

# The generalized Langevin equation with power-law memory in a nonlinear potential well

Nathan E Glatt-Holtz<sup>1</sup>, David P Herzog<sup>2</sup>, Scott A McKinley<sup>1</sup>  
and Hung D Nguyen<sup>3,4</sup>

<sup>1</sup> Department of Mathematics, Tulane University, 6823 St Charles Ave, New Orleans, Louisiana, 70118, United States of America

<sup>2</sup> Department of Mathematics, Iowa State University, 411 Morrill Rd, Ames, Iowa 50011, United States of America

<sup>3</sup> Department of Mathematics, University of California, Los Angeles. 520 Portola Plaza, Los Angeles, California 90095, United States of America

E-mail: [hungdnguyen@math.ucla.edu](mailto:hungdnguyen@math.ucla.edu)

Received 14 August 2019, revised 28 January 2020

Accepted for publication 10 February 2020

Published 14 April 2020



CrossMark

Recommended by Dr Lev Tsimring

## Abstract

The generalized Langevin equation (GLE) is a stochastic integro-differential equation that has been used to describe the velocity of microparticles in viscoelastic fluids. In this work, we consider the large-time asymptotic properties of a Markovian approximation to the GLE in the presence of a wide class of external potential wells. The qualitative behavior of the GLE is largely determined by its memory kernel  $K$ , which summarizes the delayed response of the fluid medium on the particles past movement. When  $K$  can be expressed as a finite sum of exponentials, it has been shown that long-term time-averaged properties of the position and velocity do not depend on  $K$  at all. In certain applications, however, it is important to consider the GLE with a power law memory kernel. Using the fact that infinite sums of exponentials can have power law tails, we study the infinite-dimensional version of the Markovian GLE in a potential well. In the case where the memory kernel  $K$  is integrable (i.e. in the asymptotically diffusive regime), we are able to extend previous results and show that there is a unique stationary distribution for the GLE system and that the long-term statistics of the position and velocity do not depend on  $K$ . However, when  $K$  is not integrable (i.e. in the asymptotically subdiffusive regime), we are able to show the existence of an invariant probability measure but uniqueness remains an open question. In

<sup>4</sup> Author to whom any correspondence should be addressed.

particular, the method of asymptotic coupling used in the integrable case to show uniqueness does not apply when  $K$  fails to be integrable.

Keywords: anomalous diffusion, asymptotic coupling, invariant measure

Mathematics Subject Classification numbers: 60H10

## 1. Introduction

The movement of microparticles in biological fluids is often distinct from classical Brownian motion [30]. While some particles exhibit non-Gaussian [47, 67] and/or switching behavior [56, 59], an important category of *anomalous diffusion* includes paths with stationary Gaussian increments that are negatively correlated with each other [10, 29, 41, 69]. In particular, anticorrelation can accumulate in such a way that the particle process  $\{x(t)\}_{t \geq 0}$  has a mean squared displacement (MSD,  $\mathbb{E}[|x|^2(t)]$ ) that is sublinear over a significant period of time [46, 53, 57]. Such a process is commonly called *subdiffusive*, and there have been numerous perspectives on how to model [16, 32, 34, 42, 46, 50, 54] and statistically analyze [6, 11, 33, 41, 49, 55] individual, unconstrained, subdiffusive microparticle paths.

It is also natural to investigate the behavior of subdiffusive particles when they are subjected to external forces. Commonly the magnitude and direction of these forces are expressed as the negative gradient of a space-dependent potential energy function  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ . In early works, it was common to study quadratic potentials ( $\Phi(x) = \kappa|x|^2$ ) because the equations that govern such particle motion are linear and admit exact solutions [2, 42, 46]. More recently there has been success in modeling and simulating the behavior of subdiffusive particles in nonlinear potentials as well. As we discuss below, it is possible to develop a model that exhibits ‘transient subdiffusive behavior’ (sublinear MSD over several log-decades of time [53, 54]) in the absence of a confining potential, and permits an exact expression for the stationary joint distribution of position and velocity when subjected to one of an appropriate class of nonlinear potentials [16, 60, 61]. There are a host of classical questions that can be rigorously studied for Brownian motion in such potentials, which provides a foundation to ask similar questions for subdiffusive particles. Several such questions were introduced and surveyed by Goychuk in 2009 [14] and 2012 [16]: for example, addressing escape times between minima in double-welled potentials (or more recently, escape times in stationary Gaussian potentials [20]), effective diffusivity in static periodic potentials, effective velocity induced by time-dependent potentials known as flashing or rocking ratchets (see also [15, 21, 22, 39, 40]). Since organelles and other microparticles (or macromolecules) have been observed to exhibit significant subdiffusive behavior in cytoplasm [65, 68, 72] and other biological fluids like mucus [29], applications of these theoretical models have emerged in recent years: for example in the study of magnetic nanoparticles [18, 19] and the modeling of intracellular transport by microtubule-associated molecular motors [17, 23, 24]. Moreover, it is commonplace for experimentalists to probe fluid-mechanical properties of live cells by manipulation through ‘optical tweezers’ [58, 70, 73]. However, such studies rarely take into account the subdiffusive character of the microparticle probes, or the possibly nonlinear forces exerted by the trap.

In this work, we take a step advancing the theory for a set of models that are ‘fully subdiffusive’ in a sense described below. Such models are necessarily infinite-dimensional, which makes it non-trivial to establish existence and uniqueness of stationary measures, and points towards fundamental questions about if and on what time scale a system might ‘forget’ its initial conditions.

1.1. Model development and the main question

To describe microparticle motion in *viscous* fluids, it is common to use a Langevin framework. Let  $\{(x(t), v(t))\}_{t \geq 0}$  denote the position and velocity of a particle, and let  $\Phi(x)$  denote the particle's potential energy at the position  $x$ . (Because it does not have a substantive impact on our results, we will study the dynamics in one dimension). Newton's Second Law yields [61]

$$m \dot{v}(t) = -\gamma v(t) - \Phi'(x(t))dt + \sqrt{2\gamma}dW(t), \tag{1.1}$$

where  $x'(t) = v(t)$ . Here  $m$  is the particle's mass,  $\gamma$  is the viscous drag coefficient, and  $W(t)$  is a standard Brownian motion. To be physically correct, the coefficient of the noise should be  $\sqrt{2k_B T \gamma}$ , where  $k_B$  is Boltzman's constant and  $T$  is the temperature, but for the sake of notational simplicity we will set  $k_B T = 1$  throughout. Under appropriate conditions on the potential well, this system has a unique stationary distribution with density

$$\pi(x, v) \propto \exp\left(-\left(\Phi(x) + \frac{m}{2}v^2\right)\right) \tag{1.2}$$

and is geometrically ergodic, in the sense that the law of the process converges to the stationary distribution exponentially quickly (see, for example, [5, 28, 52, 61] as well as [4, 25] for related results). Birkoff's ergodic theorem in turn implies

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(x(s), v(s))ds = \int_{\mathbb{R}^2} f(x, v)\pi(x, v)dxdv, \quad \pi\text{-a.s. and in } L^1(\pi) \tag{1.3}$$

for any  $f \in L^1(\pi)$ . Ultimately, we are interested in whether or not such a property holds for viscoelastic diffusion in a nonlinear potential well.

In order to model the memory effects arising in viscoelastic diffusion, physicists have long employed the generalized Langevin equation (GLE), which adds a memory kernel  $K : \mathbb{R} \rightarrow \mathbb{R}_+$  and a stationary Gaussian process  $\{F(t)\}_{t \in \mathbb{R}}$  to the Langevin dynamics (1.1). In particular, the GLE in a potential well  $\Phi$  can be written formally as

$$m \dot{v}(t) = -\Phi'(x(t)) - \gamma v(t) - \int_{-\infty}^t K(t-s)v(s) ds + \sqrt{2\gamma} \dot{W}(t) + F(t). \tag{1.4}$$

for  $t \geq 0$ , where we assume that  $\mathbb{E}[F(t)F(s)] = K(t-s)$  in order to satisfy the fluctuation-dissipation relation [31, 33, 43]. In general, there can be a pair of coefficients in front of the memory terms, but they do not affect our analysis. We refer the reader to [32] for a physical interpretation of those parameters. In comparison with the classical GLE version in [43], the noise term in (1.4) can be regarded as a sum of two independent noises  $\xi(t) = F(t) + \sqrt{2\gamma}\dot{W}(t)$ , such that  $\mathbb{E}[F(t)\dot{W}(t)] = 0$ , and  $\mathbb{E}[\sqrt{2\gamma}\dot{W}(t)\sqrt{2\gamma}\dot{W}(t')] = 2\gamma\delta_0(|t-t'|)$ . In this case, the memory term  $\gamma(t)$  is a sum of the Dirac delta function and  $K(t)$ , i.e.  $\gamma(t) = \gamma\delta(t) + K(t)$  and thus

$$\int_{-\infty}^t \gamma(t-s)v(s) ds = \gamma v(t) + \int_{-\infty}^t K(t-s)v(s) ds.$$

Separation of  $F(t)$  and  $\dot{W}(t)$  will help with the analysis later. We also note a distinction in the support of the memory kernel. Namely, in the physics literature, the memory kernel in the GLE is often defined on the interval  $[0, t]$  rather than  $(-\infty, t]$  as we have here. In order for a stationary integro-differential equation to be properly defined we must specify initial data on an infinite horizon. The classical definition effectively defines the velocity to be zero on the interval  $(-\infty, 0]$ . This assumption (along with a quadratic potential well), allow for use of the

Laplace transform, which can yield a host of statistical properties including an exact solution [8, 66]. Unfortunately, such an approach is not available here.

Due to the memory kernel  $K$ , in this formulation of GLE, the joint process  $(x(t), v(t))$  is non-Markovian. Therefore, it is not immediately clear what we mean by a ‘stationary distribution.’ However, if the memory kernel  $K(t)$  is a sum of  $N$  exponential functions, we can use the so-called Mori–Zwanzig formalism [14, 60, 74] to rewrite the GLE as a  $2 + N$  (or  $2 + 2N$  [11]) dimensional system of SDEs. (For a contrast on the two representations, see [33].) When  $\Phi$  is quadratic, such a representation is statistically equivalent to (1.4); otherwise, we simply call this the Markovian approximation of the GLE. This finite-dimensional Markovian version of the GLE does admit a stationary distribution and one can show that the marginal distribution of the pair  $(x, v)$  in stationarity is exactly (1.2) [16, 60, 61]. The fact that the memory kernel does not affect the stationary statistics of  $x$  and  $v$  is, in some sense, a natural generalization of the observation that the drag coefficient  $\gamma$  does not appear in  $\pi(x, v)$  for viscous diffusion. It is then reasonable to ask whether this property holds for more general forms of  $K$ .

The sum-of-exponential form for  $K$  is a very useful construct, but it turns out that restricting ourselves to *finitely* many terms neglects an important qualitative regime. Indeed, if  $K \in L^1(\mathbb{R})$ , then the associated unconstrained GLE ( $\Phi \equiv 0$ ) is always *asymptotically diffusive* in the sense that  $\mathbb{E}[x^2(t)] \sim t$  as  $t \rightarrow \infty$  [53, 60]. Here we write  $f(t) \sim g(t)$  as  $t \rightarrow \infty$  if  $\lim_{t \rightarrow \infty} f(t)/g(t) = C \in (0, \infty)$ . However, if  $K \notin L^1(\mathbb{R})$  but  $K(t) \sim t^{-\alpha}$  for some  $\alpha \in (0, 1)$  as  $t \rightarrow \infty$ , then under mild restrictions, the unconstrained GLE is *asymptotically subdiffusive*, i.e.  $\mathbb{E}[x^2(t)] \sim t^\alpha$  [53]. Moreover, a new *critical* regime is recently found when  $K(t) \sim t^{-1}$ , for which case,  $\mathbb{E}[x^2(t)] \sim t/\log t$  [9]. See also [41, 62, 64] for related results. We are primarily interested in memory kernels with power law tails, which can fall in either qualitative regime. As has been observed elsewhere [1, 14], it is possible for an *infinite* sum of exponentials to have a power law tail. Therefore, an infinite-dimensional version of the Mori–Zwanzig formalism is an appropriate way to study the GLE with power law memory.

