+\frac{1}{\Gamma(\alpha)}\sum_{j=1}^n b(x_{j-1})\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\,\mathrm{d}s+G(t_n)\\&=x_0+\frac{k^\alpha}{\Gamma(1+\alpha)}\sum_{j=1}^n b(x_{j-1})((n-j+1)^\alpha-(n-j)^\alpha)+G(t_n).\end{aligned}\tag{3.2}$$

We point out that in the memoryless case $H=1/2$, this scheme is in accordance with the Euler–Maruyama scheme for SDEs.

3.2. Strong convergence of the direct solver

In this section, we present two analysis results on the strong convergence of the direct solver. First, we prove the convergence result in Proposition 3.1 for the general case, *i.e.* for general parameters H and α and multiplicative noise. However, since here we have an additive noise, one expect better convergence rate. We are able to improve the convergence rate in the physical case of $\alpha = 2 - 2H$ and the linear $b(x)$, which is the main theorem (Thm. 3.2) of the direct solver.

Let N_G be the complexity for sampling process $G(t)$, and the notation $\mathcal{C}$ for the complexity, or cost, of an algorithm. Then, a simple estimate gives the following claim regarding the numerical scheme (3.2):

Proposition 3.1. *The scheme (3.2) has time complexity $\mathcal{C} = O(N^2 + N_G)$ and it converges to the solution of the FSDE strongly in the following sense:*

$$\sup_{n \leqslant T/k} \|x_n - x(t_n)\| \leqslant C(T) k^{H+\alpha-1}. \tag{3.3}$$

Proof. To compute the the fractional integral, we need $O(N^2)$ operations. Hence, the complexity is clearly

$$\mathcal{C} = O(N^2 + N_G).$$

Using (2.1) and (3.2), we have

$$\begin{aligned} x_n - x(t_n) &= \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x_{j-1}) - b(x(s)))\, \mathrm{d}s \\ &= \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x_{j-1}) - b(x(t_{j-1})))\, \mathrm{d}s \\ &\quad + \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x(t_{j-1})) - b(x(s)))\, \mathrm{d}s. \end{aligned} \tag{3.4}$$

Denote

$$R_n := \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x(t_{j-1})) - b(x(s)))\, \mathrm{d}s. \tag{3.5}$$

It follows from (2.14) and Lemma 2.3 that for all $n \leqslant T/k$

$$\|R_n\| \leqslant C_1(T) k^{H+\alpha-1}. \tag{3.6}$$

Hence, for any $n \leqslant T/k$, we have

$$\|x_n - x(t_n)\| \leqslant \frac{L}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\, \mathrm{d}s \|x_{j-1} - x(t_{j-1})\| + C_1(T) k^{H+\alpha-1}.$$

Applying Lemma 6.1 of [13], we find that

$$\|x_n - x(t_n)\| \leqslant u(t_n) \leqslant C(T) k^{\alpha+H-1},$$

where u solves

$$D_c^{\alpha} u = Lu, \quad u(0) = C_1(T) k^{H+\alpha-1}.$$

This then finishes the proof. □

The rate in Proposition 3.1 is only optimal for multiplicative noise and we expect better bounds for the strong order since we have additive noise. As is well known, the Euler–Maruyama scheme for usual SDE has strong order $O(k)$ for additive noise, which can be proved using the fact that $W(t_2) - W(t_1)$ is independent of the sigma algebra $\sigma(W(s) : s \leqslant t_1)$. Unfortunately, for the fractional Brownian motion, $B_H(t_2) - B_H(t_1)$ is not independent of the history. However, we note that the correlation decays and we may use this fact to improve the strong order. In fact, we are able to improve the result for the case $b(x) = Bx$ and $\alpha = 2 - 2H$.

Theorem 3.2 (Strong convergence of the direct solver)**.** *Let $\alpha = 2 - 2H$ and $b(x) = Bx$ where B is a $d \times d$ constant matrix. The scheme (3.2) has time complexity $\mathcal{C} = O(N^2 + N_G)$ and the strong error of the scheme can be controlled as*

$$\sup_{n \leqslant T/k} \|x_n - x(t_n)\| \leqslant \begin{cases} C(T,H)k^{3-3H}, & H \in (3/4, 1) \\ C(T,H)\sqrt{|\ln k|}k^{3/4}, & H = 3/4, \\ C(T,H)k^{3/2-H}, & H \in (1/2, 3/4). \end{cases} \tag{3.7}$$

Remark 3.3. Note that when $H = 1/2$, Theorem 3.2 provides a strong error of $O(k)$, which agrees with the result in the case of the standard Brownian motion with additive noise.

In fact, the key part of the proof relies on a detailed estimate of R_n as defined in (3.5), which is stated in the following lemma. The proof of the lemma is rather lengthy and technical, which is given in the appendix without interrupting the flow of the paper presentation. The theorem follows straightforwardly from a combination of the lemma and a similar proof as in Proposition 3.1.

Lemma 3.4. *Under the assumption of Theorem 3.2, we have*

$$\|R_n\| \leqslant \begin{cases} C(T,H)k^{3-3H} & H \in (3/4, 1), \\ C(T,H)\sqrt{|\ln k|}k^{3-3H} & H = 3/4, \\ C(T,H)k^{3/2-H} & H \in (1/2, 3/4). \end{cases} \tag{3.8}$$

We find that proving the strong order for FSDE is much more difficult compared with the usual SDE (Itô equations). The reason is that the increments of the fBm are not independent due to the memory effect. The key point we use is the fact that the correlation between the increments decays if the distance between them grows. In fact, we have an explicit representation of the fractional Brownian motion, which is given by (see [36])

$$B_H(t) = C_1(H)\left(\int_0^t (t-s)^{H-1/2}\,\mathrm{d}W(s) + \int_{-\infty}^0 ((t-s)^{H-1/2} - (-s)^{H-1/2})\,\mathrm{d}W(s)\right), \tag{3.9}$$

where $C_1(H)$ is a constant. Define the filtration $\mathcal{A}(t)$ by

$$\mathcal{A}(t) = \cap_{s>t}\sigma(W(\tau), \tau \in (-\infty, s], x_0). \tag{3.10}$$

It is then clear that under this representation $\mathcal{G}_t \subset \mathcal{A}_t,\ t \geqslant 0$ and $x(t) \in \mathcal{A}(t)$. It is worth pointing out that we then have the following explicit formula regarding the decay of the correlation:

Lemma 3.5. *Let $a \leqslant b < c$ and $H \in (0,1)$. We have*

$$\sqrt{\mathbb{E}\Big(\mathbb{E}(B_H(c) - B_H(b)|\mathcal{A}(a))\Big)^2} \leqslant C(H)|H - \frac{1}{2}|((c-a)^H - (b-a)^H). \tag{3.11}$$

Proof. Using the representation (3.9), we find

$$\Xi := \mathbb{E}(B_H(c) - B_H(b)|\mathcal{A}(a)) = C_1(H)\int_{-\infty}^a ((c-s)^{H-1/2} - (b-s)^{H-1/2})\,\mathrm{d}W(s).$$

The result then follows from the simple calculation below:

$$\sqrt{\mathbb{E}(\Xi)^2} = C_1(H)\sqrt{\int_{-\infty}^{a}((c-s)^{H-1/2}-(b-s)^{H-1/2})^2\,\mathrm{d}s} = C_1(H)|H-\frac{1}{2}|\sqrt{\int_{-\infty}^{a}\left(\int_b^c (r-s)^{H-3/2}\,\mathrm{d}r\right)^2\,\mathrm{d}s}$$
$$\leqslant C_1(H)|H-\frac{1}{2}|\int_b^c \|(r-\cdot)^{H-3/2}\|_{L^2(-\infty,a)}\,\mathrm{d}r = \frac{C_1(H)|H-\frac{1}{2}|}{\sqrt{2-2H}}\int_b^c (r-a)^{H-1}\,\mathrm{d}r.$$

□

For the linear $b(x)$, one can obtain an explicit expression of the correlation function as shown in Step 2 of the proof of Lemma 3.4 (see Appendix A), namely the estimate of K_2. For general $b(x)$, we expect that the strong order can also be improved compared with Proposition 3.1 by making use of the decay of correlation. However, this seems difficult and we leave this for future.

4. A fast scheme

The cost of the direct discretization is apparently large, especially to compute many sample paths. In this section, we propose a fast scheme and then prove the convergence of the fast schemes. The idea of the fast scheme is to use the approximation of the memory kernel by exponential sums in [19, 39]. The proof of the convergence is based on a stability result using comparison principle (Lem. 4.3).

4.1. The fast scheme

We first introduce the exponential sum approximation that states the algebraic decay kernel $t^{\alpha-1}$ away from the singularity point zero can be approximated by a summation of exponential functions.

Lemma 4.1 ([19]). *For $\alpha \in (0,1)$, tolerance $\epsilon > 0$, truncation $\delta > 0$,and fixed $T > 0$, there exist positive numbers s_i, ω_i, with $1 \leqslant i \leqslant M$ such that*

$$\left| t^{\alpha-1} - \sum_{i=1}^{M} \omega_i \mathrm{e}^{-s_i t} \right| \leqslant \epsilon, \ \forall t \in [\delta, T],$$

where

$$M = O\left(\log\frac{1}{\epsilon}\left(\log\log\frac{1}{\epsilon} + \log\frac{T}{\delta}\right) + \log\frac{1}{\delta}\left(\log\log\frac{1}{\epsilon} + \log\frac{1}{\delta}\right)\right).$$

Recall that $k = T/N$ is the time step where N is a positive integer. If $\epsilon = k^\sigma$ and $\delta = k$, we have

$$M = O((\log N)^2).$$

In light of this, we then break the convolution kernel into two parts

$$t^{\alpha-1} = t^{\alpha-1}\chi(t \leqslant k) + t^{\alpha-1}\chi(t > k),$$

where the second part is away from the singularity point zero and hence the approximation (4.1) can be applied. We are then led to the kernel

$$\gamma(t) = \begin{cases} \frac{1}{\Gamma(\alpha)} t^{\alpha-1} & t \leqslant k, \\ \frac{1}{\Gamma(\alpha)} \sum_{i=1}^{M} \omega_i \mathrm{e}^{-s_i t} & t > k. \end{cases} \tag{4.1}$$

This kernel is discontinuous, integrable and nonnegative. Then, the approximation (3.1) gives the following scheme:

$$x_n = x_0 + \sum_{j=1}^{n} b(x_{j-1}) \int_{t_{j-1}}^{t_j} \gamma(t_n - s)\,\mathrm{d}s + G(t_n). \tag{4.2}$$

By denoting

$$\eta_i^n = \begin{cases} 0, & n = 1, \\ \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n-1} b(x_{j-1}) \int_{t_{j-1}}^{t_j} \mathrm{e}^{-s_i(t_n - s)} \, \mathrm{d}s, & n \geqslant 2. \end{cases} \tag{4.3}$$

and some direct computations, we obtain the fast scheme:

$$x_{n+1} = x_0 + \frac{k^\alpha}{\Gamma(1+\alpha)} b(x_n) + \sum_{i=1}^{M} \omega_i \eta_i^{n+1} + G(t_{n+1}), \tag{4.4}$$

where the sequence $\{\eta_i^n\}_{i=1}^M$ satisfies the recurrence formula

$$\eta_i^{n+1} = \mathrm{e}^{-s_i k} \eta_i^n + \frac{1}{s_i \Gamma(\alpha)} (\mathrm{e}^{-s_i k} - \mathrm{e}^{-2 s_i k}) b(x_{n-1}). \tag{4.5}$$

In other words, at each time step t_n instead of summing over all the historic values of x_j for all $j \leqslant n$, here by recording the history information in the sequence of $\{\eta_i^n\}_{i=1}^M$ ($M \ll N$), at each time step one only needs get access to the past two values of x, and perform a summation of M terms (of η_i^n), instead of n terms (of x_j), where n increases with respect to time till N. In fact, (4.5) can be viewed as a consequence of the well-known fact that dynamics with exponentially decaying memory kernels can be made Markovian. To be precise, let $y(t)$ satisfy the equation of the following type for some function F:

$$\mathscr{L}(y)(t) = F\left(y_0, \int_{t_0}^{t} \mathrm{e}^{-\lambda(t-s)} b(y(s)) \, \mathrm{d}s\right),$$

where $\mathscr{L} = \sum_{m=0}^{p} a_m(t) \frac{d^m}{\mathrm{d}t^m}$ is some differential operator (including identity). Then, one can introduce

$$\eta(t) = \int_{t_0}^{t} \mathrm{e}^{-\lambda(t-s)} b(y(s)) \, \mathrm{d}s,$$

so that the equations for $(y(t), \eta(t))$ become

$$\begin{aligned} \mathscr{L}(y)(t) &= F(y_0, \eta(t)), \\ \dot{\eta} &= -\lambda \eta + b(y(t)), \end{aligned}$$

which are Markovian.

With the fast algorithm, we only need

$$O(NM) = O(N(\log N)^2)$$

efforts to compute $\{\eta_i^n\}$ for a sample path. To make the computational efficiency more transparent, we list some intuitive comparisons of the size of N and $\log(N)$: for example, when $N = 1000$, $(\log N)^2 \approx 47$; when $N = 10\,000$, $(\log N)^2 \approx 85$. In fact, M used here is a number even smaller than $(\log N)^2$, which is illustrated more specifically in later numerical Section 6.1.2.

4.2. Strong convergence of the fast algorithm

Note that even if one has the strong convergence of the direct solver, the strong convergence result for the fast solver is unclear even though the integration kernels are close on $[0, T]$. The reason is that we need some stability result to ensure that the error in the kernels will not be amplified. In this section, we first make use of the nonnegativity of the kernel to establish the stability based on comparison principles. With the stability result, we then prove strong convergence theorem (Thm. 4.4), which is the main theorem of the fast solver.

The following fractional ODE, as we shall see soon, will provide an upper bound for the solution of the Volterra type integral equation with kernel (4.1). Since the solutions of FODEs are well studied (see *e.g.* [13]), we can therefore control the solution of the Volterra integral equation. The FODE reads:

$$D_c^\alpha v = \left(1 + \frac{\epsilon}{\Gamma(\alpha)} T^{1-\alpha}\right) f(v(t)),\ v(0) = y_0. \tag{4.6}$$

The solution exists on $[0, T_b)$ where either $T_b = \infty$ or $\lim_{t \to T_b^-} v(t) = +\infty$ by the result in [13].

We now investigate the Volterra type integral equation with kernel (4.1). The solution will provide an upper bound for the discrete sequence generated by our fast solver, which therefore ensures the stability of our scheme.

Lemma 4.2. *Let $y_0 \geqslant 0$ and $f(\cdot)$ is a non-negative locally Lipschitz function. Consider*

$$y(t) = y_0 + \int_0^t \gamma(t-s) f(y(s))\, \mathrm{d}s. \tag{4.7}$$

Then, there is a unique continuous solution $y(t)$ on $[0, T] \cap [0, T_b)$ with T_b being the blowup time for (4.6). On $[0, T] \cap [0, T_b)$, $y(t)$ is a non-decreasing continuous function, and satisfies

$$y(t) \leqslant v(t)$$

where $v(\cdot)$ solves (4.6). If $f(\cdot)$ is globally Lipschtiz, then $T_b = \infty$ and $y(t)$ exists on $[0, T]$.

The proof of this lemma is rather standard, which is provided in Appendix B for readers' convenience. The following lemma gives the stability of the fast algorithm:

Lemma 4.3. *Let $y_0 \geqslant 0$. Assume $f(v)$ is a non-negative, non-decreasing, Lipschitz function on $[0, \infty)$. Let $y(t)$ be the solution to (4.7) on $[0, T]$. For a given sequence $z = \{z_m\}$, define*

$$T_n(z, y_0) = y_0 + \sum_{j=1}^{n} f(z_{j-1}) \int_{t_{j-1}}^{t_j} \gamma(t_n - s)\, \mathrm{d}s. \tag{4.8}$$

(1) *Assume that $a = \{a_n\}$ solves the induction relation $a_n = T_n(a, y_0)$. Then, $\{a_n\}$ is non-negative, non-decreasing and $a_n \leqslant y(t_n)$. In particular, if $f(v) = Lv$, we have*

$$a_n \leqslant y_0 E_\alpha \left(L \left(1 + \frac{\epsilon}{\Gamma(\alpha)}\right) t_n^\alpha\right) \leqslant C(T, H, \alpha) y_0,\ \forall \epsilon < 1,$$

where $E_\alpha(\cdot)$ is the Mittag-Leffler function defined in (2.6).

(2) *If a non-negative sequence $c = \{c_n\}$ satisfies $c_n \leqslant T_n(c, y_0)$, then*

$$c_n \leqslant a_n \leqslant y(t_n).$$

Proof. For (1), the claims follow by induction.

Indeed, $a_0 = y_0 = y(0) \geqslant 0$. Then,

$$0 \leqslant a_1 = y_0 + \frac{f(y_0)}{\Gamma(\alpha)} \int_0^k (k-s)^{\alpha-1}\, \mathrm{d}s \leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^k (k-s)^{\alpha-1} f(y(s))\, \mathrm{d}s = y(t_1)$$

by the monotonicity of f and y. Assume that for $m \leqslant n-1, n \geqslant 2$, we have $0 \leqslant a_m \leqslant y(t_m)$. By the monotonicity of f and y, $a_n = T_n(a, y_0) \geqslant 0$ is trivial. Moreover,

$$a_n = T_n(a, y_0) \leqslant y_0 + \sum_{j=1}^{n} f(a_{j-1}) \int_{t_{j-1}}^{t_j} \gamma(t_n - s)\, \mathrm{d}s \leqslant y_0 + \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \gamma(t_n - s) f(y(s))\, \mathrm{d}s = y(t_n).$$

Now, we show the monotonicity of $\{a_n\}$. $a_1 \geqslant a_0 = y_0$ is clear. For $n \geqslant 2$, we have:

$$\frac{f(a_{n-1})}{\Gamma(\alpha)} \int_{t_{n-1}}^{t_n} (t_n - s)^{\alpha-1}\,\mathrm{d}s \geqslant \frac{f(a_{n-2})}{\Gamma(\alpha)} \int_{t_{n-1}}^{t_n} (t_n - s)^{\alpha-1}\,\mathrm{d}s = \frac{f(a_{n-2})}{\Gamma(\alpha)} \int_{t_{n-2}}^{t_{n-1}} (t_{n-1} - s)^{\alpha-1}\,\mathrm{d}s. \tag{4.9}$$

Similarly,

$$f(a_{j-1}) \int_{t_{j-1}}^{t_j} \gamma(t_n - s)\,\mathrm{d}s \geqslant \begin{cases} f(a_{j-2}) \int_{t_{j-2}}^{t_{j-1}} \gamma(t_n - s)\,\mathrm{d}s, & j \geqslant 2, \\ 0, & j = 1. \end{cases} \tag{4.10}$$

Relations (4.9) and (4.10) then give $a_n \geqslant a_{n-1}$.

If $f(v) = Lv$, we have by Lemma 4.2 that

$$a_n \leqslant y(t_n) \leqslant v(t_n) = y_0 E_\alpha \left(L \left(1 + \frac{\epsilon}{\Gamma(\alpha)} \right) t_n^\alpha \right).$$

Claim (2) is straightforward by induction. To see this, we note that if we have $c_m \leqslant a_m$ for all $m \leqslant n-1$, then

$$T_n(c, y_0) \leqslant T_n(a, y_0),$$

which is just $c_n \leqslant a_n$. □

Now, we have the convergence theorem of the fast algorithm:

Theorem 4.4 (Strong convergence of the fast solver)**.** *Consider the fast scheme* (4.4).

(1) *Assume the exponential sum approximation as in Lemma 4.1 is applied with tolerance $\varepsilon > 0$, then the strong error satisfies $\|x_n - x(t_n)\| \leqslant C(T)(k^{H+\alpha-1} + \epsilon)$. In particular, if we set the tolerance $\epsilon = k^{H+\alpha-1}$, then sampling a path needs time complexity $\mathcal{C} = O(N(\log N)^2 + N_G)$, and we have the following strong error bound*

$$\sup_{n \leqslant T/k} \|x_n - x(t_n)\| \leqslant C(T) k^{H+\alpha-1}. \tag{4.11}$$

(2) *In the case $\alpha = 2 - 2H$ and $b(x) = Bx$, if we choose, $\epsilon = k^{\min(3/2-H, 3-3H)}$, then the time complexity for sampling a path is $\mathcal{C} = O(N(\log N)^2 + N_G)$ and the strong error is controlled as*

$$\sup_{n \leqslant T/k} \|x_n - x(t_n)\| \leqslant \begin{cases} C(T,H) k^{3-3H}, & H \in (3/4, 1) \\ C(T,H) \sqrt{|\ln k|} k^{3/4}, & H = 3/4, \\ C(T,H) k^{3/2-H}, & H \in (1/2, 3/4). \end{cases} \tag{4.12}$$

Proof. The complexity part is easy and we omit. We now focus on the error estimates.

Let

$$r(t) := x(t) - \left(x_0 + \int_0^t \gamma(t-s) b(x(s))\,\mathrm{d}s + G(t) \right). \tag{4.13}$$

Then, we have $r(t) = 0$ for $t \in [0, k]$ and $|r(t)| \leqslant \epsilon \int_0^t |b(x(s))|\,\mathrm{d}s$, $t \in [k, T]$ by the exponential sum approximation. It follows that

$$\sup_{t \in [0,T]} \|r(t)\| \leqslant C_1(T)\epsilon$$

It then follows from the definition of $r(t)$ that

$$\begin{aligned} x(t_n) - x_n &= \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \gamma(t_n - s)(b(x(s)) - b(x_{j-1}))\,\mathrm{d}s + r(t_n) \\ &= \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \gamma(t_n - s)(b(x(t_{j-1})) - b(x_{j-1}))\,\mathrm{d}s - R_n + r_1(t_n), \end{aligned}$$

where R_n is defined as in (3.5) and

$$r_1(t_n) = r(t_n) + \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \left(\gamma(t_n - s) - \frac{1}{\Gamma(\alpha)} (t_n - s)^{\alpha-1} \right) (b(x(s)) - b(x(t_{j-1}))) \, \mathrm{d}s. \tag{4.14}$$

By the exponential sum approximation, we again have

$$\sup_{n \leqslant T/k} \|r_1(t_n)\| \leqslant C(T)\epsilon.$$

We define $E_n = \|x(t_n) - x_n\|$. We then have

$$E_n \leqslant \sum_{j=1}^{n} L E_{j-1} \int_{t_{j-1}}^{t_j} \gamma(t_n - s) \, \mathrm{d}s + \sup_{n \leqslant T/k} \|R_n\| + \sup_{t \in [0,T]} \|r_1(t)\|. \tag{4.15}$$

Applying Lemma 4.3 for

$$y_0 = \sup_{n \leqslant T/k} \|R_n\| + \sup_{t \in [0,T]} \|r_1(t)\|$$

and $f(v) = Lv$ we have for $\epsilon < 1$,

$$\sup_{n \leqslant T/k} E_n \leqslant C(T, H, \alpha)(\sup_{n \leqslant T/k} \|R_n\| + \epsilon).$$

The result follows by the estimates of R_n in the proof of Proposition 3.1 and Theorem 3.2. □

5. Sampling fractional Brownian motion and process G

To give a complete description of the numerical scheme, we must understand how to sample fractional Brownian motion and the process $G(t)$. Since fractional Brownian motion is a Gaussian process. Using the covariant matrices, one can transform the standard Brownian motion into the desired Gaussian process.

For fractional Brownian motions, making use the time homogeneity, one has a fast algorithm to sample fractional Brownian motion, that is, the circulant method (or Wood-Chan algorithm) (see *e.g.* [52]). The idea is to sample

$$\xi_n = B_H(t_n) - B_H(t_{n-1}).$$

By the self-similarity,

$$\xi_n \overset{\mathrm{d}}{=} k^H (B_H(n) - B_H(n-1)).$$

Letting $\zeta_n = B_H(n) - B_H(n-1)$, (ζ_n) forms a Gaussian sequence whose covariance matrix satisfies

$$\Sigma_{ij} = \mathrm{Cov}(\zeta_i, \zeta_j) = \rho_H(|i-j|) \tag{5.1}$$

for some function ρ_H. This structure allows us to construct a circulant matrix M of size $2(N-1) \times 2(N-1)$ such that $\Sigma = M(1:N, 1:N)$. Then, one is able to take the square root of M using fast Fourier transform (FFT). With the square root of M, it is straightforward to transform the standard multivariable normal variables to a Gaussian sequence with covariant matrix M. The complexity is $O(N \log N)$. The first N elements will be a sample for the sequence (ζ_n). For the details, one can refer to Section 6 of [52].