In this work, we explore the infinite-dimensional Markovian version of the GLE with an eye toward addressing the fundamental question of whether  $(x(t), v(t))$  is ergodic in the sense of (1.3). In section 2 we introduce notation, explicitly define our model, and summarize the main results. In section 3 we establish well-posedness for all values of  $\alpha > 0$  (including both the asymptotically diffusive and subdiffusive cases). Using an extension of the invariant measure previously established for the finite-dimensional GLE, in section 4, we demonstrate the existence of an explicitly defined invariant probability measure in the infinite-dimensional case, see theorem 7. In section 5, we use *asymptotic coupling* [26, 51, 71] to establish uniqueness of this measure in the asymptotically diffusive case ( $\alpha > 1$ ) under a wide class of nonlinear potentials including those polynomials of even degree, see assumption 8 and theorem 10. We finish with section 6 which contains conclusions and discussions of open and related problems. In particular, we cannot extend our proof of unique ergodicity to the asymptotically nondiffusive case ( $\alpha \in (0, 1]$ ). We have yet to determine whether this is a shortcoming of current methods, or if the claim is simply not true. We hope that this work will shine some light on this interesting question.

## 2. Notation and rigorous summary of results

Suppose that a memory kernel  $K(t)$  has the form

$$K(t) = \sum_{k=1}^N c_k e^{-\lambda_k t}, \quad (2.1)$$

where  $c_k, \lambda_k > 0, k = 1, \dots, N$ . Then, following Chapter 8 of [61], which summarizes pre-existing work on the same topic [16, 60], we can use Duhamel’s formula and set

$$z_k(t) = e^{-\lambda_k t} z_k(0) + \sqrt{c_k} \int_0^t e^{-\lambda_k(t-s)} v(s) ds + \sqrt{2\lambda_k} \int_0^t e^{-\lambda_k(t-s)} dW_k(s).$$

In order to approximate equation (1.4) as an  $(N + 2)$ -dimensional Markov system, we write

$$\begin{aligned} dx(t) &= v(t) dt, \\ m dv(t) &= \left( -\gamma v(t) - \Phi'(x(t)) - \sum_{k=1}^N \sqrt{c_k} z_k(t) \right) dt + \sqrt{2\gamma} dW_0(t), \\ dz_k(t) &= (-\lambda_k z_k(t) + \sqrt{c_k} v(t)) dt + \sqrt{2\lambda_k} dW_k(t), \quad 1 \leq k \leq N, \end{aligned} \tag{2.2}$$

where  $z_k(0)$  are i.i.d.  $\mathcal{N}(0, 1)$  random variables and  $(W_0, W_1, \dots, W_N)$  is a standard  $(N + 1)$ -dimensional Brownian motion. This is the content of proposition 8.1 of [61]. It is also known (see proposition 8.2, [61]) that the system (2.2) is uniquely ergodic with an invariant probability density function  $\varrho(x, v, z_1, \dots, z_N)$  given by

$$\varrho(x, v, z_1, \dots, z_N) \propto \exp \left\{ -\Phi(x) - \frac{m}{2} v^2 - \frac{1}{2} \sum_{k=1}^N z_k^2 \right\}. \tag{2.3}$$

As discussed above, in order to study power law memory kernels we consider infinite sums of exponential functions. To this end, let  $\alpha, \beta > 0$  be given, and define constants  $c_k, \lambda_k, k = 1, 2, \dots$ , by

$$c_k = \frac{1}{k^{1+\alpha\beta}}, \quad \lambda_k = \frac{1}{k^\beta}. \tag{2.4}$$

Define the kernel  $K$  by

$$K(t) = \sum_{k \geq 1} c_k e^{-\lambda_k t}. \tag{2.5}$$

It follows (see example 3.3 of [1]) with this choice of constants  $c_k$  and  $\lambda_k$  that

$$K(t) \sim t^{-\alpha} \quad \text{as } t \rightarrow \infty, \tag{2.6}$$

thus giving the desired power law tail for the memory kernel  $K$ . This definition of the constants and exponents is not unique in yielding the power-law memory and subdiffusive behavior. For an alternate (and more efficient in the sense that few terms can yield a power law over similar time scales), see [14] and [16] for some discussion.

With this definition for  $K$ , we consider the following infinite-dimensional system of stochastic differential equations

$$\begin{aligned} dx(t) &= v(t) dt, \\ m dv(t) &= \left( -\gamma v(t) - \Phi'(x(t)) - \sum_{k \geq 1} \sqrt{c_k} z_k(t) \right) dt + \sqrt{2\gamma} dW_0(t), \\ dz_k(t) &= (-\lambda_k z_k(t) + \sqrt{c_k} v(t)) dt + \sqrt{2\lambda_k} dW_k(t), \quad k \geq 1, \end{aligned} \tag{2.7}$$

where the  $W_k$  are independent, standard Brownian motions. Throughout this work, we will assume that the potential  $\Phi$  satisfies the following growth and regularity conditions:

**Assumption 1.**  $\Phi \in C^\infty(\mathbb{R})$ ,  $\int_{\mathbb{R}} |\Phi'| e^{-\Phi} dx < \infty$  and there exists a constant  $b > 0$  such that for all  $x \in \mathbb{R}$

$$b(\Phi(x) + 1) \geq x^2.$$

A typical class of potentials  $\Phi$  that satisfies assumption 1 is the class of polynomials of even degree whose leading coefficient is positive.

**Remark 2.** The first and third parts of assumption 1 are quite standard, giving nominal regularity as well as assuring the potential grows at least as fast as a quadratic at infinity. The second condition is also a nominal growth condition on the derivative of  $\Phi$  and will be used in section 4 to check that our candidate invariant measure is indeed invariant.

In order to define a phase space for the infinite-dimensional process

$$X(t) = (x(t), v(t), z_1(t), z_2(t), \dots),$$

we will make use of the Hilbert space  $\mathcal{H}_{-s}$ ,  $s \in \mathbb{R}$ , equipped with the inner product  $\langle \cdot, \cdot \rangle_{-s}$ ,

$$\mathcal{H}_{-s} = \left\{ X = (x, v, z_1, z_2, \dots) : x^2 + v^2 + \sum_{k \geq 1} k^{-2s} z_k^2 < \infty \right\}, \tag{2.8}$$

and

$$\langle X, \tilde{X} \rangle_{-s} = x\tilde{x} + v\tilde{v} + \sum_{k \geq 1} k^{-2s} z_k \tilde{z}_k. \tag{2.9}$$

We denote by  $\| \cdot \|_{-s}$  the norm in  $\mathcal{H}_{-s}$  given by

$$\|X\|_{-s}^2 = x^2 + v^2 + \sum_{k \geq 1} k^{-2s} z_k^2. \tag{2.10}$$

The canonical basis  $\mathcal{D} = \{e_x, e_v, e_1, e_2, \dots\}$  in  $\mathcal{H}_{-s}$  is then given by

$$\begin{aligned} e_x &= (1, 0, 0, 0, \dots), \\ e_v &= (0, 1, 0, 0, \dots), \\ e_k &= (0, 0, 0, \dots, k^s, 0, \dots), \quad k \geq 1. \end{aligned} \tag{2.11}$$

From now on, for simplicity, we omit the subscript  $-s$  in the inner product  $\langle \cdot, \cdot \rangle$ . In view of (2.11), for  $X = (x, v, z_1, z_2, \dots)$ ,  $X$  we may write

$$X = \langle X, e_x \rangle e_x + \langle X, e_v \rangle e_v + \sum_{k \geq 1} \langle X, e_k \rangle e_k = x e_x + v e_v + \sum_{k \geq 1} k^{-s} z_k e_k. \tag{2.12}$$

Next, we collect several formulas involving Fréchet derivatives that will be useful later, especially in section 4. For  $\psi : \mathcal{H}_{-s} \rightarrow \mathbb{R}$ , let  $D\psi : \mathcal{H}_{-s} \rightarrow \mathcal{H}_{-s}$  be the first Fréchet derivative, if it exists. Then, the derivative of  $\psi$  in the direction of  $e \in \mathcal{D}$  is given by

$$\langle D\psi(X), e \rangle = \lim_{\varepsilon \rightarrow 0} \frac{\psi(X + \varepsilon e) - \psi(X)}{\varepsilon} = \frac{\partial \psi}{\partial \langle X, e \rangle}(X).$$

In view of representation (2.12), substituting  $e$  with  $e_x, e_v, e_k, k \geq 1$  in the above formula gives

$$\langle D\psi(X), e_x \rangle = \frac{\partial \psi}{\partial x}(X), \quad \langle D\psi(X), e_v \rangle = \frac{\partial \psi}{\partial v}(X), \tag{2.13}$$

$$\text{and } \langle D\psi(X), e_k \rangle = k^s \frac{\partial \psi(X)}{\partial z_k}$$

Similarly, if  $\psi$  is twice Fréchet differentiable, let  $D^2\psi : \mathcal{H}_{-s} \rightarrow L(\mathcal{H}_{-s}, \mathcal{H}_{-s})$  be the second Fréchet derivative, where  $L(\mathcal{H}_{-s}, \mathcal{H}_{-s})$  denotes the space of linear bounded maps from  $\mathcal{H}_{-s}$  to itself. Then, for  $e \in \mathcal{D}$ , we have

$$\langle D^2\psi(X)(e), e \rangle = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \langle D\psi(X + \varepsilon e) - D\psi(X), e \rangle = \frac{\partial^2 \psi}{\partial \langle X, e \rangle^2}(X).$$

Thus

$$\langle D^2\psi(X)(e_x), e_x \rangle = \frac{\partial^2 \psi}{\partial x^2}(X), \quad \langle D^2\psi(X)(e_v), e_v \rangle = \frac{\partial^2 \psi}{\partial v^2}(X), \quad (2.14)$$

$$\text{and } \langle D^2\psi(X)(e_k), e_k \rangle = k^{2s} \frac{\partial^2 \psi}{\partial z_k^2}(X).$$

Throughout, unless otherwise stated, we make the following assumptions about kernel parameters  $\alpha, \beta$  as in (2.4) and (2.5) and the phase space regularity parameter  $s$ .

**Assumption 3.** Let  $\alpha, \beta > 0$  be as in (2.4) and  $s$  as in (2.8). We assume that they satisfy either the asymptotically diffusive condition

$$(D) \alpha > 1, \beta > \frac{1}{\alpha - 1} \text{ and } \frac{1}{2} < s < \frac{(\alpha - 1)\beta}{2};$$

or the critical condition

$$(C) \alpha = 1, \beta > 1 \text{ and } \frac{1}{2} < s < \frac{\beta}{2};$$

or the asymptotically subdiffusive condition

$$(SD) 0 < \alpha < 1, \beta > \frac{1}{\alpha} \text{ and } \frac{1}{2} < s < \frac{\alpha\beta}{2}.$$

**Remark 4.** The assumption above really concerns the parameters  $\alpha, \beta$  only. Indeed, the particular choice of  $s$  in either part is the natural phase space range for the process for those particular choices of  $\alpha, \beta$ . It is also worth remarking that, so long as  $\beta > 0$  is large enough, the above simply splits the dynamics in the diffusive ( $\alpha > 1$ ) and the other two ( $0 < \alpha < 1$ , and  $\alpha = 1$ ) regimes.

**Remark 5.** In our context, to relate the infinite-dimensional system (2.7) to the original equation (1.4), the initial data  $z_k(0)$  for (2.7) is necessarily i.i.d. with  $\mathcal{N}(0, 1)$  distribution. To ensure consistency, we want  $(z_1(0), z_2(0), \dots)$  to live in  $\mathcal{H}_{-s}$  almost surely. This means that we must restrict to the phase space  $\mathcal{H}_{-s}$  for  $s > 1/2$ . It turns out that this constraint on  $s$  is also required to compute the density of the invariant measure, see section 4.

Fixing a stochastic basis  $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W)$  satisfying the usual conditions, see [38], where  $W$  is the cylindrical Wiener process defined later in (3.3), we state the following result giving existence and uniqueness of solutions of (2.7).