As stated in [32], the case $\alpha = 2 - 2H$ is the physical case. As mentioned above, if $\alpha = 2 - 2H$,

$$G(t) \overset{\mathrm{d}}{=} \beta_H B_{1-H}$$

is a fractional Brownian motion up to a factor, so $G(t)$ can be sampled directly using the method here. This is a key observation for the simulation of overdamped GLE with fractional noise. Hence we conclude

Theorem 5.1. *For the overdamped GLE with fractional noise (i.e. $\alpha = 2-2H$, $\sigma = \sqrt{2/\Gamma(2H+1)}$), we can sample $G(t)$ with complexity*

$$N_G = O(N \log N). \tag{5.2}$$

Consequently, the total complexity of the direct scheme (3.2) is

$$\mathcal{C} = O(N^2), \tag{5.3}$$

and the total complexity of our fast algorithm (4.4) is

$$\mathcal{C} = O(N(\log N)^2). \tag{5.4}$$

For general α, the FSDE model is not physical, but may be used in other situations. The covariance matrix of $G(t)$ has been given in [32], for which the above trick fails, so that we do not have fast algorithms for sampling $G(t)$. Another option is to consider the following equivalent form of the FSDE:

$$J_{1-\alpha}(x-x_0)(t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha}(x(s)-x_0)\,\mathrm{d}s = \int_0^t b(x(s))\,\mathrm{d}s + \sigma B_H(t). \tag{5.5}$$

This is like integrate the differential form formally. It is known that a discretization of $J_{1-\alpha}(x-x_0)(t_n) - J_{1-\alpha}(x-x_0)(t_{n-1})$ is the L^1 scheme [33,44]. Hence, a possible numerical scheme is

$$(\mathcal{D}^\alpha x)_n = b(x_{n-1}) + \frac{\sigma}{k}\xi_n, \tag{5.6}$$

where $\mathcal{D}^\alpha$ refers to the L^1 scheme in [33]. This numerical scheme is like a Euler scheme for the differential form (1.3). For this scheme, though we can sample ξ_n fast, we do not have a fast algorithm for the L^1 scheme. Moreover, proving the convergence of this scheme is challenging. Hence, developing a fast algorithm for the general α case is left for future.

6. Numerical study of the overdamped GLE

As we have mentioned, the overdamped GLE with fractional noise is equation (2.1) with

$$\alpha = 2-2H, \quad \sigma = \frac{\sqrt{2}}{\sqrt{\Gamma(2H+1)}}.$$

We aim to study the ergodicity of the overdamped GLE

$$D^{2-2H}x = -\nabla V(x) + \sqrt{\frac{2}{\Gamma(2H+1)}}\dot{B}_H.$$

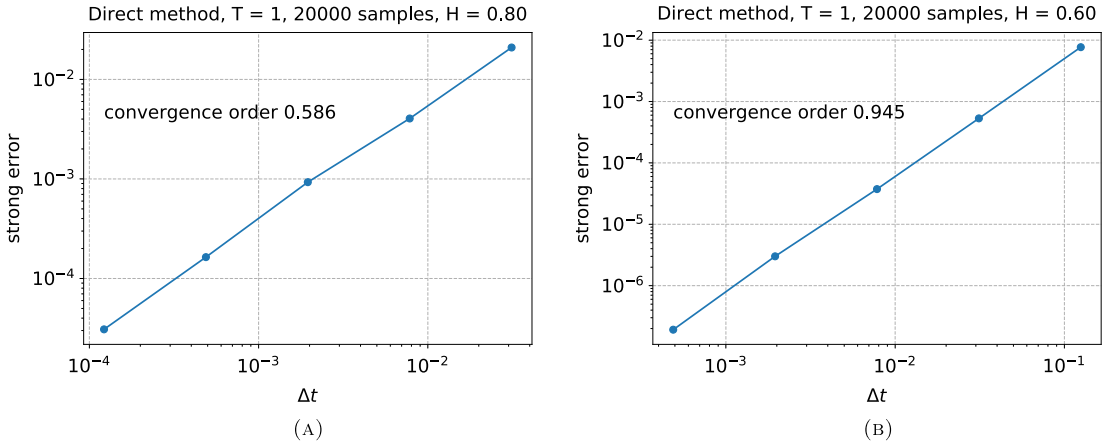
6.1. Example 1 (harmonic potential)

In this subsection, we aim to validate the order of strong convergence and study numerically the weak convergence of the schemes. The example we use is in the 1D case and

$$\nabla V(x) = x.$$

In [32], the formula of the exact solution in terms of $G(t)$ is given:

$$x(t) = x_0 e_{\alpha,1}(t) + G(t) + \int_0^t G(t-s)\dot{e}_{\alpha,1}(s)\,\mathrm{d}s. \tag{6.1}$$



(A) (B)

FIGURE 1. Example 1 (Sect. 6.1.1): The strong convergence of the direct method for $H = 0.8$ and $H = 0.6$, respectively in terms of various Δt in log–log scale. The convergence orders match the theoretical result $\min\{\frac{3}{2} - H, 3 - 3H\}$, as is proved in Theorem 3.2. (A) $H = 0.8$ and (B) $H = 0.6$.

Given a sample of $G(t)$, the integral here can be evaluated numerically with small time steps:

$$\int_0^t G(t-s)\dot{e}_{\alpha,1}(s)\,\mathrm{d}s \approx \sum_i G(t-t_i)(e_{\alpha,1}(t_i) - e_{\alpha,1}(t_{i-1}))$$

as the reference solution. For strong convergence, we must use the same sample of G, hence, we obtain ξ_n^m for the small step k_m. Then, for $k_{m-1} = 2k_m$,

$$\xi_n^{m-1} = \xi_{2n-1}^m + \xi_{2n}^m.$$

(k_m should be small so that the error from numerical integral is much smaller than the error from the scheme.)

6.1.1. The strong convergence of the direct solver

We first test the strong convergence of the direct solver. The time step for the reference solution is chosen as $k_m = 2^{-14}$, and $x_0 = 1$. The strong order of convergence is tested for various $H = 0.5, 0.55, 0.6, 0.65, 0.7, 0.73, 0.75, 0.8, 0.84$ over 20 000 sample paths. Note that $H = 0.5$ is the memoryless case with the normal Brownian motion. In Figure 1, we first plot the strong error of the solution

$$\mathbb{E}|x_n - x(t_n)|^2$$

in terms of different time steps for some values of H to get a sense of convergence. Here cases of $H = 0.8, 0.6$ are plotted, respectively: (1) in the case of $H = 0.8$, the strong order reads $\min\{\frac{3}{2} - H, 3 - 3H\} = 0.6$. As can be seen in Figure 1, the slope is approximately 1.1718, and hence the convergence order is 0.586. (2) similarly in the case $H = 0.6$, the convergence rate numerically reads 0.945, which approximately matches the analytical results $\min\{\frac{3}{2} - H, 3 - 3H\} = 0.9$. Then Figure 2 shows the plot of convergence rate from our numerical results in terms of H, which matches the order proved in Theorem 3.2.

6.1.2. The strong convergence of the fast solver

For the strong convergence of the fast solver. Similarly as previous section, the time step for the reference solution is chosen as $k_m = 2^{-14}$, and $x_0 = 1$. We first plot the strong order of convergence for $H = 0.8$ and $H = 0.6$ in Figure 3, where both cases match the order $\min\{\frac{3}{2} - H, 3 - 3H\}$ as proved in Theorem 4.4. Next,

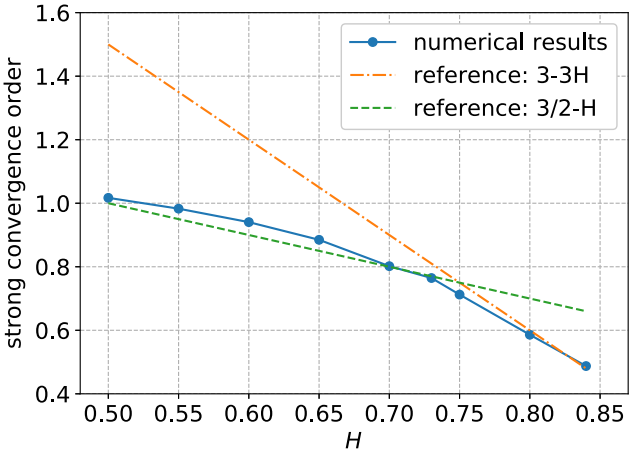
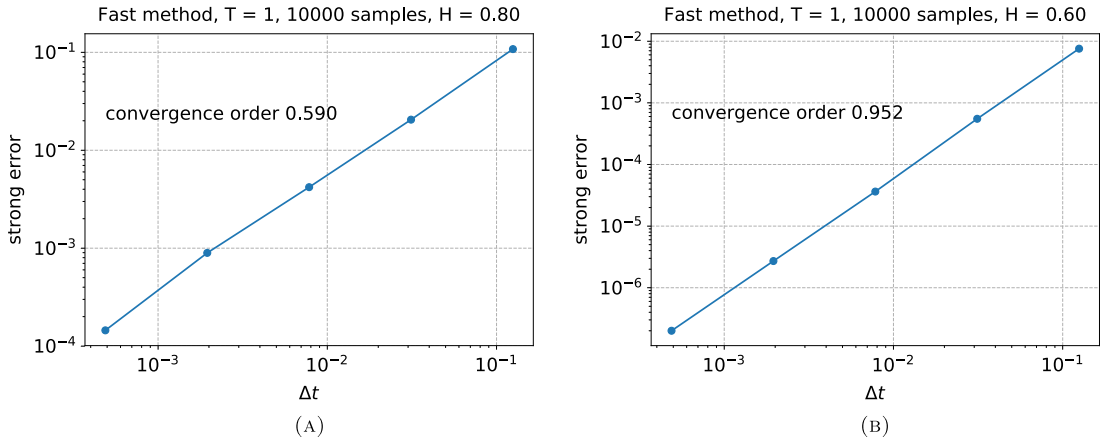


FIGURE 2. Example 1 (Sect. 6.1.1): The strong convergence of the direct method for $H = 0.5, 0.55, 0.6, 0.65, 0.7, 0.73, 0.75, 0.8, 0.84$ computed over 20 000 sample paths in terms of $\Delta t = 2^{-11}, 2^{-9}, 2^{-7}, 2^{-5}, 2^{-3}$. As can be seen in the figure, the convergence orders match the reference as is proved in Theorem 3.2.



(A) (B)

FIGURE 3. Example 1 (Sect. 6.1.2): The strong convergence of the direct method for $H = 0.8$ and $H = 0.6$, respectively in terms of various Δt in log-log scale. The convergence orders match the theoretical result $\min\{\frac{3}{2} - H, 3 - 3H\}$, as is proved in Theorem 3.2. (A) $H = 0.8$ and (B) $H = 0.6$.

the strong convergence order for various $H = 0.55, 0.6, 0.65, 0.7, 0.73, 0.75, 0.8, 0.84$ is plotted in Figure 4. The numerical orders agree with theoretical results.

Note that for the computational cost of the fast solver, as is shown in previous discussion is $O(NM)$, instead of $O(N^2)$, where M is the number of terms used in the approximation by exponential sums. For this example with $H = 0.8$, $\Delta t = 2^{-11}, T = 1$ and the tolerance $\epsilon = 10^{-9}$, we only need $M = 36$ while $N = 2048$. To compute longer time behavior as is interested here, say, $T = 128$, for $H = 0.8$ using $\Delta t = 2^{-9}$, one only needs $M = 44$ while $N = 65\,536$. This drastically reduced the computational cost especially with many sample paths to be simulated.

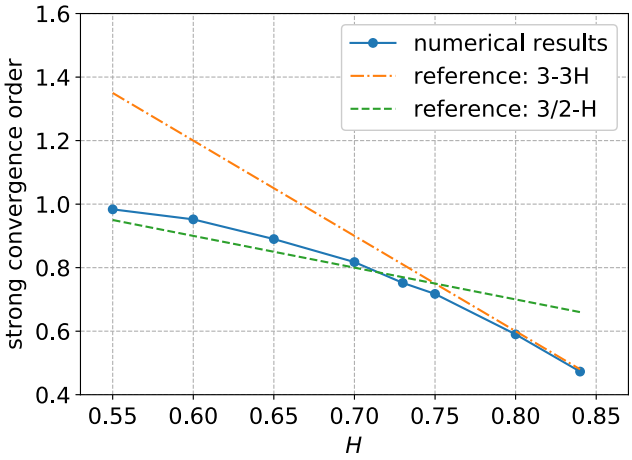


FIGURE 4. Example 1 (Sect. 6.1.2): The strong convergence of the fast method for $H = 0.55, 0.6, 0.65, 0.7, 0.73, 0.75, 0.8, 0.84$ computed over 10 000 sample paths in terms of $\Delta t = 2^{-11}, 2^{-9}, 2^{-7}, 2^{-5}, 2^{-3}$. The convergence orders match those proved in Theorem 4.4.

6.1.3. The test of ergodicity

As has been proved rigorously in Theorem 2 of [32] that for the linear force case, the process has ergodicity and converges algebraically to the Gibbs measure

$$\mu(\mathrm{d}x) \sim \exp\left(-\frac{1}{2}x^2\right)\mathrm{d}x. \tag{6.2}$$

Consider the initial data $x_0 = 0$, $H = 0.75$ and computed over 50 000 sample paths with $\Delta t = 2^{-9}$. Figure 5 plots the empirical distribution at different times $t = 0, 0.25, 1, 2, 8, 32$. Next, we plot the variance of $x(t)$ (also called the mean square displacement in physical literature) and its difference between its equilibrium. As can be seen in Figure 6, the variance of $x(t)$ convergences to its equilibrium $E(x^2(\infty)) = 1$ algebraically.

6.2. Example 2 (confining potential)

Here, we consider a 1D example, but with general potential $V(x)$ with confining structure. To be specific,

$$V(x) = ax^4 + bx^3 + cx^2,$$

is considered for its ergodicity, where $b = 0$ gives rise to the symmetric case whereas $b \neq 0$ corresponds to the asymmetric case.

6.2.1. Ergodicity of symmetric confining potential

Consider the symmetric double well potential

$$V(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2, \quad V'(x) = x^3 - x.$$

Consider the initial data $x_0 = 1$, $H = 0.6$ and computed over 50 000 sample paths with $\Delta t = 2^{-5}$ till the final time $T = 512$ using the fast algorithm. Figure 7 shows the empirical distribution at different times. It can be seen that the empiral distribution of x concentrates at $x = 1$ initially, then gradually expands and shifts to the left, and finally presents a symmetric double-well shape that matches the reference Gibbs measure

$$\mu(\mathrm{d}x) \sim \exp(-V(x))\,\mathrm{d}x.$$

Note that to consider the long time behavior, the initial values of x does not matter much.

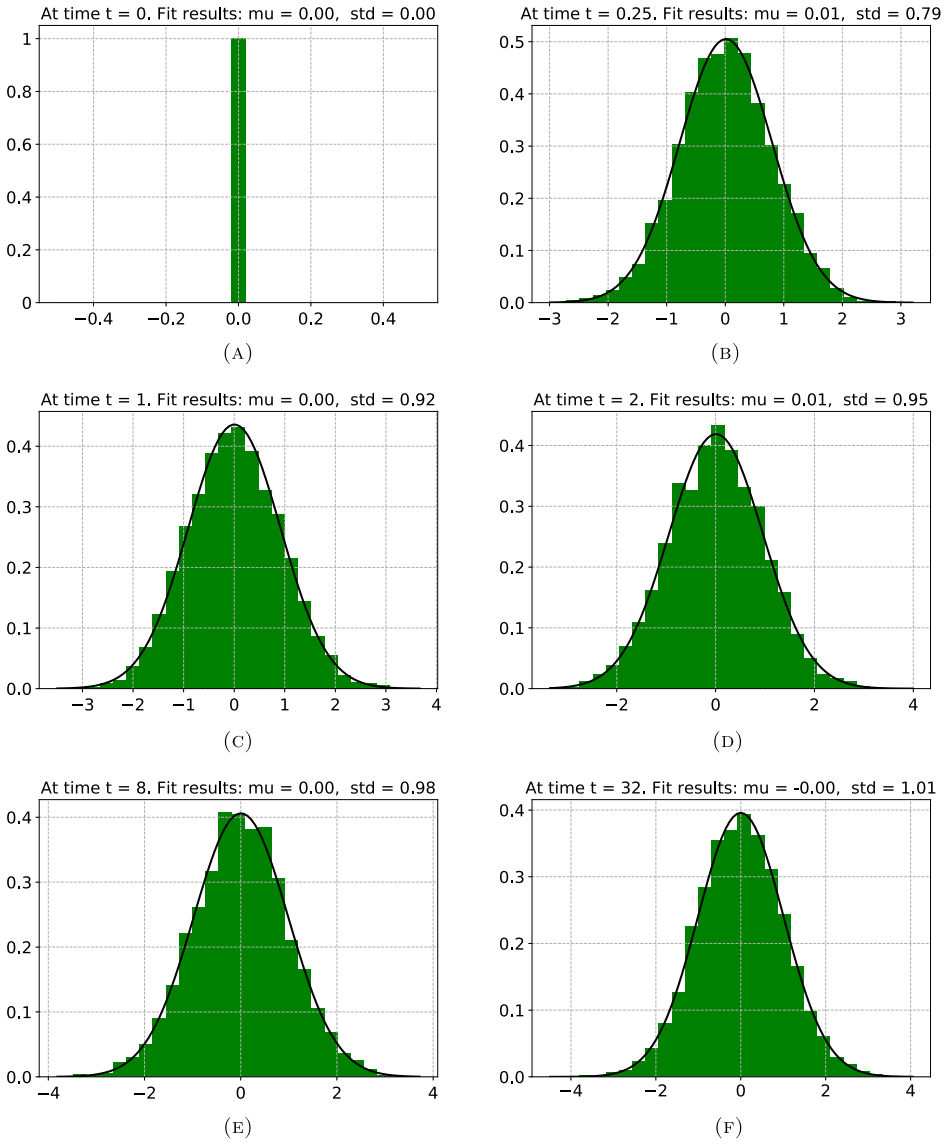


FIGURE 5. Example 1 (Sect. 6.1.3): the empirical distribution of x at different times $t = 0, 0.25, 1, 2, 8, 32$. It can be seen that the distribution stays as a gaussian shape and converges to the Gibbs measure (6.2) with mean 0 and variance 1. (A) $t = 0$, (B) $t = 0.25$, (C) $t = 1$, (D) $t = 2$, (E) $t = 8$ and (F) $t = 32$.

6.2.2. Ergodicity of asymmetric confining potential

Consider the asymmetric double well potential

$$V(x) = \frac{1}{4}x^4 + \frac{1}{3}x^3 - x^2, \quad V'(x) = x^3 + x^2 - 2x.$$

Consider the initial data $x_0 = 1$, $H = 0.6$ and computed over 50 000 sample paths with $\Delta t = 2^{-5}$ till the final time $T = 512$. Figure 8 plot the empirical distribution at different times. Since the initial data are all assigned

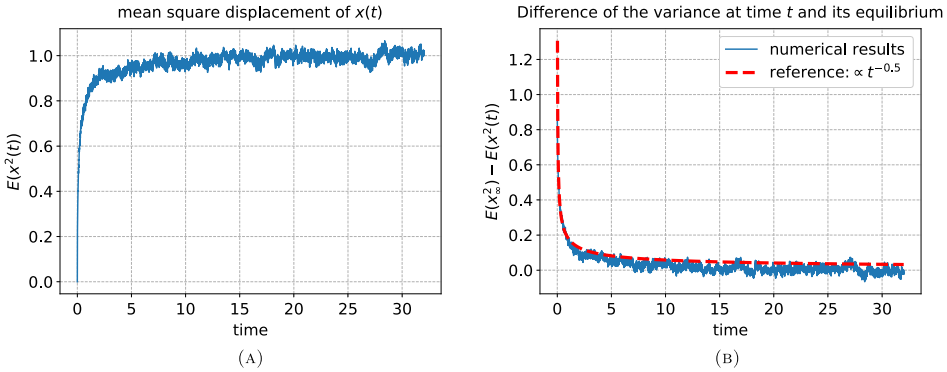


FIGURE 6. Example 1 (Sect. 6.1.3): The mean-square displacement of $x(t)$, *i.e.* the variance of $x(t)$ in the case $H = 0.75$. The left plot shows its tendency of approaching its equilibrium. The right figure plots its different between its equilibrium. As can be seen the decay rate is algebraic. (A) variance of $x(t)$ and (B) $E(x^2(\infty)) - E(x^2(t))$.

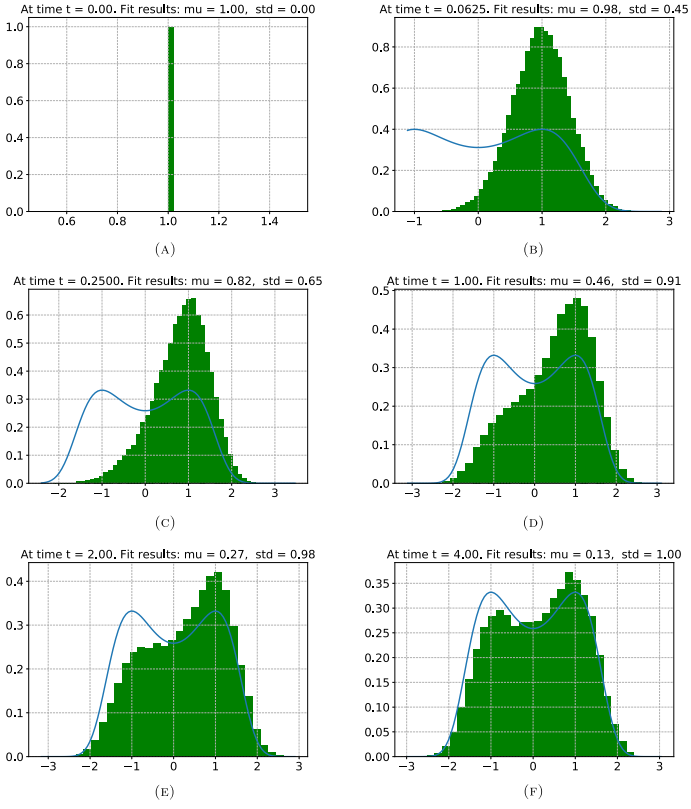


FIGURE 7. Example 2 (Sect. 6.2.1): the empirical distribution of x at different times $t = 0, 0.0625, 0.25, 1, 2, 4, 8, 16, 32, 512$. The solid line is the reference gibbs measure $\sim \exp(-V(x))$. It can be seen that given the intial data concentrating at $x = 0$, the distribution of x expands, and moves gradually to create a symmetric double-well shape. (A) $t = 0$, (B) $t = 0.0625$, (C) $t = 0.25$, (D) $t = 1$, (E) $t = 2$, (F) $t = 4$, (G) $t = 8$, (H) $t = 16$, (I) $t = 32$ and (J) $t = 512$.

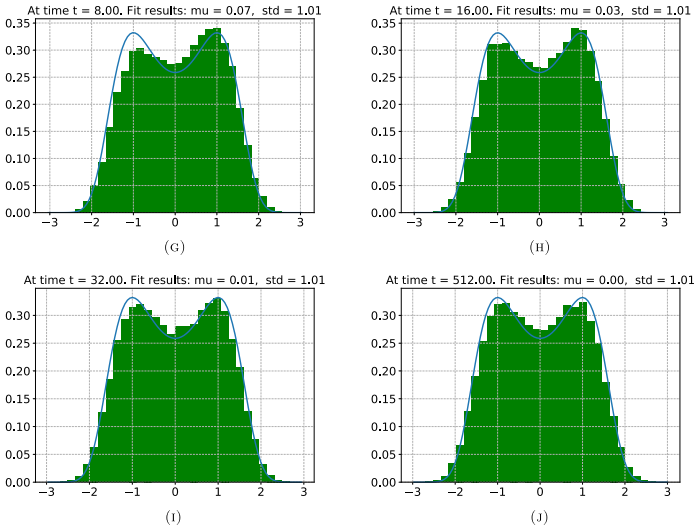


FIGURE 7. continued.