**Proposition 6.** Suppose that  $\Phi$  satisfies assumption 1 and the constants  $\alpha, \beta, s$  satisfy assumption 3. Then for all initial conditions  $X_0 \in \mathcal{H}_{-s}$ , there exists a unique pathwise solution  $X(\cdot, X_0) : \Omega \times [0, \infty) \rightarrow \mathcal{H}_{-s}$  of (2.7) in the following sense:  $X(\cdot, X_0)$  is  $\mathcal{F}_t$ -adapted,

$X(\cdot, X_0) \in C([0, \infty); \mathcal{H}_{-s})$  almost surely and that if  $\tilde{X}(\cdot, X_0)$  is another solution then for every  $T \geq 0$ ,

$$\mathbb{P} \left\{ \forall t \in [0, T], X(t, X_0) = \tilde{X}(t, X_0) \right\} = 1.$$

Moreover, for every  $X_0 \in \mathcal{H}_{-s}$  and  $T \geq 0$ , there exists  $C(T, X_0) > 0$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \|X(t)\|_{-s}^2 \right] \leq C(T, X_0), \tag{2.15}$$

and that for any bounded set  $B \subset \mathcal{H}_{-s}$ , we have

$$\sup_{X_0 \in B} C(T, X_0) < \infty. \tag{2.16}$$

The proof of proposition 6 will be carried out in section 3.

Our candidate stationary measure for the system (2.7) is an infinite-dimensional analogue of the one defined in (2.3). To write it down, let  $\mu_x, \mu_v$  and  $\nu$  denote the probability measures on  $\mathbb{R}$  defined by

$$\mu_x(\mathrm{d}y) = \frac{1}{\int_{\mathbb{R}} e^{-\Phi(z)} \mathrm{d}z} e^{-\Phi(y)} \mathrm{d}y, \quad \mu_v(\mathrm{d}y) = \frac{\sqrt{m}}{\sqrt{2\pi}} e^{-\frac{my^2}{2}} \mathrm{d}y, \tag{2.17}$$

$$\text{and } \nu(\mathrm{d}y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \mathrm{d}y.$$

Note that  $\mu_x$  is indeed a probability measure on  $\mathbb{R}$  by assumption 1. We denote by  $\mu$  the product probability measure on  $\mathbb{R}^\infty$  given by

$$\mu = \mu_x \times \mu_v \times \prod_{k \geq 1} \nu. \tag{2.18}$$

Observe that since  $s > 1/2$  by way of assumption 3

$$\int_{\mathbb{R}^\infty} \|X\|_{-s}^2 \mu(\mathrm{d}X) = \int_{\mathbb{R}^\infty} x^2 + v^2 + \sum_{k \geq 1} k^{-2s} z_k^2 \mu(\mathrm{d}X) < \infty. \tag{2.19}$$

Thus the restriction of  $\mu$  to  $\mathcal{H}_{-s}$  is a probability measure as the above calculation shows that  $\|X\|_{-s} < \infty$   $\mu$ -almost surely.

Let  $X(t) = (x(t), v(t), z_1(t), z_2(t), \dots)$  be the solution of (2.7) and define the operator  $\mathcal{P}(t) : \mathcal{B}_b(\mathcal{H}_{-s}) \rightarrow \mathcal{B}_b(\mathcal{H}_{-s})$  by

$$\mathcal{P}(t)\varphi(X_0) = \mathbb{E}[\varphi(X(t, X_0))]. \tag{2.20}$$

Here  $\mathcal{B}_b(\mathcal{H}_{-s})$  is the space of bounded Borel measurable  $\varphi : \mathcal{H}_{-s} \rightarrow \mathbb{R}$  and  $C_b(\mathcal{H}_{-s})$  is the space of bounded continuous  $\varphi : \mathcal{H}_{-s} \rightarrow \mathbb{R}$ . It will be shown later in proposition 15 that  $\{\mathcal{P}(t)\}_{t \geq 0}$  is a Feller Markov semigroup on  $C_b(\mathcal{H}_{-s})$ . We recall that a finite measure  $\xi$  on  $\mathcal{H}_{-s}$  is invariant for  $\{\mathcal{P}(t)\}_{t \geq 0}$  if for every  $\varphi \in C_b(\mathcal{H}_{-s})$  and  $t > 0$

$$\int_{\mathcal{H}_{-s}} \mathcal{P}(t)\varphi(X)\xi(\mathrm{d}X) = \int_{\mathcal{H}_{-s}} \varphi(X)\xi(\mathrm{d}X).$$

We have the following result.

**Theorem 7.** *Suppose that assumptions 1 and 3 are satisfied. Then  $\mu$  defined by (2.18) is an invariant probability measure for the Markov semigroup  $\{\mathcal{P}(t)\}_{t \geq 0}$  on  $C_b(\mathcal{H}_{-s})$  defined by (2.7).*

The idea behind the proof of theorem 7 is simple but the details are non-trivial. This is because one tries to ‘integrate by parts’ in infinite-dimensions aiming to show that  $\mathcal{L}^* \mu = 0$  where  $\mathcal{L}^*$  is some very formal adjoint of the Markov generator  $\mathcal{L}$ . The way that we circumnavigate this is by showing  $\mu$  is ‘approximately invariant’ for a sequence  $X^R(t)$ ,  $R \geq 1$ , of processes which approximate  $X(t)$  as  $R \rightarrow \infty$ . This turns out to be enough to show that  $\mu$  is invariant for the original process. Finally, our last result concerns unique ergodicity in the diffusive regime. To state it, we impose the following additional condition on the potential  $\Phi(x)$ .

**Assumption 8.** *There exist a function  $f(x) : \mathbb{R} \rightarrow \mathbb{R}^+$  that is bounded on bounded sets and a positive number  $q$  such that for all  $x, y \in \mathbb{R}$ ,*

$$|\Phi'(x) - \Phi'(y)| \leq |x - y| (f(x - y) + \Phi(x)^q). \tag{2.21}$$

**Remark 9.** Assumption 8 is essentially a growth bound on the second derivative of  $\Phi$ . In particular, if we assume further that there exist  $c, q_1, q_2 > 0$  such that for all  $x, y \in \mathbb{R}$  and  $t \in [0, 1]$ ,

$$\begin{aligned} |\Phi''(x)| &\leq c(\Phi(x)^{q_1} + 1), \\ \text{and } \Phi((1 - t)x + ty) &\leq c(\Phi(x)^{q_2} + \Phi(|x - y|)^{q_2} + 1), \end{aligned} \tag{2.22}$$

then  $\Phi$  satisfies assumption 8. From this observation, it is a short exercise to see that Condition 2.22 includes not only the class of non-negative polynomials of even degree, but also functions that even grow faster than a polynomial at infinity, e.g.  $\Phi(x) = e^{x^2}$ .

**Theorem 10.** *Suppose assumptions 1, 8 and Condition (D) of assumption 3 are satisfied. Then  $\mu$  is the unique invariant measure for the Markov process defined by (2.7).*

The proof of theorem 10, which will be given in section 5, uses an asymptotic coupling argument following the ideas and results in the works of [12, 26, 44]. For a similar rigorous study of finite-dimensional Langevin Equation, we refer the reader to [52, 60].

### 3. Well-posedness

For notational convenience, throughout we write (2.7) more compactly as the following semi-linear stochastic evolution equation

$$dX(t) = (LX(t) + F(X(t))) dt + B dW(t), \quad X(0) = X_0 \in \mathcal{H}_{-s}, \tag{3.1}$$

where  $L$  is a linear map given by

$$LX = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & -\frac{\gamma}{m} & -\frac{\sqrt{c_1}}{m} & -\frac{\sqrt{c_2}}{m} & \dots \\ 0 & \sqrt{c_1} & -\lambda_1 & 0 & \dots \\ 0 & \sqrt{c_2} & 0 & -\lambda_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} x \\ v \\ z_1 \\ z_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} v \\ -\frac{\gamma}{m}v - \frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} z_k \\ -\lambda_1 z_1 + \sqrt{c_1} v \\ -\lambda_2 z_2 + \sqrt{c_2} v \\ \vdots \end{pmatrix}, \tag{3.2}$$

where  $c_k, \lambda_k$  are as in (2.4) and the force  $F$  is defined as

$$F(X) = (0, -\Phi'(x)/m, 0, 0, \dots)^T.$$

Regarding the stochastic term  $B \, dW$ , we may formally write

$$B \, dW(t) = \begin{pmatrix} 0 & 0 & 0 & \dots \\ \frac{\sqrt{2\gamma}}{m} & 0 & 0 & \dots \\ 0 & \sqrt{2\lambda_1} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} dW_0 \\ dW_1 \\ dW_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\sqrt{2\gamma}}{m} dW_0(t) \\ \sqrt{2\lambda_1} dW_1(t) \\ \sqrt{2\lambda_2} dW_2(t) \\ \vdots \end{pmatrix}.$$

Following the formulation in, for example, [7], fix an auxiliary Hilbert space  $\mathcal{W}$  and pick a complete orthonormal basis  $\{e_k^{\mathcal{W}}\}_{k \geq 0}$ . The cylindrical Wiener process  $W(t)$  on the Hilbert space  $\mathcal{W}$  is then defined as

$$W(t) = W_0(t)e_0^{\mathcal{W}} + W_1(t)e_1^{\mathcal{W}} + W_2(t)e_2^{\mathcal{W}} + \dots, \tag{3.3}$$

where the sequence  $\{W_k(t)\}_{k \geq 0}$  are independent one-dimensional Brownian Motions. We can then define  $B : \mathcal{W} \rightarrow \mathcal{H}_{-s}$  by its action

$$Be_0^{\mathcal{W}} = \frac{\sqrt{2\gamma}}{m} e_v, \quad \text{and} \quad Be_k^{\mathcal{W}} = \sqrt{2\lambda_k} k^{-s} e_k, \quad k \geq 1, \tag{3.4}$$

where  $\{e_x, e_v, e_1, e_2, \dots\}$  is the canonical basis of  $\mathcal{H}_{-s}$ , see (2.11). In view of (3.4), we have

$$BB^* e_x = 0, \quad BB^* e_v = \frac{2\gamma}{m^2} e_v, \quad \text{and} \quad BB^* e_k = 2\lambda_k k^{-2s} e_k, \quad k \geq 1. \tag{3.5}$$

In order to prove well-posedness of equation (3.1), we need the following basic fact.

**Proposition 11.** *Suppose that  $\alpha, \beta$  as in (2.4) and  $s$  as in (2.8) satisfy  $0 \leq 2s < \alpha\beta$ . Then,  $L : \mathcal{H}_{-s} \rightarrow \mathcal{H}_{-s}$  defined in (3.2) is a bounded linear operator.*

**Proof.** Recalling (2.4), (2.10), (3.2) and invoking Cauchy–Schwarz inequality, we estimate

$$\begin{aligned} \|LX\|_{-s}^2 &= v^2 + \left( -\frac{\gamma}{m} v - \frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} z_k \right)^2 + \sum_{k \geq 1} k^{-2s} (-\lambda_k z_k + \sqrt{c_k} v)^2 \\ &\leq v^2 + \frac{2\gamma^2}{m^2} v^2 + \frac{2}{m^2} \left( \sum_{k \geq 1} \sqrt{c_k} z_k \right)^2 + \sum_{k \geq 1} 2\lambda_k^2 k^{-2s} z_k^2 + \sum_{k \geq 1} 2k^{-2s} c_k v^2 \\ &\leq \left( 1 + \frac{2\gamma^2}{m^2} + 2 \sum_{k \geq 1} k^{-2s} c_k \right) v^2 + \frac{2}{m^2} \sum_{k \geq 1} c_k k^{2s} \sum_{k \geq 1} k^{-2s} z_k^2 + 2 \sum_{k \geq 1} \lambda_k^2 k^{-2s} z_k^2 \\ &= \left( 1 + \frac{2\gamma^2}{m^2} + 2 \sum_{k \geq 1} k^{-2s} k^{-(1+\alpha\beta)} \right) v^2 + \frac{2}{m^2} \sum_{k \geq 1} k^{-(1+\alpha\beta)} k^{2s} \sum_{k \geq 1} k^{-2s} z_k^2 \\ &\quad + 2 \sum_{k \geq 1} k^{-\beta} k^{-2s} z_k^2 \\ &\leq \left( 1 + \frac{2\gamma^2}{m^2} + 2 \sum_{k \geq 1} k^{-(1+\alpha\beta+2s)} + \frac{2}{m^2} \sum_{k \geq 1} k^{-(1+\alpha\beta-2s)} + 2 \right) \|X\|_{-s}^2. \end{aligned}$$

Since  $\sum_{k \geq 1} k^{-(1+\alpha\beta-2s)}$  converges whenever  $\alpha\beta > 2s$ , the desired result follows. □

The proof of proposition 6 follows a Lyapunov argument which we now explain. Let  $\theta_R \in C^\infty(\mathbb{R}; [0, 1])$  satisfy

$$\theta_R(x) = \begin{cases} 1 & \text{if } |x| \leq R \\ 0 & \text{if } |x| \geq R + 1 \end{cases} \tag{3.6}$$

and consider ‘cutoff’ equation corresponding to (3.1)

$$dX^R(t) = [LX^R(t) + F(X^R(t))\theta_R(x^R(t))] dt + B dW(t), \quad X^R(0) = X_0. \tag{3.7}$$

For each  $R > 0$ , it is not hard to prove that the global (in time) solution  $X^R$  exists and is unique, giving local (up until the time of explosion) pathwise existence and uniqueness of (3.1). Then, using a Lyapunov function  $\Psi(X)$  that dominates the norm of  $X$  in  $\mathcal{H}_{-s}$ , we show a global bound on these solutions that does not depend on  $R$ , thereby obtaining global solutions of (3.1).