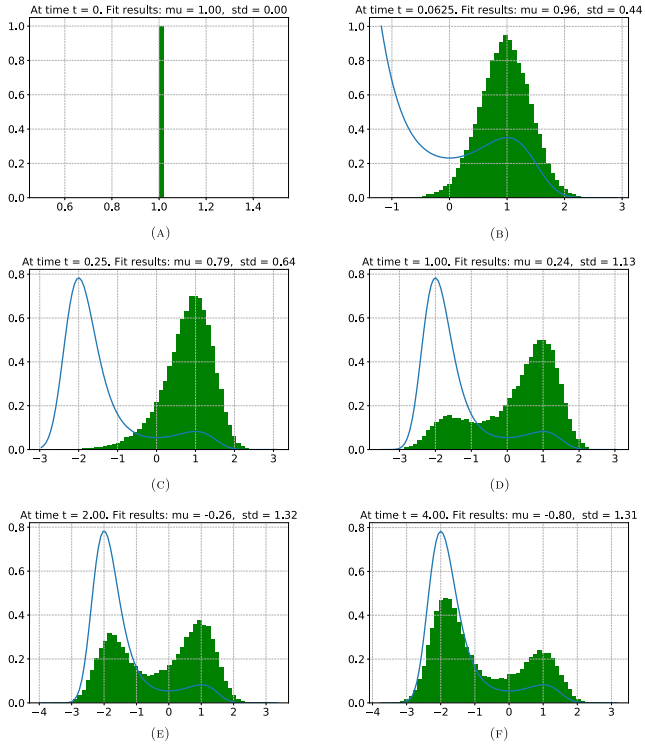


FIGURE 8. Example 2 (Sect. 6.2.2): the empirical distribution of x at different times $t = 0, 0.0625, 0.25, 1, 2, 4, 8, 16, 32, 512$. The solid line is the reference Gibbs measure $\sim \exp(-V(x))$. It can be seen that provided the intial data all gathering at $x = 1$, the distribution of x expands, and moves towards the left, and finally creates a double-well shape. (A) $t = 0$, (B) $t = 0.0625$, (C) $t = 0.25$, (D) $t = 1$, (E) $t = 2$, (F) $t = 4$, (G) $t = 8$, (H) $t = 16$, (I) $t = 32$ and (J) $t = 512$.

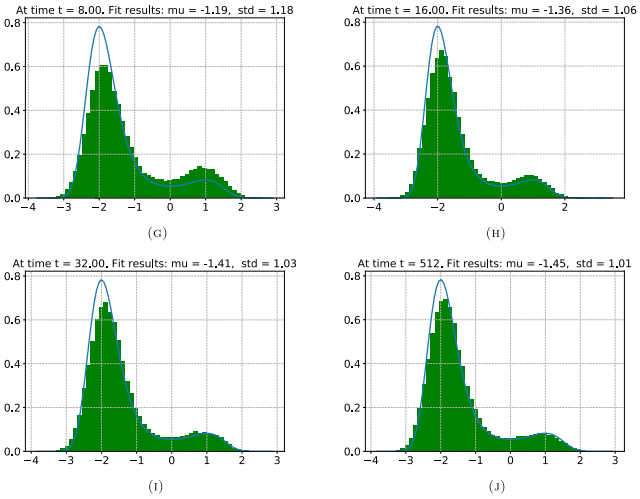


FIGURE 8. continued.

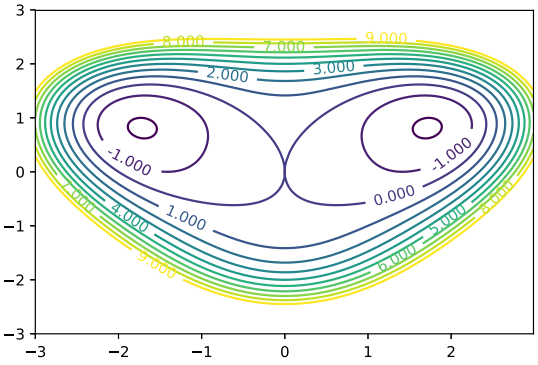


FIGURE 9. Contour plot of the double-well potential in Example 3 (Sect. 6.3).

as $x_0 = 1$, at first the empiral distribution of x concentrates at $x = 1$, then expands according to the time evolution, gradually shifts to the left and presents an asymmetric double-well shape in the end, which resembles the reference Gibbs measure

$$\mu(\mathrm{d}x) \sim \exp(-V(x))\,\mathrm{d}x.$$

6.3. Example 3 (2D double well potential)

For the 2D case, we should use the driven process as $B_H := (B_H^1, B_H^2)$ where B_H^i are two independent fractional Brownian motions. To see this, using the FDT, we find that the kernel is simply $\frac{1}{\Gamma(1-\alpha)}\tau^{-\alpha} I_d$ where I_d is the 2×2 identity matrix. Hence, the equation (2.1) still holds. The fractional Brownian motion B_H here is different from the 2D fractional Brownian random field (see *e.g.* [46]).

It is of interests to study double well potential in literature due to its application to describe the chemical phenomena, such as vibrionic spectra [34], proton transfer [24] and etc. Here in the numerical test, we consider the double-well potential

$$V(x, y) = \frac{1}{4}(x^2 + y^2)^2 - x^2 - x^2 y,$$

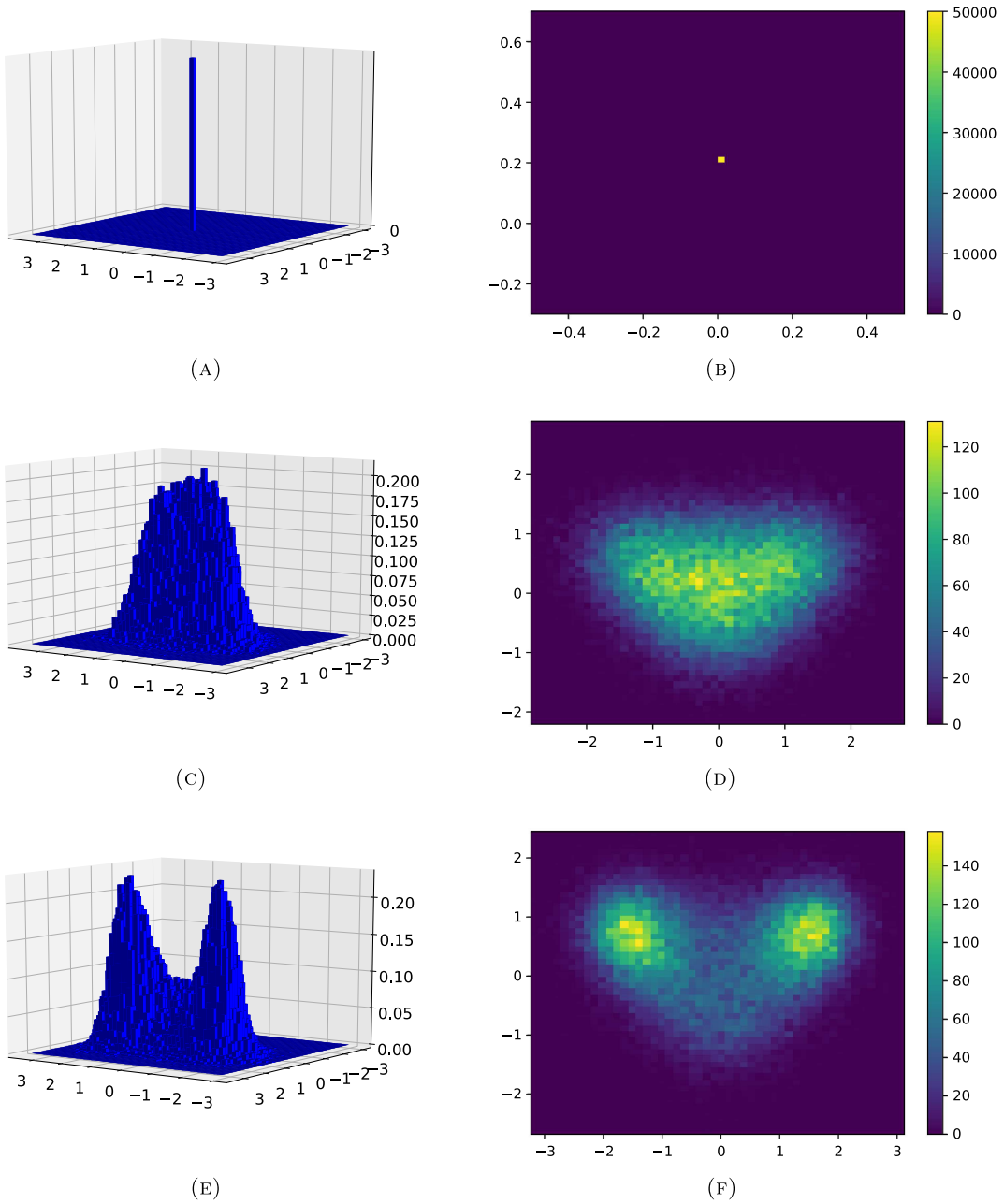


FIGURE 10. Example 3 (Sect. 6.3): the empirical distribution of x in 3D histogram and 2D contour, respectively, in at various moments $t = 0, 0.25, 1, 2, 32, 512$. The solid line is the reference Gibbs measure $\sim \exp(-V(x))$. It can be seen that provided the intial data all gathering at $x = (0, 0.2)$, the distribution of x expands, and moves towards a double-well shape. (A) $t = 0$, (B) $t = 0$, (C) $t = 0.25$, (D) $t = 0.25$, (E) $t = 1$, (F) $t = 1$, (G) $t = 2$, (H) $t = 2$, (I) $t = 32$, (J) $t = 32$, (K) $t = 512$ and (L) Reference Gibbs measure.

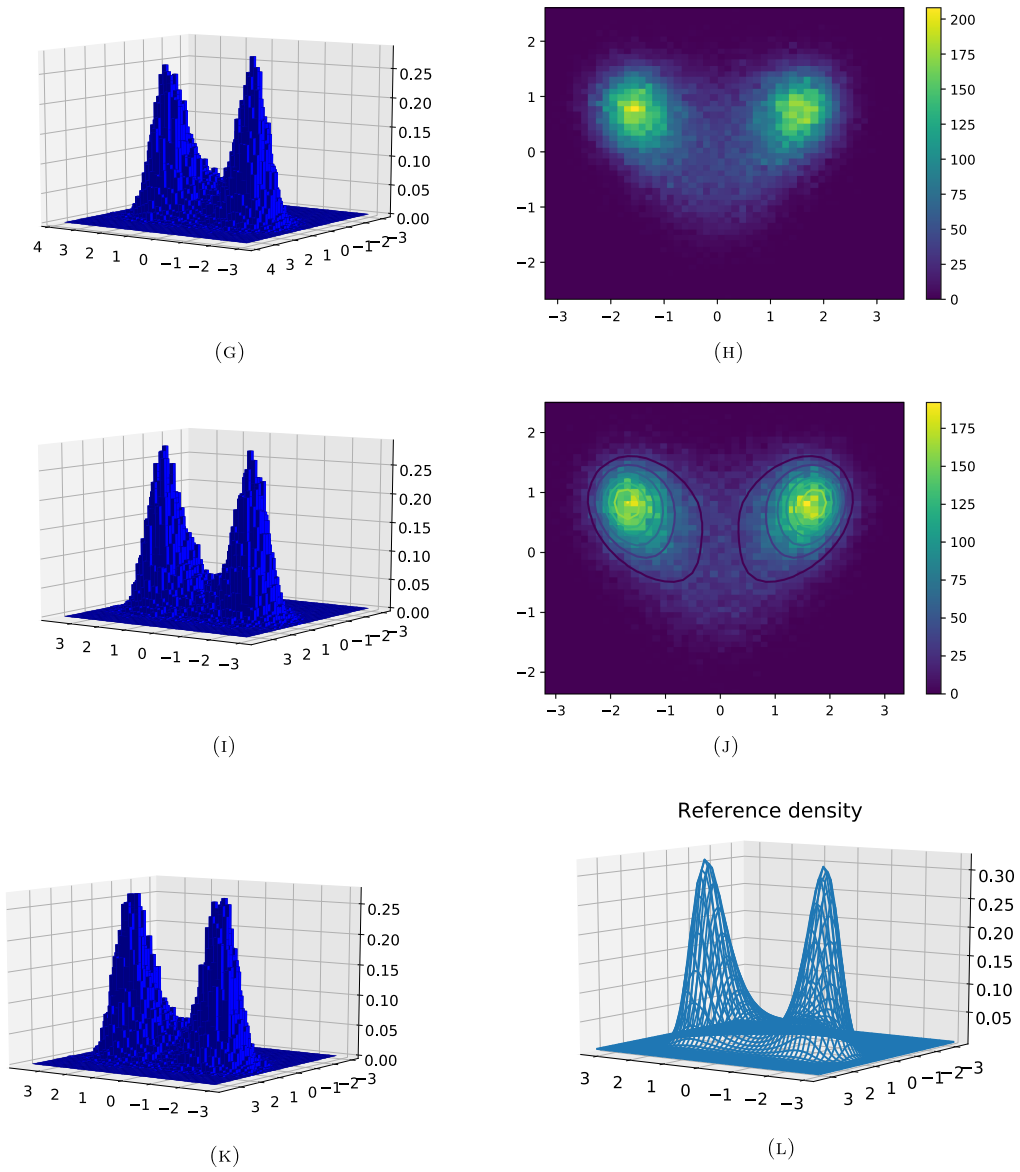
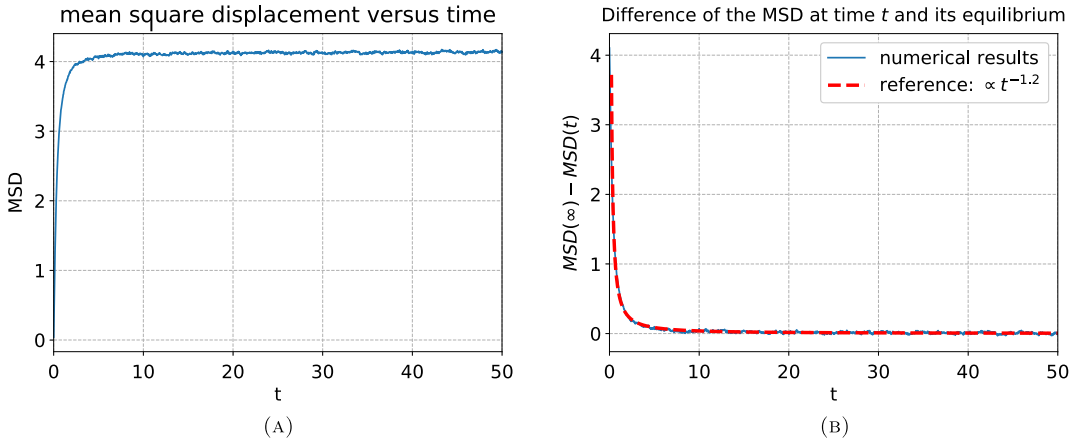


FIGURE 10. continued.

the quadratic coupling cases as is considered in, say, [48, 55], which can be visualized in Figure 9. In this example, consider intial datum $(x, y) = (0, 0.2)$ and compute till $T = 512$ with $\Delta t = 2^{-5}$ *via* the fast solver. We plot the empirical distribution at different time $t = 0, 0.25, 1, 2, 32, 512$ in Figure 10, where both 3D histogram and its 2D contour are plotted for the convenience of visualization. In Figure 11, the mean square displacement of $x(t)$, namely,

$$\mathrm{MSD}(t) := \mathbb{E}|x(t) - x(0)|^2, \tag{6.3}$$

is plotted in terms of time. The numerics indicate that the mean square displacement approaches to an equilibrium in an algebraic rate, instead of exponential.



(A) (B)

FIGURE 11. Example 3 (Sect. 6.3): The mean square displacement is defined as (6.3). It can be seen that the mean square displacement approaches an equilibrium algebraically, instead of exponentially. (A) Mean square displacement and (B) $\mathrm{MSD}(\infty) - \mathrm{MSD}(t)$.

APPENDIX A. PROOF OF LEMMA 3.4

In order to prove Lemma 3.4, we need the following lemma:

Lemma A.1 ([47], Thm. 3.1)**.** *Suppose $f \in C^{0,\beta}([0,T];B)$ for a Banach space B and $0 \leqslant \beta \leqslant 1$. Let $\alpha \in (0,1)$ and*

$$u(t) = u_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s)\, \mathrm{d}s.$$

Then,

$$u(t) = u_0 + \frac{f(0)}{\Gamma(1+\alpha)} t^\alpha + \psi(t), \tag{A.1}$$

where

$$\psi \in \begin{cases} C^{0,\beta+\alpha}([0,T];B), & \beta+\alpha < 1, \\ C^{1,\beta+\alpha-1}([0,T];B), & \beta+\alpha > 1, \\ C^{0,1;1}([0,T];B), & \beta+\alpha = 1. \end{cases} \tag{A.2}$$

Now, we are ready to prove Lemma 3.4:

Proof of Lemma 3.4. To simplify the notation, we denote

$$R(n,k) := \begin{cases} k^{6-6H} & H > \frac{3}{4}, \\ |\ln n| k^{6-6H} & H = \frac{3}{4}, \\ t_n^{3-4H} k^{3-2H} & H \in (\frac{1}{2}, \frac{3}{4}). \end{cases} \tag{A.3}$$

As in the proof of Proposition 3.1, we only need to estimate

$$R_n = \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x(t_{j-1})) - b(x(s)))\, \mathrm{d}s. \tag{A.4}$$

Recall

$$x(t) = \left(x_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} b(x(s))\, \mathrm{d}s\right) + G(t) =: \zeta(t) + \beta_H B_{1-H}(t),$$

where we have identified $G(t)$ with $\beta_H B_{1-H}$ since they have the same distribution.

Since $b(x) = Bx$, we have

$$R_n = \frac{B}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (\zeta(t_{j-1}) - \zeta(s))\, \mathrm{d}s + \frac{B\beta_H}{\Gamma(\alpha)} \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (B_{1-H}(t_{j-1}) - B_{1-H}(s))\, \mathrm{d}s =: R_{n,1} + R_{n,2}. \tag{A.5}$$

Step 1. We first of all estimate $R_{n,1}$. We denote

$$I_{n,1}^j := \frac{B}{\Gamma(\alpha)} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (\zeta(t_{j-1}) - \zeta(s))\, \mathrm{d}s,$$

and

$$R_{n,1}^m := \sum_{j=1}^{m} I_{n,1}^j. \tag{A.6}$$

Clearly, $R_{n,1} = R_{n,1}^n$. By the definition of $R_{n,1}^m$, we easily get that

$$\|R_{n,1}^m\|^2 = 2 \sum_{j=1}^{m-1} \mathbb{E}(R_{n,1}^j I_{n,1}^{j+1}) + \sum_{j=1}^{m} \|I_{n,1}^j\|^2,$$

since $R_{n,1}^1 = I_{n,1}^1$.

$$\|I_{n,1}^j\| \leqslant C \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \|\zeta(s) - \zeta(t_{j-1})\|\, \mathrm{d}s \leqslant Ck^{2-2H} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\, \mathrm{d}s, \tag{A.7}$$

where the estimate of $\|\zeta(s) - \zeta(t_{j-1})\|$ is due to that ζ is α-Hölder continuous by Lemma A.1 (see also the estimates of I_1, I_2 in Lem. 2.3).

Hence, we have

$$\sum_{j=1}^{n} \|I_{n,1}^j\|^2 \leqslant Ck^{4(1-H)} \left(k^{2\alpha} + \sum_{j=1}^{n-1} \left(\int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\, \mathrm{d}s \right)^2 \right) \leqslant Ck^{4(1-H)} \left(k^{2\alpha} + k^2 \sum_{j=1}^{n-1} (t_n - t_j)^{2\alpha-2} \right)$$
$$= Ck^{4(1-H)} \left(k^{2\alpha} + k^{2\alpha} \sum_{m=1}^{n-1} m^{2\alpha-2} \right) = Ck^{8-8H} \left(1 + \sum_{m=1}^{n-1} m^{2-4H} \right).$$

Depending on the behavior of $\sum_{m=1}^{n-1} m^{2-4H}$, the estimates branches into three cases. Clearly, when $4H - 2 > 1$, *i.e.* $H > 3/4$, the finite sum is bounded by a constant, while $H = 3/4$, it is bounded by $\ln n$. When $H < 3/4$, $\sum_{m=1}^{n-1} m^{2-4H} \leqslant \int_0^n x^{2-4H}\, \mathrm{d}x = \frac{n^{3-4H}}{3-4H}$. Hence, it is easily bounded by $CR(n,k)$.

Now turning to the first term in $R_{n,1}$ which is present for $n \geqslant 2$. Before starting the estimate, we shall sharpen the estimate on $\|\zeta(s) - \zeta(t_{j-1})\|$, since the same bound as in (A.7) is not enough for the desired results of these cross terms. To do so, we point out that $x \in C^{0,1-H}[0,T; L^2(\Omega)]$, so is $b(x(s))$. Hence one could apply Lemma A.1 to ζ with $f(s) = b(x(s))$ and $B = L^2(\Omega)$, and get

$$\zeta(s) = x_0 + \frac{b(x(0))}{\Gamma(1+\alpha)} s^{2-2H} + \psi(s),$$

where $\psi \in C^{0,3-3H}$ for $H > 2/3$ or $C^{1,2-3H}$ for $H < 2/3$. It follows that $\|\psi(s) - \psi(t_j)\| \leqslant Ck^{\min(1,3-3H)}$, $H \neq 2/3$. For $H = 2/3$, there is $|\log k|$ factor but overall it is bounded by $k^{\min(3/2-H,3-3H)}$. Now we are ready for

the estimation of $2\sum_{j=1}^{m-1}\mathbb{E}(R_{n,1}^j I_{n,1}^{j+1})$. We first apply Cauchy–Schwarz Inequality inequality and split ζ into ψ and s^{2-2H}.

$$\begin{aligned}2\sum_{j=1}^{m-1}\mathbb{E}\left(R_{n,1}^j I_{n,1}^{j+1}\right) &\leqslant C\sum_{j=1}^{m-1}\|R_{n,1}^j\|\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}\left(k^{\min(3/2-H,3-3H)}+\left(s^{2-2H}-t_j^{2-2H}\right)\right)\mathrm{d}s\\ &\leqslant C\sum_{j=1}^{m-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}\|R_{n,1}^j\|^2\,\mathrm{d}s + C\sum_{j=1}^{m-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}k^{2\min(3/2-H,3-3H)}\,\mathrm{d}s\\ &+ C\sum_{j=1}^{m-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}\left(s^{2-2H}-t_j^{2-2H}\right)^2\,\mathrm{d}s =: J_1+J_2+J_3.\end{aligned}$$

For J_1, we simply estimate it as

$$J_1 \leqslant C\int_0^{t_m}(t_n-s)^{\alpha-1}u_n(s)\,\mathrm{d}s \leqslant C\int_0^{t_m}(t_m-s)^{\alpha-1}u_n(s)\,\mathrm{d}s$$

where

$$\begin{aligned}&u_n(s) := \|R_{n,1}^j\|^2,\ s\in[t_j,t_{j+1}),\\ &R_{n,1}^0 = 0.\end{aligned}\tag{A.8}$$

J_2 is easily bounded:

$$J_2 \leqslant Ck^{2\min(3/2-H,3-3H)}t_n^\alpha \leqslant C(T)R(n,k).$$

For J_3, we first of all note

$$(s^{2-2H}-t_j^{2-2H})^2 \leqslant Ck^2t_j^{2-4H}.$$

Hence,

$$J_3 \leqslant Ck^2\sum_{j=1}^{m-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}t_j^{2-4H}\,\mathrm{d}s \leqslant Ck^2\left(k^{2-4H}k^{2-2H}+\sum_{j=2}^{n-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}t_j^{2-4H}\,\mathrm{d}s\right).$$

For the last term, we do a simple change of variables $s-k\to s$, and obktain

$$\sum_{j=2}^{n-1}\int_{t_j}^{t_{j+1}}(t_n-s)^{\alpha-1}t_j^{2-4H}\,\mathrm{d}s \leqslant \sum_{j=2}^{n-1}\int_{t_{j-1}}^{t_j}(t_{n-1}-s)^{\alpha-1}t_j^{2-4H}\,\mathrm{d}s \leqslant \int_k^{t_{n-1}}(t_{n-1}-s)^{\alpha-1}s^{2-4H}\,\mathrm{d}s.$$

If $H<3/4$, this integral is bounded by t_{n-1}^{4-6H}. If $H=3/4$, it is bounded by $\ln(n)t_{n-1}^{4-6H}$. If $H>3/4$, we have it bounded by

$$\int_k^{t_{n-1}}(t_{n-1}-s)^{\alpha-1}s^{2-4H}\,\mathrm{d}s = t_{n-1}^{4-6H}\int_{1/(n-1)}^1(1-s)^{\alpha-1}s^{2-4H}\,\mathrm{d}s \leqslant Ct_{n-1}^{4-6H}\left(\frac{1}{n-1}\right)^{3-4H} \leqslant Ct_{n-1}^{1-2H}k^{3-4H}.$$

Hence, we find that $J_3\leqslant CR(n,k)$ still holds.