To prove proposition 6, we begin with the following proposition.

**Proposition 12 (Local Existence).** *Suppose that  $\alpha, \beta$  as in (2.4) and  $s$  as in (2.8) satisfy  $1 < 2s < \alpha\beta$ . Let  $X_0 \in \mathcal{H}_{-s}$  be given. For each  $R > 0$ , there exists a unique  $X^R(t) \in L^2(\Omega, C([0, \infty); \mathcal{H}_{-s}))$  satisfying (3.7).*

**Remark 13.** We see that the conditions stated in assumption 3 meet the hypothesis  $1 < 2s < \alpha\beta$  of proposition 12.

**Proof of proposition 12.** The linear map  $L$  is bounded by lemma 11 and the nonlinear term in (3.7) is globally Lipschitz in  $\|\cdot\|_{-s}$  by construction. Moreover, the additive noise term lives in  $\mathcal{H}_{-s}$  almost surely since

$$\begin{aligned} \mathbb{E} \left\| \int_0^T B dW(t) \right\|_{-s}^2 &= \mathbb{E} \left| \int_0^T \frac{\sqrt{2\gamma}}{m} dW_0(t) \right|^2 + \sum_{k \geq 1} k^{-2s} \mathbb{E} \left| \int_0^T \sqrt{2\lambda_k} dW_k(t) \right|^2 \\ &= \frac{2\gamma T}{m^2} + 2T \sum_{k \geq 1} k^{-2s} \lambda_k < \infty. \end{aligned}$$

The corresponding solution hence exists and is unique by a standard Banach fixed point argument. □

Next, inspired by [52, 61], we introduce a Lyapunov function  $\Psi : \mathcal{H}_{-s} \rightarrow [0, \infty)$  given by

$$\Psi(X) := \frac{1}{m} \Phi(x) + \frac{1}{2} v^2 + \frac{1}{2} \sum_{k \geq 1} k^{-2s} z_k^2. \tag{3.8}$$

Define  $\mathcal{L} : C^2(\mathcal{H}_{-s}) \rightarrow \mathbb{R}$  to be the operator given by

$$\mathcal{L}\varphi(X) := \langle D\varphi(X), LX + F(X) \rangle + \frac{1}{2} \text{Tr}(D^2\varphi BB^*).$$

In view of (2.13), (2.14), (3.2) and (3.5),  $\mathcal{L}$  can be explicitly written as

$$\begin{aligned} \mathcal{L}\varphi(X) &= v \frac{\partial \varphi(X)}{\partial x} + \left( -\frac{\gamma}{m}v - \frac{1}{m}\Phi'(x) - \frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} z_k \right) \frac{\partial \varphi(X)}{\partial v} \\ &+ \sum_{k \geq 1} (-\lambda_k z_k + \sqrt{c_k} v) \frac{\partial \varphi(X)}{\partial z_k} + \frac{\gamma}{m^2} \frac{\partial^2 \varphi(X)}{\partial v^2} + \sum_{k \geq 1} \lambda_k \frac{\partial^2 \varphi(X)}{\partial z_k^2} \end{aligned} \tag{3.9}$$

where  $\varphi \in C^2(\mathcal{H}_{-s})$ . Note that, once we establish proposition 6,  $\mathcal{L}$  is in fact the infinitesimal generator of the Markov semigroup  $\{\mathcal{P}(t)\}_{t \geq 0}$  associated with (2.7). We assert the following proposition.

**Proposition 14 (Global bound).** *Suppose that assumptions 1 and 3 are satisfied. Let  $\Psi(X)$  be defined in (3.8) and  $\mathcal{L}$  be the operator as in (3.9). Then, for every  $X \in \mathcal{H}_{-s}$ ,*

$$\mathcal{L}\Psi(X) \leq a_1 \Psi(X) + a_2, \tag{3.10}$$

where  $a_1, a_2$  are finite constants that can be explicitly given as

$$a_1 = \max \left\{ 1 + \frac{1}{m}, \frac{1}{m} \sum_{k \geq 1} c_k k^{2s} + \sum_{k \geq 1} c_k k^{-2s} \right\} \text{ and } a_2 = \frac{\gamma}{m^2} + \sum_{k \geq 1} k^{-2s} \lambda_k.$$

**Proof.** Applying assumption 3, first note that  $a_1$  and  $a_2$  are both finite since

$$\sum_{k \geq 1} c_k k^{2s} = \sum_{k \geq 1} \frac{1}{k^{1+\alpha\beta-2s}} < \infty \quad \text{and} \quad \sum_{k \geq 1} k^{-2s} \lambda_k = \sum_{k \geq 1} \frac{1}{k^{\beta+2s}} < \infty.$$

We now apply  $\mathcal{L}$  to  $\Psi$  to see that

$$\begin{aligned} \mathcal{L}\Psi(X) &= -\frac{\gamma}{m}v^2 - \sum_{k \geq 1} \lambda_k k^{-2s} z_k^2 \\ &- \frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} z_k v + \sum_{k \geq 1} \sqrt{c_k} k^{-2s} z_k v + \frac{\gamma}{m^2} + \sum_{k \geq 1} k^{-2s} \lambda_k. \end{aligned} \tag{3.11}$$

The cross terms between  $z_k$  and  $v$  can be bounded using Hölder’s inequality as follows:

$$\begin{aligned} &-\frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} z_k v + \sum_{k \geq 1} \sqrt{c_k} k^{-2s} z_k v \\ &\leq \frac{1}{2m} \left( \sum_{k \geq 1} c_k k^{2s} \sum_{k \geq 1} k^{-2s} z_k^2 + v^2 \right) + \frac{1}{2} \left( \sum_{k \geq 1} c_k k^{-2s} \sum_{k \geq 1} k^{-2s} z_k^2 + v^2 \right) \\ &= \frac{1}{2} \left( 1 + \frac{1}{m} \right) v^2 + \frac{1}{2} \left( \frac{1}{m} \sum_{k \geq 1} c_k k^{2s} + \sum_{k \geq 1} c_k k^{-2s} \right) \sum_{k \geq 1} k^{-2s} z_k^2 \\ &\leq a_1 \Psi(X), \end{aligned} \tag{3.12}$$

where  $a_1 > 0$  is as in the statement of the result. We finally combine (3.11) with (3.12) to obtain (3.10). □

We are now ready to prove the main existence and uniqueness result for equation (3.1). The argument is classical and can be found in literature, e.g. [3, 13, 35].

**Proof of proposition 6.** For every  $R > 0$ , let  $X^R(t)$  be the unique solution of the cutoff system (3.7) given to us by proposition 12. Define the stopping time

$$\tau_R = \inf \{t > 0 : \|X(t)\|_{-s} > R\}.$$

Note that, for all times  $t < \tau_R$ ,  $X^R$  solves (3.1). Consequently, the solution (3.1) exists and is unique up until the *time of explosion*  $\tau_\infty = \lim_{R \rightarrow \infty} \tau_R$ , which is possibly finite on a set of positive probability. We show using the Lyapunov function  $\Psi$  above, see (3.8) that  $\tau_\infty = \infty$  a.s.

By Ito’s Formula we have that

$$\begin{aligned} d\Psi(X(t \wedge \tau_R)) &= \mathcal{L}\Psi(X(t \wedge \tau_R))dt + \frac{\sqrt{2\gamma}}{m}v(t \wedge \tau_R)dW_0(t) \\ &\quad + \sum_{k \geq 1} \sqrt{2\lambda_k}k^{-2s}z_k(t \wedge \tau_R)dW_k(t), \end{aligned} \tag{3.13}$$

where  $\mathcal{L}$  is the operator defined in (3.9). We then infer the following bound

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \Psi(X(t \wedge \tau_R)) \right] &\leq \Psi(X_0) + \mathbb{E} \sup_{0 \leq t \leq T} \int_0^t \mathcal{L}\Psi(X(r \wedge \tau_R))dr \\ &\quad + \mathbb{E} \sup_{0 \leq t \leq T} \left| \int_0^t \frac{\sqrt{2\gamma}}{m}v(r \wedge \tau_R)dW_0(r) + \sum_{k \geq 1} \sqrt{2\lambda_k}k^{-2s}z_k(t \wedge \tau_R)dW_k(r)dr \right|. \end{aligned} \tag{3.14}$$

Applying proposition 14 on  $\mathcal{L}\Psi(X(r \wedge \tau_R))$  and the Burkholder–Davis–Gundy inequality on the martingale term on the above RHS yields

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \Psi(X(t \wedge \tau_R)) \right] &\leq \Psi(X_0) + a_2T + a_1\mathbb{E} \int_0^T \sup_{0 \leq r \leq t} \Psi(X(r \wedge \tau_R))dt \\ &\quad + c \left[ \mathbb{E} \int_0^T \frac{2\gamma}{m^2}v(t \wedge \tau_R)^2 + \sum_{k \geq 1} 2\lambda_k k^{-4s}z_k(t \wedge \tau_R)^2 dt \right]^{1/2}, \end{aligned} \tag{3.15}$$

where  $a_1, a_2$  are as in (3.10) and  $c > 0$  is the constant from Burkholder–Davis–Gundy’s inequality and is independent of  $R, T, X_0$ . We observe now that the last integrand in (3.15) is dominated by  $\Psi(X(t \wedge \tau_R))$ . We thus infer that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \Psi(X(t \wedge \tau_R)) \right] \leq \Psi(X_0) + c_1 + c_2 \int_0^T \mathbb{E} \left[ \sup_{0 \leq r \leq t} \Psi(X(r \wedge \tau_R)) \right] dt, \tag{3.16}$$

where  $c_1, c_2 > 0$  are constants independent of  $R$  and  $X_0$ . Gronwall’s inequality then implies that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \Psi(X(t \wedge \tau_R)) \right] \leq (\Psi(X_0) + c_1) e^{c_2 T}. \tag{3.17}$$

Also note that there exists a constant  $c > 0$  such that for all  $X \in \mathcal{H}_{-s}$ ,  $c(\Psi(X) + 1) \geq \|X\|_{-s}^2$ . We thus infer the existence of a constant  $C(T, X_0) > 0$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \|X(t \wedge \tau_R)\|_{-s}^2 \right] \leq C(T, X_0). \tag{3.18}$$

Sending  $R$  to infinity in the above, we obtain by Fatou’s lemma

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \|X(t \wedge \tau_\infty)\|_{-s}^2 \right] \leq C(T, X_0). \tag{3.19}$$

Hence, this implies  $\mathbb{P}\{T < \tau_\infty\} = 1$  for any  $T > 0$ . By sending  $T$  to infinity, we see that  $\mathbb{P}\{\tau_\infty = \infty\} = 1$ , which implies the well-posedness of the global solution for any fixed  $X_0 \in \mathcal{H}_{-s}$ .