Overall, we have

$$u_n(t_m) \leqslant C\int_0^{t_m}(t_m-s)^{\alpha-1}u_n(s)\,\mathrm{d}s + CR(n,k).$$

Applying Lemma 6.1 of [13], we find

$$\|R_{n,1}\|^2 = u_n(t_n) \leqslant C(T,H)R(n,k).$$

Step 2. We now estimate $R_{n,2}$. We similarly define

$$I_{n,2}^j = \frac{B\beta_H}{\Gamma(\alpha)} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (B_{1-H}(t_{j-1}) - B_{1-H}(s))\, \mathrm{d}s.$$

Then, it is clear that

$$\|R_{n,2}\|^2 = \left(\sum_j \|I_{n,2}^j\|^2 + 2 \sum_{i<j:j\leqslant i+3} \mathbb{E} I_{n,2}^i I_{n,2}^j \right) + 2 \sum_{i<j:j\geqslant i+4} \mathbb{E} I_{n,2}^i I_{n,2}^j =: K_1 + K_2.$$

Here we split the terms into two part so that the i and j in the second term are separated enough, which helps its estimate.

$$\|I_{n,2}^j\| \leqslant C \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \|B_{1-H}(t_{j-1}) - B_{1-H}(s)\|\, \mathrm{d}s \leqslant Ck^{1-H} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\, \mathrm{d}s,$$

where the increments of the fBm are estimated as term I_3 of Lemma 2.3. Then a straightforward calculation shows,

$$\begin{aligned} \sum_{j=1}^n \|I_{n,2}^j\|^2 &\leqslant Ck^{2(1-H)} \left(k^{2\alpha} + \sum_{j=1}^{n-1} \left(\int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\, \mathrm{d}s \right)^2 \right) \leqslant Ck^{2(1-H)} \left(k^{2\alpha} + k^2 \sum_{j=1}^{n-1} (t_n - t_j)^{2\alpha-2} \right) \\ &= Ck^{2(1-H)} \left(k^{2\alpha} + k^{2\alpha} \sum_{m=1}^{n-1} m^{2\alpha-2} \right) = Ck^{6-6H} \left(1 + \sum_{m=1}^{n-1} m^{2-4H} \right). \end{aligned}$$

Similar as the first part of Step 1, the estimates henceforth branch according to the behavior of $\sum_{m=1}^{n-1} m^{2-4H}$. Clearly, when $4H - 2 > 1$, *i.e.* $H > 3/4$, the finite sum is bounded by a constant, while $H = 3/4$, it is bounded by $\ln n$. When $H < 3/4$,

$$\sum_{m=1}^{n-1} m^{2-4H} \leqslant \int_0^n x^{2-4H}\, \mathrm{d}x = \frac{n^{3-4H}}{3-4H}.$$

Hence we then have

$$\sum_{j=1}^n \|I_{n,2}^j\|^2 \leqslant CR(n,k). \tag{A.9}$$

Clearly,

$$K_1 \leqslant 7 \sum_j \|I_{n,2}^j\|^2 \leqslant CR(n,k)$$

by (A.9).

For K_2, which is present only if $n \geqslant 8$, we use (2.7) and have

$$\begin{aligned} \mathbb{E} I_{n,2}^i I_{n,2}^j = C \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (t_n - \tau)^{\alpha-1} \big(&|s - t_{j-1}|^{2-2H} - |s - \tau|^{2-2H} \\ &+ |\tau - t_{i-1}|^{2-2H} - |t_{i-1} - t_{j-1}|^{2-2H} \big)\, \mathrm{d}\tau\, \mathrm{d}s. \end{aligned}$$

where

$$t_{i-1} \leqslant s \leqslant t_i < t_i + 3k \leqslant t_{j-1} \leqslant \tau \leqslant t_j.$$

We first apply mean value theorem for $t \mapsto |t-t_{j-1}|^{2-2H} - |t-\tau|^{2-2H}$ so that there exists $\xi \in (t_{i-1}, s)$ such that

$$|s-t_{j-1}|^{2-2H}-|s-\tau|^{2-2H}+|\tau-t_{i-1}|^{2-2H}-|t_{i-1}-t_{j-1}|^{2-2H} = (2-2H)(s-t_{i-1})\left(|\xi-t_{j-1}|^{1-2H} - |\xi-\tau|^{1-2H}\right)$$

Then, we apply mean value theorem again so that there exists $\eta \in (t_{j-1}, \tau)$ and have the bound

$$\begin{aligned}|s-t_{j-1}|^{2-2H} - |s-\tau|^{2-2H} + |\tau-t_{i-1}|^{2-2H} - |t_{i-1}-t_{j-1}|^{2-2H}\\ \leqslant (2-2H)(2H-1)(s-t_{i-1})(\tau-t_{j-1})|\xi-\eta|^{-2H} \leqslant Ck^2|\tau-s-2k|^{-2H}.\end{aligned}$$

We do change of variables $\tau - k \to \tau$ and $s+k \to s$ to find

$$\begin{aligned}\mathbb{E}I^i_{n,2}I^j_{n,2} &\leqslant Ck^2\int_{t_i}^{t_{i+1}}\int_{t_{j-2}}^{t_{j-1}}(t_n+k-s)^{\alpha-1}(t_n-k-\tau)^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau\,\mathrm{d}s\\ &\leqslant Ck^2\int_{t_i}^{t_{i+1}}\int_{t_{j-2}}^{t_{j-1}}(t_{n-1}-s)^{\alpha-1}(t_{n-1}-\tau)^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau\,\mathrm{d}s.\end{aligned}$$

It follows that

$$\begin{aligned}K_2 &\leqslant Ck^2\sum_{i\leqslant j-2, j\leqslant n-1}\int_{t_{i-1}}^{t_i}\int_{t_{j-1}}^{t_j}(t_{n-1}-s)^{\alpha-1}(t_{n-1}-\tau)^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau\,\mathrm{d}s\\ &\leqslant Ck^2\int_0^{t_{n-2}}\int_{s+k}^{t_{n-1}}(t_{n-1}-s)^{\alpha-1}(t_{n-1}-\tau)^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau\,\mathrm{d}s\\ &= Ck^2\int_k^{t_{n-1}}\int_0^{s-k}s^{\alpha-1}\tau^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau\,\mathrm{d}s,\end{aligned}$$

where in the last equality, we have made the change of variables $(t_{n-1}-s, t_{n-1}-\tau) \to (s,\tau)$. Since

$$\begin{aligned}\int_0^{s-k}\tau^{\alpha-1}|s-\tau|^{-2H}\,\mathrm{d}\tau &= s^{2-4H}\int_0^{1-k/s}\tau^{\alpha-1}|1-\tau|^{-2H}\,\mathrm{d}\tau\\ &\leqslant Cs^{2-4H}k^{1-2H}s^{2H-1} = Ck^{1-2H}s^{1-2H},\end{aligned}$$

we thus have

$$\begin{aligned}K_2 &\leqslant Ck^2k^{1-2H}\int_k^{t_{n-1}}s^{2-4H}\,\mathrm{d}s = Ck^{3-2H}t_{n-1}^{3-4H}\int_{1/(n-1)}^1 s^{2-4H}\,\mathrm{d}s\\ &\leqslant Ck^{3-2H}t_{n-1}^{3-4H}(1+k^{3-4H}t_{n-1}^{4H-3}) \leqslant Ck^{\min(3-2H,6-6H)}.\end{aligned}$$

Hence, we have

$$\|R^n_n\| \leqslant \begin{cases}Ck^{3-3H} & H > \frac{3}{4},\\ C\sqrt{|\ln k|}k^{3-3H} & H = \frac{3}{4},\\ Ck^{3/2-H} & H \in \left(\frac{1}{2}, \frac{3}{4}\right)\end{cases} =: \tilde{R}(T,k,H).$$

Using (3.4), we find

$$\|x_n - x(t_n)\| \leqslant \frac{L}{\Gamma(\alpha)}\sum_{j=1}^n\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\,\mathrm{d}s\|x_{j-1}-x(t_{j-1})\| + \tilde{R}(T,k,H).$$

Applying Lemma 6.1 of [13] again, we obtain the desired error bound. □

Appendix B. Proof of Lemma 4.2

Proof of Lemma 4.2. Recall the kernel $\gamma(t)$ of exponential sum approximation in (4.1), which is positive everywhere. Recall (4.7):

$$y(t) = y_0 + \int_0^t \gamma(t-s) f(y(s))\,\mathrm{d}s. \tag{B.1}$$

Assume that there are two continuous solutions $y_1(t)$ and $y_2(t)$ on some interval $I = [0,T] \cap [0,T_b)$. Let $T_1 = \min(T, T_b)$. We clearly have

$$|y_1(t) - y_2(t)| \leqslant \int_0^t \gamma(t-s)|f(y_1(s)) - f(y_2(s))|\,\mathrm{d}s.$$

Assume that $y_1(t) = y_2(t)$ for all $t < t^*$ for some $t^* \in [0, T_1)$. Then, we can pick δ small enough and then $|y_i(t)| \leqslant M$ for some $M > 0$ and all $t \in [0, t^* + \delta]$. Let L be the Lipschitz constant for f on the interval $[0, M]$. Then, for any $t \in [t^*, t^* + \delta]$,

$$|y_1(t) - y_2(t)| \leqslant \int_{t^*}^t \gamma(t-s)|f(y_1(s)) - f(y_2(s))|\,\mathrm{d}s \leqslant L \left(\sup_{s\in(t^*,t^*+\delta)} |y_1(s) - y_2(s)| \right) \sup_{t\in[t^*,t^*+\delta]} \int_{t^*}^t \gamma(t-s)\,\mathrm{d}s.$$

If δ is sufficiently small, we have

$$\nu = L \sup_{t\in[t^*,t^*+\delta]} \int_{t^*}^t \gamma(t-s)\,\mathrm{d}s < 1,$$

then we have

$$(1-\nu) \sup_{s\in[t^*,t^*+\delta]} |y_1(s) - y_2(s)| \leqslant 0.$$

This means the set of all such t^* is open in I with the inherited topology from $\mathbb{R}$. This set is clearly also closed under the inherited topology by the continuity of y_i. Hence, the set of all such t^* is I, which means $y_1(t) = y_2(t)$ on I.

Now, we establish the existence result and the desired properties. Consider the following standard Picard sequence,

$$y_{n+1}(t) = y_0 + \int_0^t \gamma(t-s) f(y_n(s))\,\mathrm{d}s, \quad y_0(t) = y_0.$$

By induction, it is not hard to see

$$y_{n+1}(t) \geqslant y_n(t), \ t \in [0, \infty),$$

and for each n $y_n(t) \geqslant y_0$ and is non-decreasing.

Clearly, $y_0 \leqslant v(t)$ for $t \in [0,T] \cap [0,T_b)$. Assume this is true for n. Consider y_{n+1}. For $t \in [0,k] \cap [0,T_b)$, it is easy to see

$$\begin{aligned} y_{n+1}(t) &\leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(y_n(s))\,\mathrm{d}s \\ &\leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(v(s))\,\mathrm{d}s + \frac{\epsilon T^{1-\alpha}}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(v(s))\,\mathrm{d}s \\ &= v(t). \end{aligned}$$

For $t \in (k, T] \cap [0, T_b)$, we have

$$\begin{aligned} y_{n+1}(t) &\leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(y_n(s))\,\mathrm{d}s + \frac{\epsilon}{\Gamma(\alpha)} \int_0^{t-k} f(y_n(s))\,\mathrm{d}s \\ &\leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(y_n(s))\,\mathrm{d}s + \frac{\epsilon T^{1-\alpha}}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(y_n(s))\,\mathrm{d}s \\ &\leqslant y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(v(s))\,\mathrm{d}s + \frac{\epsilon T^{1-\alpha}}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(v(s))\,\mathrm{d}s = v(t). \end{aligned}$$

Hence, on $[0, T] \cap [0, T_b)$, we have

$$y_{n-1}(t) \leqslant y_n(t) \leqslant \ldots \leqslant v(t).$$

This means $y_n(t)$ increases to a non-decreasing function $y(t)$ pointwise on $[0, T] \cap [0, T_b)$. By the monotone convergence theorem, we have

$$y(t) = y_0 + \int_0^t \gamma(t-s) f(y(s))\,\mathrm{d}s, \ \ \forall t \in [0, T] \cap [0, T_b).$$

Since $y(t) \leqslant v(t)$, we conclude that $y(t)$ must be continuous. This means that $y(\cdot)$ is a continuous solution with the desired properties.

If f is Lipschitz, it is well known that $v(\cdot)$ exists globally, or $T_b = \infty$ (see *e.g.* [13]) and the claim follows. □

Acknowledgements. The work of D. Fang is supported in part by by National Science Foundation under the grant DMS-1619778, and National Science Foundation Research Networks in Mathematical Sciences KI-Net under DMS-1107291 and DMS-1107444. The work of L. Li was partially sponsored by Shanghai Sailing Program 19YF1421300.

References

[1] B. Alpert, L. Greengard and T. Hagstrom, Rapid evaluation of nonreflecting boundary kernels for time-domain wave propagation. *SIAM J. Numer. Anal.* **37** (2000) 1138–1164.
[2] G. Beylkin and L. Monzón, On approximation of functions by exponential sums. *Appl. Comput. Harmon. Anal.* **19** (2005) 17–48.
[3] G. Beylkin and L. Monzón, Approximation by exponential sums revisited. *Appl. Comput. Harmon. Anal.* **28** (2010) 131–149.
[4] S. Burov and E. Barkai, Critical exponent of the fractional langevin equation. *Phys. Rev. Lett.* **100** (2008) 070601.
[5] S. Burov and E. Barkai, Fractional langevin equation: overdamped, underdamped, and critical behaviors. *Phys. Rev. E* **78** (2008) 031112.
[6] H.B. Callen and T.A. Welton, Irreversibility and generalized noise. *Phys. Rev.* **83** (1951) 34.
[7] W. Chu and X. Li, The mori-zwanzig formalism for the derivation of a fluctuating heat conduction model from molecular dynamics. Preprint arXiv:1709.05928 (2017).
[8] H. Ding, A high-order numerical algorithm for two-dimensional time-space tempered fractional diffusion-wave equation. *Appl. Numer. Math.* **135** (2019) 30–46.
[9] H. Ding and C. Li, A high-order algorithm for time-Caputo-tempered partial differential equation with Riesz derivatives in two spatial dimensions. *J. Sci. Comput.* **80** (2019) 81–109.
[10] G. Drazer and D.H. Zanette, Experimental evidence of power-law trapping-time distributions in porous media. *Phys. Rev. E* **60** (1999) 5858.
[11] T.E. Duncan, Y. Hu and B. Pasik-Duncan, Stochastic calculus for fractional Brownian motion. I. Theory. *SIAM J. Control Optim.* **38** (2000) 582–612.
[12] B.U. Felderhof, On the derivation of the fluctuation-dissipation theorem. *J. Phys. A* **11** (1978) 921–927.
[13] Y. Feng, L. Li, J.-G. Liu and X. Xu, Continuous and discrete one dimensional autonomous fractional ODEs. *Discrete Contin. Dyn. Syst. Ser. B* **23** (2018) 3109–3135.
[14] R. Gorenflo and F. Mainardi, Fractional calculus: integral and differential equations of fractional order. In: *Mathematical Physics*. Springer Verlag, New York (1997) 223–276.
[15] P. Guo, C. Zeng, C. Li and Y.Q. Chen, Numerics for the fractional Langevin equation driven by the fractional Brownian motion. *Fract. Calc. Appl. Anal.* **16** (2013) 123–141.
[16] C. Hijón, P. Español, E. Vanden-Eijnden and R. Delgado-Buscalioni, Mori–Zwanzig formalism as a practical computational tool. *Faraday Discuss.* **144** (2010) 301–322.

[17] S. Jiang and L. Greengard, Efficient representation of nonreflecting boundary conditions for the time-dependent Schrödinger equation in two dimensions. *Comm. Pure Appl. Math.* **61** (2008) 261–288.
[18] S. Jiang, L. Greengard and S. Wang, Efficient sum-of-exponentials approximations for the heat kernel and their applications. *Adv. Comput. Math.* **41** (2015) 529–551.
[19] S. Jiang, J. Zhang, Q. Zhang and Z. Zhang, Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun. Comput. Phys.* **21** (2017) 650–678.
[20] Y. Kantor and M. Kardar, Anomalous dynamics of forced translocation. *Phys. Rev. E* **69** (2004) 021806.
[21] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo, Theory and applications of fractional differential equations. In: Vol. 204 of *North-Holland Mathematics Studies*. Elsevier Science B.V., Amsterdam (2006).
[22] V. Kobelev and E. Romanov, Fractional langevin equation to describe anomalous diffusion. *Prog. Theor. Phys., Suppl.* **139** (2000) 470–476.
[23] S.C. Kou and X.S. Xie, Generalized langevin equation with fractional gaussian noise: subdiffusion within a single protein molecule. *Phys. Rev. Lett.* **93** (2004) 180603.
[24] P.M. Krasilnikov, Two-dimensional model of a double-well potential: proton transfer upon hydrogen bond deformation. *Biophysics* **59** (2014) 189–198.
[25] R. Kubo, The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **29** (1966) 255.
[26] B. Leimkuhler and M. Sachs, Ergodic properties of quasi-Markovian generalized Langevin equations with configuration dependent noise and non-conservative force. Preprint arXiv:1804.04029 (2018).
[27] J.-R. Li, A fast time stepping method for evaluating fractional integrals. *SIAM J. Sci. Comput.* **31** (2009/2010) 4696–4714.
[28] L. Li and J.-G. Liu, A generalized definition of Caputo derivatives and its application to fractional ODEs. *SIAM J. Math. Anal.* **50** (2018) 2867–2900.
[29] X. Li and C. Xu, A space-time spectral method for the time fractional diffusion equation. *SIAM J. Numer. Anal.* **47** (2009) 2108–2131.
[30] Z. Li, X. Bian, X. Li and G.E. Karniadakis, Incorporation of memory effects in coarse-grained modeling via the Mori–Zwanzig formalism. *J. Chem. Phys.* **143** (2015) 243128.
[31] Z. Li, H.S. Lee, E. Darve and G.E. Karniadakis, Computing the non-Markovian coarse-grained interactions derived from the Mori–Zwanzig formalism in molecular systems: application to polymer melts. *J. Chem. Phys.* **146** (2017) 014104.
[32] L. Li, J.-G. Liu and J. Lu, Fractional stochastic differential equations satisfying fluctuation-dissipation theorem. *J. Stat. Phys.* **169** (2017) 316–339.
[33] Y. Lin and C. Xu, Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.* **225** (2007) 1533–1552.
[34] C.-K. Lin, H.-C. Chang and S.H. Lin, Symmetric double-well potential model and its application to vibronic spectra: studies of inversion modes of ammonia and nitrogen-vacancy defect centers in diamond. *J. Phys. Chem. A* **111** (2007) 9347–9354.
[35] H.P. Lu, L. Xun and X.S. Xie, Single-molecule enzymatic dynamics. *Science* **282** (1998) 1877–1882.
[36] B.B. Mandelbrot and J.W. Van Ness, Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** (1968) 422–437.
[37] Z. Mao, S. Chen and J. Shen, Efficient and accurate spectral method using generalized Jacobi functions for solving Riesz fractional differential equations. *Appl. Numer. Math.* **106** (2016) 165–181.
[38] U. Marconi, A. Puglisi, L. Rondoni and A. Vulpiani, Fluctuation–dissipation: response theory in statistical physics. *Phys. Rep.* **461** (2008) 111–195.
[39] W. McLean, Exponential Sum Approximations for $t^{-\beta}$. Springer, Cham (2018) 911–930.
[40] R. Metzler and J. Klafter, The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Phys. A Math. Gen.* **37** (2004) R161.
[41] R. Metzler and J. Klafter, When translocation dynamics becomes anomalous. *Biophys. J.* **85** (2003) 2776.
[42] H. Mori, A continued-fraction representation of the time-correlation functions. *Prog. Theoret. Phys.* **34** (1965) 399–416.
[43] H. Mori, Transport, collective motion, and brownian motion. *Prog. Theor. Phys.* **33** (1965) 423–455.
[44] R.H. Nochetto, E. Otárola and A.J. Salgado, A PDE approach to space-time fractional parabolic problems. *SIAM J. Numer. Anal.* **54** (2016) 848–873.
[45] D. Nualart, Fractional Brownian motion: stochastic calculus and applications. In: Vol. III of *International Congress of Mathematicians*. Eur. Math. Soc., Zürich (2006) 1541–1562.
[46] H. Qian, G.M. Raymond and J.B. Bassingthwaighte, On two-dimensional fractional Brownian motion and fractional Brownian random field. *J. Phys. A Math. Gen.* **31** (1998) L527.
[47] S.G. Samko, A.A. Kilbas and O.I. Marichev, Fractional Integrals and Derivatives: Theory and Applications. Gordon and Breach, Yverdon Publishers (1993).
[48] J.C. Sankey, C. Yang, B.M. Zwickl, A.M. Jayich and J.G.E. Harris, Strong and tunable nonlinear optomechanical coupling in a low-loss system. *Nat. Phys.* **6** (2010) 707.
[49] G.J. Schütz, H. Schindler and T. Schmidt, Single-molecule microscopy on model membranes reveals anomalous diffusion. *Biophys. J.* **73** (1997) 1073–1080.
[50] P. Schwille, J. Korlach and W.W. Webb, Fluorescence correlation spectroscopy with single-molecule sensitivity on cell and model membranes. *Cytometry A* **36** (1999) 176–182.
[51] C. Sheng and J. Shen, A space-time Petrov–Galerkin spectral method for time fractional diffusion equation. *Numer. Math. Theory Methods Appl.* **11** (2018) 854–876.

[52] G. Shevchenko, Fractional Brownian motion in a nutshell. In: Vol. 36 of *International Journal of Modern Physics: Conference Series*. World Scientific (2015) 1560002.

[53] A. Taloni, A. Chechkin and J. Klafter, Generalized elastic model yields a fractional langevin equation description. *Phys. Rev. Lett.* **104** (2010) 160602.

[54] I.M. Tolić-Nørrelykke, E.-L. Munteanu, G. Thon, L. Oddershede and K. Berg-Sørensen, Anomalous diffusion in living yeast cells. *Phys. Rev. Lett.* **93** (2004) 078102.

[55] M. Topaler and N. Makri, Quantum rates for a double well coupled to a dissipative bath: accurate path integral results and comparison with approximate theories. *J. Chem. Phys.* **101** (1994) 7500–7519.

[56] D. Venturi, H. Cho and G.E. Karniadakis, Mori–Zwanzig approach to uncertainty quantification. In: Vols. 1–3 of *Handbook of Uncertainty Quantification*. Springer, Cham (2017) 1037–1073.

[57] M. Weiss, H. Hashimoto and T. Nilsson, Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys. J.* **84** (2003) 4043–4052.

[58] M. Weiss, M. Elsner, F. Kartberg and T. Nilsson, Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophys. J.* **87** (2004) 3518–3524.

[59] Y. Yan, Z.-Z. Sun and J. Zhang, Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations: a second-order scheme. *Commun. Comput. Phys.* **22** (2017) 1028–1048.

[60] H. Yang, G. Luo, P. Karnchanaphanurach, T.-M. Louie, I. Rech, S. Cova, L. Xun and X.S. Xie, Protein conformational dynamics probed by single-molecule electron transfer. *Science* **302** (2003) 262–266.

[61] B. Yuttanan and M. Razzaghi, Legendre wavelets approach for numerical solutions of distributed order fractional differential equations. *Appl. Math. Model.* **70** (2019) 350–364.

[62] J. Zhang, D. Li and X. Antoine, Efficient numerical computation of time-fractional nonlinear Schrödinger equations in unbounded domain. *Commun. Comput. Phys.* **25** (2019) 218–243.

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