We finally note that  $C(T, X_0)$  is actually dominated by  $\Psi(X_0)e^{c_2 T}$  following from estimates (3.17) and (3.18). It is also clear that  $\Psi(X)$  is bounded on bounded sets in  $\mathcal{H}_{-s}$ . We therefore obtain the bound in (2.16), which concludes the proof.  $\square$

In addition to pathwise existence and uniqueness of the solution of equation (3.1), we will need the following basic properties of the Markov semigroup  $\mathcal{P}(t) : \mathcal{B}_b(\mathcal{H}_{-s}) \rightarrow \mathcal{B}_b(\mathcal{H}_{-s})$ . We recall that  $\mathcal{P}(t)$  defined as in (2.20) possesses the Markov property; namely, for every  $X \in \mathcal{H}_{-s}$ ,  $\varphi \in C_b(\mathcal{H}_{-s})$ ,  $t, r \geq 0$ ,

$$\mathcal{P}(t+r)\varphi(X) = \mathcal{P}(t)(\mathcal{P}(r)\varphi)(X).$$

**Proposition 15.** *Under the Hypothesis of proposition 6, let  $X(t)$  be the unique strong solution of (2.7) and  $\mathcal{P}(t)$  be the corresponding Markov semigroup. We have the following:*

(a) Whenever  $X_k \rightarrow X_0$  in  $\mathcal{H}_{-s}$

$$\lim_{k \rightarrow \infty} \mathbb{E} \|X(t, X_k) - X(t, X_0)\|_{-s} = 0. \tag{3.20}$$

(b)  $\mathcal{P}(t)$  has the Feller property:  $\mathcal{P}(t)\varphi \in C_b(\mathcal{H}_{-s})$  whenever  $\varphi \in C_b(\mathcal{H}_{-s})$ .

**Proof.** (a) For notational simplicity, throughout this proof, we shall omit the subscript  $-s$  in the norm  $\|\cdot\|_{-s}$ . Denote by  $X_{(k)}(t)$  the solution of (2.7) with initial data  $X_{(k)}(0) = X_k$ . Fixing  $R > 0$  to be chosen later, define the stopping time

$$\tau_R^k = \inf_{t \geq 0} \{ \|X_{(k)}(t)\| + \|X_{(0)}(t)\| \geq R \},$$

and observe that by Chebychev’s inequality,

$$\begin{aligned} \mathbb{P} \{ \tau_R^k \leq t \} &= \mathbb{P} \left\{ \sup_{0 \leq \ell \leq t} \|X_{(k)}(\ell)\| + \|X_{(0)}(\ell)\| \geq R \right\} \\ &\leq \frac{\mathbb{E} [\sup_{0 \leq \ell \leq t} \|X_{(k)}(\ell)\| + \|X_{(0)}(\ell)\|]}{R} \\ &\leq \frac{\sqrt{C(t, X_k)} + \sqrt{C(t, X_0)}}{R}, \end{aligned}$$

where we used (2.15) in the last inequality. It follows from (2.16) that

$$\sup_k \mathbb{P} \{ \tau_R^k \leq t \} \leq \frac{C(t)}{R}, \tag{3.21}$$

for a finite constant  $C(t) > 0$  independent of  $R$ . Next, let  $X_{(k)}^R$  and  $X_{(0)}^R$  be the local solutions of (3.7) from proposition 12. Since the drift term of (3.7) is Lipschitz, there exists a constant  $c(R, t) > 0$  such that (see theorem 9.1, [7]).

$$\mathbb{E} \|X_{(k)}^R(t) - X_{(0)}^R(t)\| \leq C(R, t) \|X_k - X_0\|. \tag{3.22}$$

Now we have a chain of implications

$$\begin{aligned} \mathbb{E} \|X_{(k)}(t) - X_{(0)}(t)\| &\leq \mathbb{E} \left[ (\|X_{(k)}(t)\| + \|X_{(0)}(t)\|) \mathbf{1}_{\{\tau_R^k \leq t\}} \right] + \mathbb{E} \left[ \|X_{(k)}(t) - X_{(0)}(t)\| \mathbf{1}_{\{\tau_R^k > t\}} \right] \\ &= \mathbb{E} \left[ (\|X_{(k)}(t)\| + \|X_{(0)}(t)\|) \mathbf{1}_{\{\tau_R^k \leq t\}} \right] + \mathbb{E} \left[ \|X_{(k)}^R(t) - X_{(0)}^R(t)\| \mathbf{1}_{\{\tau_R^k > t\}} \right] \\ &\leq \left( \mathbb{E} \left[ (\|X_{(k)}(t)\| + \|X_{(0)}(t)\|)^2 \right] \right)^{1/2} \left( \mathbb{P} \{ \tau_R^k \leq t \} \right)^{1/2} + \mathbb{E} \|X_{(k)}^R(t) - X_{(0)}^R(t)\| \\ &\leq \left( \mathbb{E} \left[ (\|X_{(k)}(t)\| + \|X_{(0)}(t)\|)^2 \right] \right)^{1/2} \left( \mathbb{P} \{ \tau_R^k \leq t \} \right)^{1/2} + C(R, t) \|X_k - X_0\| \\ &\leq \frac{C_1(t)}{R^{1/2}} + C(R, t) \|X_k - X_0\|, \end{aligned}$$

where note carefully that  $C_1(t)$  is independent of  $k$  and  $R$  and  $C(R, t)$  is independent of  $k$ . The above RHS now tends to zero by taking  $R$  sufficiently large first and then letting  $X_k$  sufficiently close to  $X_0$ . This establishes (a).

To prove (b), let  $X_k \rightarrow X_0$  and  $\varphi \in C_b(\mathcal{H}_{-s})$ , we have to show  $\mathcal{P}(t)\varphi(X_k) \rightarrow \mathcal{P}(t)\varphi(X_0)$ . It suffices to show that for every subsequence  $\{k_i\}$ , there exists a further subsequence  $\{k_{ij}\}$  such that  $\mathcal{P}(t)\varphi(X_{k_{ij}}) \rightarrow \mathcal{P}(t)\varphi(X_0)$ . In view of part (a), the sequence  $X_{(k_i)}(t)$  converges to  $X_{(0)}(t)$  in  $L^1(\Omega; \mathcal{H}_{-s})$ . We thus can extract a subsequence  $X_{(k_{ij})}(t)$  converging to  $X_{(0)}(t)$  a.s. Since  $\varphi$  is bounded, applying the dominated convergence theorem yields

$$\mathbb{E} [\varphi(X_{(k_{ij})}(t))] \rightarrow \mathbb{E} [\varphi(X_{(0)}(t))] \text{ as } j \rightarrow \infty,$$

which implies (b) and thus completes the proof. □

#### 4. Invariance of $\mu$

In this section, we show that  $\mu$  defined in (2.18) is invariant for the Markov semigroup  $\{\mathcal{P}(t)\}_{t \geq 0}$ . We first sketch briefly the structure of the proof before diving into details.

The goal is to show that for every  $\varphi \in C_b(\mathcal{H}_{-s})$  and  $t \geq 0$  we have

$$\int_{\mathcal{H}_{-s}} \mathcal{P}(t)\varphi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \varphi(X)\mu(dX). \tag{4.1}$$

Let  $C_b^2(\mathcal{H}_{-s})$  denote the space of real-valued functions on  $\mathcal{H}_{-s}$  that have bounded first and second Fréchet derivatives. Approximating  $\varphi$  by functions in  $C_b^2(\mathcal{H}_{-s})$  if necessary, it thus suffices to show that (4.1) holds for any  $\psi \in C_b^2(\mathcal{H}_{-s})$

$$\int_{\mathcal{H}_{-s}} \mathcal{P}(t)\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX). \tag{4.2}$$

In order to show (4.2), it is helpful to make use of the cutoff system (3.7) and the semigroup  $\mathcal{P}^R(t)$  where for  $R > 0$ ,  $\mathcal{P}^R(t)$  is defined analogously to (2.20) by replacing  $X(t)$  with  $X^R(t)$  solving (3.7). The advantage of using the cutoff systems is that, because they have globally Lipschitz coefficients, they immediately satisfy the Kolmogorov backward equation, see

theorem 9.23 of [7]. This fact we will need later in the proof of proposition 17. Specifically, we will prove that  $\mu$  is *almost invariant* for the cutoff semigroup  $\mathcal{P}^R(t)$ ; namely,

$$\int_{\mathcal{H}_{-s}} \mathcal{P}^R(t)\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX) + \varepsilon^R(\psi, t), \tag{4.3}$$

where  $\varepsilon^R(\psi, t)$  is a remainder term that (possibly) depends on  $\psi$  and  $t$ , and satisfies  $\varepsilon^R(\psi, t) \rightarrow 0$  as  $R \rightarrow \infty$ . We will see that this then implies the desired equality (4.2).

Before proving theorem 7, we collect several properties about Gaussian measures on  $\mathbb{R}$  which follow simply by using integration by parts. Let  $\mu_\nu, \nu$  be as in (2.17). Then, for every  $\varrho_1 \in C_b^2(\mathbb{R})$ , it holds that

$$\int_{\mathbb{R}} -y\varrho_1'(y) + \frac{1}{m}\varrho_1''(y)\mu_\nu(dy) = 0, \tag{4.4}$$

and

$$\int_{\mathbb{R}} -y\varrho_1'(y) + \varrho_1''(y)\nu(dy) = 0. \tag{4.5}$$

Also, for every  $\varrho_2 \in C_b^1(\mathbb{R}^2)$ , we have

$$\int_{\mathbb{R}^2} \left( \frac{1}{m}z\partial_y\varrho_2(y, z) - y\partial_z\varrho_2(y, z) \right) (\mu_\nu \times \nu) (dy, dz) = 0. \tag{4.6}$$

With these observations, we have the following result:

**Lemma 16.** *Given  $R > 0$ , let  $\mathcal{L}^R$  be the infinitesimal generator of the Markov semigroup  $\{\mathcal{P}^R(t)\}_{t \geq 0}$  associated with  $X^R$  solving (3.7) and let  $\mu$  be as in (2.18). Then, for every  $\psi \in C_b^2(\mathcal{H}_{-s})$ , we have the following equality*

$$\int_{\mathcal{H}_{-s}} \mathcal{L}^R\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} v\Phi'(x) (1 - \theta^R(x)) \psi(X)\mu(dX). \tag{4.7}$$

**Proof.** Similar to (3.9),  $\mathcal{L}^R\psi$  is given by

$$\begin{aligned} \mathcal{L}^R\psi(X) &= v\frac{\partial\psi(X)}{\partial x} + \left( -\frac{\gamma}{m}v - \frac{1}{m}\Phi'(x)\theta^R(x) - \frac{1}{m}\sum_{k \geq 1} \sqrt{c_k z_k} \right) \frac{\partial\psi(X)}{\partial v} \\ &\quad + \sum_{k \geq 1} (-\lambda_k z_k + \sqrt{c_k}v) \frac{\partial\psi(X)}{\partial z_k} + \frac{\gamma}{m^2} \frac{\partial^2\psi(X)}{\partial v^2} + \sum_{k \geq 1} \lambda_k \frac{\partial^2\psi(X)}{\partial z_k^2}. \end{aligned} \tag{4.8}$$

We integrate both sides against  $\mu$  in  $\mathcal{H}_{-s}$  and rearrange the above RHS appropriately to obtain

$$\begin{aligned}
 \int_{\mathcal{H}_{-s}} \mathcal{L}^R \psi(X) \mu(dX) &= \int_{\mathcal{H}_{-s}} \left[ v \frac{\partial \psi(X)}{\partial x} - \frac{1}{m} \Phi'(x) \theta^R(x) \frac{\partial \psi(X)}{\partial v} \right] \mu(dX) \\
 &\quad + \int_{\mathcal{H}_{-s}} \left[ -\frac{\gamma}{m} v \frac{\partial \psi(X)}{\partial v} + \frac{\gamma}{m^2} \frac{\partial^2 \psi(X)}{\partial v^2} \right] \mu(dX) \\
 &\quad + \sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left[ -\frac{\sqrt{c_k}}{m} z_k \frac{\partial \psi(X)}{\partial v} + \sqrt{c_k} v \frac{\partial \psi(X)}{\partial z_k} \right] \mu(dX) \\
 &\quad + \sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left[ -\lambda_k z_k \frac{\partial \psi(X)}{\partial z_k} + \lambda_k \frac{\partial^2 \psi(X)}{\partial z_k^2} \right] \mu(dX) \\
 &= I_{0,1} + I_{0,2} + \sum_{k \geq 1} I_{k,1} + \sum_{k \geq 1} I_{k,2}.
 \end{aligned} \tag{4.9}$$

At this point (4.9) is still formal. We need to show that  $\mathcal{L}^R \psi \in L^1(\mathcal{H}_{-s}, \mu)$  and that the above rearrangement is possible. To this end, we claim that the RHS after the first equality of (4.9) is absolutely convergent. Since  $\psi \in C_b^2(\mathcal{H}_{-s})$ , (2.13) and Parseval’s identity imply a bound on first partial derivatives

$$\begin{aligned}
 &\left| \frac{\partial \psi(X)}{\partial x} \right|^2 + \left| \frac{\partial \psi(X)}{\partial v} \right|^2 + \sum_{k \geq 1} k^{2s} \left| \frac{\partial \psi(X)}{\partial z_k} \right|^2 \\
 &= \langle D\psi(X), e_x \rangle^2 + \langle D\psi(X), e_v \rangle^2 + \sum_{k \geq 1} \langle D\psi(X), e_k \rangle^2 \\
 &= \|D\psi(X)\|_{-s}^2 \\
 &\leq \|D\psi\|_{\infty}^2,
 \end{aligned} \tag{4.10}$$

where for  $D\psi : \mathcal{H}_{-s} \rightarrow \mathcal{H}_{-s}$ ,  $\|D\psi\|_{\infty} = \sup_{Y \in \mathcal{H}_{-s}} \|D\psi(Y)\|_{-s}$ . Similarly, (2.14) implies bounds on second partial derivatives,

$$\frac{\partial^2 \psi(X)}{\partial x^2} = \langle D^2 \psi(X)(e_x), e_x \rangle \leq \|D^2 \psi\|_{\infty}, \quad \frac{\partial^2 \psi(X)}{\partial v^2} \leq \|D^2 \psi\|_{\infty}, \tag{4.11}$$

and

$$k^{2s} \frac{\partial^2 \psi(X)}{\partial z_k^2} \leq \|D^2 \psi\|_{\infty}, \tag{4.12}$$

where for  $D^2 \psi(X) : \mathcal{H}_{-s} \rightarrow L(\mathcal{H}_{-s}, \mathcal{H}_{-s})$ ,

$$\|D^2 \psi\|_{\infty} = \sup_{Y \in \mathcal{H}_{-s}} \|D^2 \psi(Y)\|_{L(\mathcal{H}_{-s}, \mathcal{H}_{-s})}.$$

In the RHS after the first equality of (4.9), the first four terms are bounded by, using (4.10), (4.11),

$$\begin{aligned} & \int_{\mathcal{H}_{-s}} \left| v \frac{\partial \psi(X)}{\partial x} \right| + \frac{1}{m} \left| \Phi'(x) \theta^R(x) \frac{\partial \psi(X)}{\partial v} \right| \\ & \quad + \frac{\gamma}{m} \left| v \frac{\partial \psi(X)}{\partial v} \right| + \frac{\gamma}{m^2} \left| \frac{\partial^2 \psi(X)}{\partial v^2} \right| \mu(dX) \\ & \leq \|D\psi\|_\infty \left[ \left(1 + \frac{\gamma}{m}\right) \int_{\mathbb{R}} |v| \mu_v(dv) + \frac{1}{m} \int_{\mathbb{R}} |\Phi'(x)| \theta_R(x) e^{-\Phi(x)} dx \right] + \frac{\gamma}{m^2} \|D^2\psi\|_\infty, \end{aligned} \tag{4.13}$$

which is finite, by the definition of  $\mu_v$  from (2.17) and the fact that  $\theta_r$ , as in (3.6) has compact support. For the first sum on the third line of (4.9), we estimate as follows.

$$\begin{aligned} & \sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left| \frac{\sqrt{c_k}}{m} z_k \frac{\partial \psi(X)}{\partial v} \right| \mu(dX) \\ & \leq \frac{\|D\psi\|_\infty}{m} \left( \sum_{k \geq 1} c_k k^{2s} \right)^{1/2} \int_{\mathcal{H}_{-s}} \left( \sum_{k \geq 1} k^{-2s} z_k^2 \right)^{1/2} \mu(dX) \\ & \leq \frac{\|D\psi\|_\infty}{m} \left( \sum_{k \geq 1} c_k k^{2s} \right)^{1/2} \int_{\mathcal{H}_{-s}} \|X\|_{-s} \mu(dX) < \infty, \end{aligned} \tag{4.14}$$

since by assumption 3,  $\sum_{k \geq 1} c_k k^{2s}$  is finite and so is  $\int_{\mathcal{H}_{-s}} \|X\|_{-s} \mu(dX)$ , by the definition of  $\mu$ , see (2.19). Similarly, for the second sum on the third line of (4.9), using (4.10) again, we infer

$$\begin{aligned} & \sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left| \sqrt{c_k} v \frac{\partial \psi(X)}{\partial z_k} \right| \mu(dX) \\ & \leq \int_{\mathcal{H}_{-s}} \left( \sum_{k \geq 1} \frac{c_k}{k^{2s}} \right)^{1/2} \left( \sum_{k \geq 1} k^{2s} \left| \frac{\partial \psi(X)}{\partial z_k} \right|^2 \right)^{1/2} |v| \mu(dX) \\ & \leq \|D\psi\|_\infty \left( \sum_{k \geq 1} \frac{c_k}{k^{2s}} \right)^{1/2} \int_{\mathcal{H}_{-s}} |v| \mu(dX) < \infty. \end{aligned} \tag{4.15}$$

For the first sum on the fourth line of (4.9), similar to (4.15), we invoke (4.10) again to see that

$$\sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left| \lambda_k z_k \frac{\partial \psi(X)}{\partial z_k} \right| \mu(dX) \leq \|D\psi\|_\infty \int_{\mathcal{H}_{-s}} \|X\|_{-s} \mu(dX) < \infty. \tag{4.16}$$

Lastly, we employ (4.12) to estimate the latter sum on the fourth line of (4.9),

$$\begin{aligned} \sum_{k \geq 1} \int_{\mathcal{H}_{-s}} \left| \lambda_k \frac{\partial^2 \psi(X)}{\partial z_k^2} \right| \mu(dX) &= \int_{\mathcal{H}_{-s}} \left| \sum_{k \geq 1} \lambda_k k^{-2s} k^{2s} \frac{\partial^2 \psi(X)}{\partial z_k^2} \right| \mu(dX) \\ &\leq \|D^2\psi\|_\infty \int_{\mathcal{H}_{-s}} \sum_{k \geq 1} \lambda_k k^{-2s} \mu(dX) < \infty. \end{aligned} \tag{4.17}$$

We can now apply Fubini theorem on the Hilbert space  $\mathcal{H}_{-s}$ , see e.g. [45]. For  $X \in \mathcal{H}_{-s}$ , we write

$$X = P_{x,v} X + P_{x,v}^\perp X = x e_x + v e_v + Z,$$

where  $P_{x,v}X = xe_x + ve_v$  is the projection on the subspace  $\langle\{e_x, e_v\}\rangle$  and  $P_{x,v}^\perp X = Z$  is the projection on  $\langle\{e_x, e_v\}\rangle^\perp$ . Then,  $\mu$  can be decomposed as  $\mu = \mu_{x,v} \times \mu_{x,v}^\perp$ , where  $\mu_{x,v} = \mu_x \times \mu_v$  is a measure on  $P_{x,v}\mathcal{H}_{-s}$  and  $\mu_{x,v}^\perp = \prod_{k \geq 1} \nu$  is a measure on  $P_{x,v}^\perp\mathcal{H}_{-s}$ . It follows that

$$\begin{aligned} \int_{\mathcal{H}_{-s}} I_{0,1}(X)\mu(dX) &= \int_{\mathcal{H}_{-s}} v \frac{\partial\psi(X)}{\partial x} - \frac{1}{m} \Phi'(x)\theta^R(x) \frac{\partial\psi(X)}{\partial v} \mu(dX) \\ &= \int_{P_{x,v}^\perp\mathcal{H}_{-s}} \int_{\mathbb{R}^2} v \frac{\partial\psi(X)}{\partial x} - \frac{1}{m} \Phi'(x)\theta^R(x) \frac{\partial\psi(X)}{\partial v} \mu_{x,v}(dx, dv) \mu_{x,v}^\perp(dZ), \end{aligned} \tag{4.18}$$

where we use Fubini’s theorem in the last implication. This is possible since we already established the absolute convergence in (4.13). Integrating by parts the first integral against  $\mu_{x,v}$  yields

$$\begin{aligned} \int_{\mathcal{H}_{-s}} I_{0,1}(X)\mu(dX) &= \int_{P_{x,v}^\perp\mathcal{H}_{-s}} \int_{\mathbb{R}^2} v\Phi'(x) (1 - \theta^R(x)) \psi(X) \mu_{x,v}(dx, dv) \mu_{x,v}^\perp(dZ) \\ &= \int_{\mathcal{H}_{-s}} v\Phi'(x) (1 - \theta^R(x)) \psi(X) \mu(dX). \end{aligned} \tag{4.19}$$

Similarly for  $I_{0,2}$ , we have

$$\begin{aligned} \int_{\mathcal{H}_{-s}} I_{0,2}(X)\mu(dX) &= \int_{\mathcal{H}_{-s}} \left[ -\frac{\gamma}{m} v \frac{\partial\psi(X)}{\partial v} + \frac{\gamma}{m^2} \frac{\partial^2\psi(X)}{\partial v^2} \right] \mu(dX) \\ &= \int_{P_{x,v}^\perp\mathcal{H}_{-s}} \int_{\mathbb{R}} \left[ -\frac{\gamma}{m} v \frac{\partial\psi(X)}{\partial v} + \frac{\gamma}{m^2} \frac{\partial^2\psi(X)}{\partial v^2} \right] \mu_v(dv) \mu_v^\perp(d(xe_x + Z)) \\ &= 0, \end{aligned} \tag{4.20}$$

where we have employed (4.4) in the last implication. For the last two terms  $I_{k,1}, I_{k,2}$ , after integration by parts, we invoke (4.6), (4.5) respectively to obtain

$$\int_{\mathcal{H}_{-s}} I_{k,1}(X)\mu(dX) = 0, \quad \int_{\mathcal{H}_{-s}} I_{k,2}(X)\mu(dX) = 0, \quad k \geq 1. \tag{4.21}$$

Formula (4.7) now follows immediately from (4.9), (4.19), (4.20) and (4.21), thus completing the proof.  $\square$

We now show that  $\mu$  is almost invariant under the cutoff system (3.7).

**Proposition 17.** *Let  $R > 0$ . For every  $\psi \in C_b^2(\mathcal{H}_{-s})$ , there exists  $\epsilon^R(\psi, t) > 0$  such that*

$$\int_{\mathcal{H}_{-s}} \mathcal{P}^R(t)\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX) + \epsilon^R(\psi, t). \tag{4.22}$$

Furthermore,  $\epsilon^R(\psi, t) \rightarrow 0$  as  $R \rightarrow \infty$ .

**Proof.** Since equation (3.7) has a globally Lipschitz drift term, in view of theorem 9.23 from [7], for every  $\psi \in C_b^2(\mathcal{H}_{-s})$ ,  $\mathcal{P}^R(t)\psi \in C_b^2(\mathcal{H}_{-s})$  satisfies the Kolmogorov backward equation, namely

$$\mathcal{P}^R(t)\psi(X) = \psi(X) + \int_0^t \mathcal{L}^R \mathcal{P}^R(r)\psi(X) dr. \tag{4.23}$$

Integrating both sides on  $\mathcal{H}_{-s}$  with respect to  $\mu$  gives

$$\int_{\mathcal{H}_{-s}} \mathcal{P}^R(t)\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX) + \int_0^t \int_{\mathcal{H}_{-s}} \mathcal{L}^R \mathcal{P}^R(r)\psi(X)\mu(dX) dr. \tag{4.24}$$

We note that Fubini’s theorem was applied to switch the order of integration in the double-integral term above. Indeed, from (4.13)-(4.16), we see that for all  $r \in [0, t]$

$$\int_{\mathcal{H}_{-s}} |\mathcal{L}^R \mathcal{P}^R(r)\psi(X)| \mu(dX) \leq c (\|D\mathcal{P}^R(r)\psi\|_\infty + \|D^2\mathcal{P}^R(r)\psi\|_\infty),$$

where  $c > 0$  is a constant independent of  $R > 0$ . Furthermore, in view of theorem 9.8 and 9.9 from [7],  $\sup_{0 \leq r \leq t} \|D\mathcal{P}^R(r)\psi\|_\infty$  and  $\sup_{0 \leq r \leq t} \|D^2\mathcal{P}^R(r)\psi\|_\infty$  are both finite. We thus infer that  $\int_0^t \int_{\mathcal{H}_{-s}} |\mathcal{L}^R \mathcal{P}^R(r)\psi(X)| \mu(dX) dr < \infty$ , which guarantees that the Fubini’s theorem is applicable. Now, since  $\mathcal{P}(t)\psi \in C_b^2(\mathcal{H}_{-s})$  for all  $t \geq 0$ , lemma 16 implies that

$$\begin{aligned} \int_{\mathcal{H}_{-s}} \mathcal{P}^R(t)\psi(X)\mu(dX) &= \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX) \\ &+ \int_0^t \int_{\mathcal{H}_{-s}} v\Phi'(x) (1 - \theta^R(x)) \mathcal{P}^R(r)\psi(X)\mu(dX) dr. \end{aligned} \tag{4.25}$$

Let  $\varepsilon^R(\psi, t)$  be given by

$$\varepsilon^R(\psi, t) := \int_0^t \int_{\mathcal{H}_{-s}} v\Phi'(x) (1 - \theta^R(x)) \mathcal{P}^R(r)\psi(X)\mu(dX) dr. \tag{4.26}$$

It is clear that the integrand on the above RHS is dominated by  $\|\psi\|_\infty |v\Phi'(x)|$  and that

$$\|\psi\|_\infty \int_0^t \int_{\mathcal{H}_{-s}} |v\Phi'(x)| \mu(dX) dr = t \|\psi\|_\infty \int_{\mathbb{R}} |v| \mu_v(dv) \int_{\mathbb{R}} |\Phi'(x)| \mu_x(dx) < \infty,$$

since  $\mu_v$  is Gaussian and by assumption 1,  $\Phi'(x)e^{-\Phi(x)}$  is integrable. We additionally note that by the construction of local solutions,  $X^R(r) \rightarrow X(r)$  as  $R \rightarrow \infty$  a.s. It follows that  $\mathcal{P}^R(r)\psi(X) \rightarrow \mathcal{P}(r)\psi(X)$ , implying  $v\Phi'(x) (1 - \theta^R(x)) \mathcal{P}^R(r)\psi(X) \rightarrow 0$ , since  $\theta^R(x) \rightarrow 1$ . We therefore apply the dominated convergence theorem to infer that  $\varepsilon^R(\psi, t) \rightarrow 0$ , which completes the proof.  $\square$

With proposition 17 in hand, we are ready to give the proof of theorem 7.

**Proof of theorem 7.** . By taking  $R \rightarrow \infty$  on both sides of (4.22), we obtain for all  $\psi \in C_b^2(\mathcal{H}_{-s})$

$$\int_{\mathcal{H}_{-s}} \mathcal{P}(t)\psi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \psi(X)\mu(dX).$$

For  $\varphi \in C_b(\mathcal{H}_{-s})$ , approximating  $\varphi$  by a sequence  $\{\psi_k\} \subset C_b^2(\mathcal{H}_{-s})$ , we apply the dominated convergence theorem to arrive at

$$\int_{\mathcal{H}_{-s}} \mathcal{P}(t)\varphi(X)\mu(dX) = \int_{\mathcal{H}_{-s}} \varphi(X)\mu(dX).$$

The proof is complete. □

### 5. Uniqueness of the invariant measure in the diffusive regime

In order to prove uniqueness of  $\mu$ , we will construct an *asymptotic coupling* using an appropriate Girsanov shift argument, following and applying the methods and ideas developed in [26, 51, 71]. Intuitively, this means that solutions started from different initial data have a positive probability of converging to one another as  $t \rightarrow \infty$ . theorem 1.1 of [26] will then allow us to conclude there is only one ergodic invariant measure, thus uniqueness of  $\mu$  follows by ergodic decomposition. The idea of using Girsanov shift to construct *asymptotic coupling* first appeared in the work of [71] and was later developed in [26]. For some more recent applications of this theory to SPDEs, we refer the reader to [12, 44].

For the reader’s convenience, we briefly explain the framework of the *asymptotic coupling method* adapted to our setting, following [12, 26]. To begin, we denote by  $\mathcal{H}_{-s}^{\mathbb{N}}$  the pathspace over  $\mathcal{H}_{-s}$ ,

$$\mathcal{H}_{-s}^{\mathbb{N}} = \{\mathbf{U} : \mathbb{N} \rightarrow \mathcal{H}_{-s}\} = \{\mathbf{U} = (U_0, U_1, U_2, \dots) : U_i \in \mathcal{H}_{-s}\},$$

and let  $\mathcal{P}(\mathcal{H}_{-s}^{\mathbb{N}} \times \mathcal{H}_{-s}^{\mathbb{N}})$  be the set of probability measures on  $\mathcal{H}_{-s}^{\mathbb{N}} \times \mathcal{H}_{-s}^{\mathbb{N}}$ . For any two measures  $M_1, M_2$  on  $\mathcal{H}_{-s}^{\mathbb{N}}$ , we denote by  $\tilde{\mathcal{C}}(M_1, M_2)$  the collection of *asymptotically equivalent coupling* for  $M_1, M_2$ ,

$$\tilde{\mathcal{C}}(M_1, M_2) = \{\Gamma \in \mathcal{P}(\mathcal{H}_{-s}^{\mathbb{N}} \times \mathcal{H}_{-s}^{\mathbb{N}}) : \Gamma \Pi_i^{-1} \ll M_i, i = 1, 2\}, \tag{5.1}$$

where  $\Pi_1(u, v) = u$  and  $\Pi_2(u, v) = v$ . For any initial condition  $X_0 \in \mathcal{H}_{-s}$ , let  $\mathbf{X} = (X_0, X(1), X(2), \dots)$  be the corresponding solution path on  $\mathcal{H}_{-s}^{\mathbb{N}}$  where  $X(t)$  solves (3.1). Then the law of  $\mathbf{X}$ , denoted by  $\delta_{X_0} \mathcal{P}^{\mathbb{N}}$ , defines a probability measure on  $\mathcal{H}_{-s}^{\mathbb{N}}$ . Next, we introduce the set  $\mathcal{D}$  given by

$$\mathcal{D} = \{(\mathbf{U}, \mathbf{V}) \in \mathcal{H}_{-s}^{\mathbb{N}} \times \mathcal{H}_{-s}^{\mathbb{N}} : \lim_{n \rightarrow \infty} \|U_n - V_n\|_{-s} = 0\}. \tag{5.2}$$

Having introduced the above, we will seek to apply the following result (see corollary 2.2 of [26] and corollary 2.1 [12]).

**Theorem 18.** *If for every pair  $X_0, \tilde{X}_0 \in \mathcal{H}_{-s}$ , there exists an element  $\Gamma \in \tilde{\mathcal{C}}(\delta_{X_0} \mathcal{P}^{\mathbb{N}}, \delta_{\tilde{X}_0} \mathcal{P}^{\mathbb{N}})$  such that  $\Gamma(\mathcal{D}) > 0$ , then there exists at most one ergodic invariant measure for (3.1).*

The problem thus reduces to constructing such a coupling  $\Gamma$ . To this end, we introduce another process  $\tilde{X}(t)$  on  $\mathcal{H}_{-s}$  satisfying the following shifted version of equation (3.1)

$$d\tilde{X}(t) = L\tilde{X}(t) dt + F(\tilde{X}(t)) dt + B dW(t) + B U(X(t), \tilde{X}(t)) \mathbf{1}\{t \leq \tau\} dt. \tag{5.3}$$

In the above,  $\tilde{X}(0) = \tilde{X}_0 \in \mathcal{H}_{-s}$ ,  $\tau$  is a stopping time and  $U(X(t), \tilde{X}(t)) \in L^2(\Omega, \mathcal{W})$  is an adapted control depending on both  $\tilde{X}$  and the process  $X$  satisfying (3.1) with  $X(0) = X_0 \in \mathcal{H}_{-s}$ . Here we recall that  $\mathcal{W}$  is the auxiliary Hilbert space on which  $W(t)$  evolves, [7]. Now notice that if we set

$$\tilde{W}(t) = W(t) + \int_0^t U(X(r), \tilde{X}(r)) \mathbf{1}\{r \leq \tau\} dr \tag{5.4}$$

and the control  $U$  and stopping time  $\tau$  are such that, for some deterministic constant  $C > 0$ ,

$$\mathbb{P}\left\{ \int_0^\infty \|U(X(t), \tilde{X}(t)) \mathbf{1}\{t \leq \tau\}\|_{\mathcal{W}}^2 dt < C \right\} = 1, \tag{5.5}$$

then  $W$  and  $\tilde{W}$  are equivalent on  $C([0, \infty); \mathcal{W})$ . As a consequence, the processes  $X$  and  $\tilde{X}$  with  $X(0) = \tilde{X}(0) = \tilde{X}_0 \in \mathcal{H}_{-s}$  are mutually absolutely continuous on the infinite time horizon  $[0, \infty)$ . As shown later in the proof of theorem 10, our coupling  $\Gamma$  is essentially the law of the pair  $(X(\cdot, X_0), \tilde{X}(\cdot, \tilde{X}_0))$ . However, in order for  $\Gamma$  to meet the requirement  $\Gamma(\mathcal{D}) > 0$  from theorem 18, by introducing the difference

$$\bar{X}(t) = X(t) - \tilde{X}(t) = (\bar{x}(t), \bar{v}(t), \bar{z}_1(t), \dots), \tag{5.6}$$

we have to pick  $U$  and  $\tau$  such that (5.5) is satisfied,  $\mathbb{P}\{\tau = \infty\} > 0$  and

$$\|\bar{X}(t)\|_{-s} \rightarrow 0 \text{ as } t \rightarrow \infty \text{ on the event } \{\tau = \infty\}.$$

Thus, reemphasizing what was discussed above, we are constructing the control  $U$  and the stopping time  $\tau$  such that we can drive two solutions of (3.1) with different initial data to one another as  $t \rightarrow \infty$  on a set of positive probability.

To introduce our choice of  $U$  and  $\tau$ , first observe that  $\bar{X}$  satisfies  $\bar{X}(0) = X_0 - \tilde{X}_0$  and

$$d\bar{X}(t) = L\bar{X}(t) dt + [F(X(t)) - F(\tilde{X}(t))] dt - B U(X(t), \tilde{X}(t)) \mathbf{1}\{t \leq \tau\} dt. \tag{5.7}$$

Writing  $X(t) = (x(t), v(t), z_1(t), \dots)$ ,  $\tilde{X}(t) = (\tilde{x}(t), \tilde{v}(t), \tilde{z}_1(t), \dots)$  and recalling the notation (5.6), we define for given  $\lambda > 0$

$$u_0(X(t), \tilde{X}(t)) = \frac{m}{\sqrt{2\gamma}} \left[ \left(3\lambda - \frac{\gamma}{m}\right) \bar{v}(t) + 2\lambda^2 \bar{x}(t) - \frac{1}{m} (\Phi'(x(t)) - \Phi'(\tilde{x}(t))) - \frac{1}{m} \sum_{k \geq 1} \sqrt{c_k} \bar{z}_k(t) \right], \tag{5.8}$$

and

$$U(X(t), \tilde{X}(t)) = (0, u_0(X(t), \tilde{X}(t)), 0, 0, \dots). \tag{5.9}$$

Note that  $BU$  only possibly enacts control over the velocity difference  $\bar{v}(t) = v(t) - \tilde{v}(t)$ , and this is essentially done to gain control over nonlinear difference  $\Phi'(x) - \Phi'(\tilde{x})$ .

For a given  $\kappa > 0$ , we define the stopping time  $\tau = \tau(\kappa)$  by

$$\tau = \inf_{t \geq 0} \left\{ \int_0^t |u_0(X(s), \tilde{X}(s))|^2 ds \geq \kappa \right\}. \tag{5.10}$$

With these choices, note that on the event  $\{t < \tau\}$ ,  $\bar{X}(t) = (\bar{x}(t), \bar{v}(t), \bar{z}_1(t), \dots)$  satisfies the following system of equations

$$\begin{aligned}
 \frac{d\bar{x}(t)}{dt} &= \bar{v}(t), & \bar{x}(0) &= \bar{x}_0, \\
 \frac{d\bar{v}(t)}{dt} &= -3\lambda\bar{v}(t) - 2\lambda^2\bar{x}(t), & \bar{v}(0) &= \bar{v}_0, \\
 \frac{d\bar{z}_k(t)}{dt} &= -\lambda_k\bar{z}_k(t) + \sqrt{c_k}\bar{v}(t), & \bar{z}_k(0) &= (\bar{z}_k)_0.
 \end{aligned}
 \tag{5.11}$$

Intuitively, the coefficient  $\lambda > 0$  will be picked so that that  $\|\bar{X}(t)\|_{\mathcal{H}_{-s}} \rightarrow 0$  as  $t \rightarrow \infty$  on the event  $\{\tau = \infty\}$ . Hence the control induces the requisite dissipation, but we still need to see that we can pick  $\kappa > 0$  so that (5.5) is satisfied and  $\mathbb{P}\{\tau = \infty\} > 0$ . Before turning to this issue, we make the following remark.

**Remark 19.** (a) There is a significantly flexibility in the choice of  $u_0$  in (5.8). One can of course choose other formulas for the coefficients of  $\bar{x}(t)$  and  $\bar{v}(t)$  in (5.8) as long as  $\|\bar{X}(t)\|_{-s} \rightarrow 0$  as  $t \rightarrow \infty$  on  $\{\tau = \infty\}$ .

(b). The appearance of  $u_0$  requires the drag constant  $\gamma$  be strictly positive. We note that for well-posedness (see proposition 6) and the existence of invariant measures (see theorem 7),  $\gamma$  can be zero.

With these observations, we state the following proposition which outlines the needed details to deduce unique ergodicity.

**Proposition 20.** *Under the Hypothesis of theorem 10 and recalling  $m, \gamma > 0$  from (2.7) and  $\alpha, \beta$  as in (2.4), let  $\lambda > 0$  be as in (5.8) and  $\kappa > 0$  as in (5.10). Then there exist  $\lambda = \lambda(\alpha, \beta) > 0, \kappa = \kappa(X_0, \tilde{X}_0, \gamma, m, \alpha, \beta) > 0$  such that  $\tau = \tau(\kappa)$  and  $U$  are such that*

- (a) Condition (5.5) is satisfied.
- (b)  $\|\bar{X}(t)\|_{-s} \rightarrow 0$  as  $t \rightarrow \infty$  on  $\{\tau = \infty\}$ .
- (c)  $\mathbb{P}\{\tau = \infty\} > 0$ .

Before presenting the proof of proposition 20, we now show how one can deduce unique ergodicity of (3.1) by combining proposition 20 and theorem 18, (see [12, 26] for further details).

**Proof of theorem 10.** In view of proposition 20 (a), the Novikov’s condition is verified

$$\mathbb{E} \left[ \exp \int_0^\infty \|U(X(t), \tilde{X}(t))\mathbf{1}\{t \leq \tau\}\|_{\mathcal{W}}^2 dt \right] < e^C.$$

The process  $\tilde{W}(t)$  defined in (5.4) is thus equivalent to the Wiener process  $W(t)$  on  $C([0, \infty); \mathcal{W})$  by Girsanov’s theorem, (theorem 10.4 from [7]). It follows that the process  $\tilde{X}(\cdot, \tilde{X}_0)$  solving (5.3) is absolutely continuous to the process  $X(\cdot, \tilde{X}_0)$  on  $C([0, \infty); \mathcal{H}_{-s})$ , (see e.g. [63]). It follows that the law  $\Gamma$  induced by

$$\{(X(nt, X_0), \tilde{X}(nt, \tilde{X}_0)) : n = 0, 1, 2, \dots\}$$

belongs to  $\tilde{\mathcal{C}}(\delta_{X_0} \mathcal{P}^{\mathbb{N}}, \delta_{\tilde{X}_0} \mathcal{P}^{\mathbb{N}})$ . In addition, proposition 20 (b) and (c) imply that  $\Gamma(\mathcal{D}) > 0$  where  $D$  is given by (5.2). We therefore conclude unique ergodicity by virtue of theorem 18, thus completing the proof. □

We now turn to the proof of proposition 20. Parts (a) and (b) essentially follow by construction. Establishing part (c), however, requires a bit more work. To complete the proof, we need a crucial estimate on the potential  $\Phi$  which relies on Lyapunov methods. To this end, for  $N \in \mathbb{N}$  and  $s \in \mathbb{R}$ , we introduce  $\Theta : \mathcal{H}_{-s} \rightarrow \mathbb{R}$  defined by

$$\Theta(X; s, N) = \frac{1}{m} \Phi(x) + \frac{1}{2} v^2 + \frac{1}{2m} \sum_{k=1}^N z_k^2 + \frac{1}{2} \sum_{k>N} k^{-2s} z_k^2. \tag{5.12}$$

In the diffusive regime, it turns out that  $\Theta(X; s, N)$  can be chosen such that it satisfies a Lyapunov bound that is stronger than the bound on  $\Psi$  from proposition 14. That is:

**Proposition 21.** *Let  $\Theta(X; s, N)$  be defined as in (5.12). Then, under assumptions 1 and Condition (D) of assumption 3, there exists  $N = N(m, \gamma, \alpha, \beta, s) \in \mathbb{N}$  sufficiently large such that, for some  $a > 0$ ,  $\Theta(X) := \Theta(X; s, N)$  satisfies*

$$\sup_{X \in \mathcal{H}_{-s}} \mathcal{L}\Theta(X) \leq a.$$

Proposition 21 will be established at the end of the section, but the proof follows a similar line of reasoning to that employed in the proof of proposition 14.

**Remark 22.**

- (a) In assumption 3, the diffusive regime (D) requires  $\alpha > 1$  as opposed to  $\alpha \in (0, 1)$  and  $\alpha = 1$  in, respectively, the subdiffusive (SD) and critical (C) regimes. Recalling  $c_k, \lambda_k$  from (2.4), the condition  $\alpha > 1$  is needed so that the infinite sum

$$\sum_{k \geq 1} \frac{c_k}{2\lambda_k k^{-2s}} = \sum_{k \geq 1} k^{-1-(\alpha-1)\beta+s}$$

converges, as shown later in (5.18) and (5.21). This convergence is critically employed in the proofs of propositions 20 (c) and 21.

- (b) The asymptotic behavior of  $\lambda_k$  as  $k \rightarrow \infty$  presents a barrier to obtaining a stronger Lyapunov bound of the form

$$\mathcal{L}\Theta(X) \leq -c\Theta(X) + a$$

in the proof of proposition 21, where  $c > 0$  is a constant. The above inequality, however, can be readily achieved in the finite-dimensional system (2.2), see [60]. With the appropriate support properties of the diffusion, such a bound implies geometric ergodicity. However, because we cannot see immediately why (2.2) holds in our infinite-dimensional system, suggests that perhaps the system relaxes to equilibrium slower than an exponential rate.

By combining the previous Proposition with the exponential martingale inequality, we obtain the following corollary.

**Corollary 23.** *Under assumptions 1 and Condition (D) of assumption 3, let  $X(t) = (x(t), v(t), \dots)$  be the solution of (2.7) with initial condition  $X_0 \in \mathcal{H}_{-s}$ . Let  $\Theta$  be the Lyapunov function defined in (5.12). Then there exists  $\varepsilon = \varepsilon(m, \gamma, \alpha, \beta) > 0$  such that for every  $\eta, r > 0$ ,*

$$\mathbb{P}\left\{\sup_{t \geq 0} \frac{e^{-\eta t} \Phi(x(t))}{m} - \Theta(X_0) - \frac{a}{\eta} \geq r\right\} \leq e^{-\varepsilon r}, \tag{5.13}$$

where  $a$  is as in the statement of proposition 21.

The proof of corollary 23 will also be given at the end of this section.

We now conclude proposition 20 assuming the previous two results.

**Proof of proposition 20.** We begin by showing part (a) of the result. In view of formulas (3.3) and (3.4), the norm of the control  $U(t)$  in  $\mathcal{W}$  satisfies

$$\|U(t)\|_{\mathcal{W}}^2 = u_0(t)^2.$$

It thus follows by definition of  $\tau$  that

$$\int_0^\infty \|U(X(t), \tilde{X}(t))\mathbf{1}\{t \leq \tau\}\|_{\mathcal{W}}^2 dt = \int_0^\infty |u_0(t)|^2 \mathbf{1}\{t \leq \tau\} dt = \kappa \quad \mathbb{P} - \text{almost surely.}$$

Applying theorem 10.4 from [7] finishes the proof of part (a).

To conclude part (b), for  $t \leq \tau$  one can readily verify that the exact solution of (5.11) is given by

$$\begin{aligned} \bar{x}(t) &= \left(2\bar{x}_0 + \frac{\bar{v}_0}{\lambda}\right)e^{-\lambda t} - \left(\bar{x}_0 + \frac{\bar{v}_0}{\lambda}\right)e^{-2\lambda t} \\ \bar{v}(t) &= -(2\lambda\bar{x}_0 + \bar{v}_0)e^{-\lambda t} + 2(\lambda\bar{x}_0 + \bar{v}_0)e^{-2\lambda t} \\ \bar{z}_k(t) &= e^{-\lambda_k t} \left[ (\bar{z}_k)_0 + \sqrt{c_k} \int_0^t e^{\lambda_k r} \bar{v}(r) dr \right]. \end{aligned}$$

From this, it follows that

$$|\bar{x}(t)| \leq C_1 e^{-\lambda t}, \quad |\bar{v}(t)| \leq C_2 e^{-\lambda t}, \tag{5.14}$$

where

$$C_1 := 3|\bar{x}_0| + \frac{3|\bar{v}_0|}{\lambda}, \quad C_2 := 4\lambda|\bar{x}_0| + 4|\bar{v}_0|.$$

Combining these bounds, we thus obtain the following bound on  $\bar{z}_k(t)$

$$|\bar{z}_k(t)| \leq e^{-\lambda_k t} \left[ |(\bar{z}_k)_0| + C_2 \sqrt{c_k} \int_0^t e^{(\lambda_k - \lambda)r} dr \right].$$

Choosing  $\lambda = \lambda_1 + 1$  note that  $\lambda - \lambda_k \geq 1$  for all  $k \geq 1$  since  $\lambda_k \downarrow 0$ . With this choice of  $\lambda$ , it follows from the inequality above that for all  $k \geq 1$ ,

$$|\bar{z}_k(t)| \leq e^{-\lambda_k t} \left( |(\bar{z}_k)_0| + C_2 \sqrt{c_k} \right), \tag{5.15}$$

and hence by Young's inequality,

$$\bar{z}_k(t)^2 \leq 2e^{-2\lambda_k t} \left[ (\bar{z}_k)_0^2 + C_2^2 c_k \right].$$

Thus putting it all together we find that

$$\begin{aligned} \|\bar{X}(t)\|_{-s}^2 &= \bar{x}(t)^2 + \bar{v}(t)^2 + \sum_{k \geq 1} k^{-2s} \bar{z}_k(t)^2 \\ &\leq (C_1^2 + C_2^2) e^{-\lambda t} + 2 \sum_{k \geq 1} k^{-2s} (\bar{z}_k)_0^2 e^{-2\lambda k t} + 2C_2^2 \sum_{k \geq 1} k^{-2s} e^{-2\lambda k t} c_k. \end{aligned}$$

Thus on the event  $\{\tau = \infty\}$ , it is now evident that  $\|\bar{X}\|_{-s}^2 \rightarrow 0$  as  $t \rightarrow \infty$  by applying the Monotone convergence theorem.

Turning to part (c) of the result, for any  $R > 0$  consider the event  $E_R$  given by

$$E_R = \left\{ \sup_{t \geq 0} \frac{e^{-\lambda t/2q} \Phi(x(t))}{m} - \Theta(X_0) - \frac{2qa}{\lambda} < R \right\}, \tag{5.16}$$

where  $q$  is the constant from Assumption 8. In view of corollary 23 with  $\eta = \lambda/2q$ ,  $E_R$  has positive probability provided  $R = R(\gamma, m, \alpha, \beta) > 0$  is sufficiently large. We first claim that on  $E_R$ ,

$$\int_0^\infty \|U(t)\|_{\mathcal{W}}^2 dt \text{ is bounded almost surely.}$$

To see this, recall by definition of the control  $U$  that

$$\int_0^\infty \|U(t)\|_{\mathcal{W}}^2 dt = \int_0^\infty |u_0(t)|^2 dt.$$

Thus estimating  $u_0(t)^2$ , from (5.8) we have

$$\begin{aligned} u_0(t)^2 &\leq \frac{2m^2}{\gamma} \left[ \left(3\lambda - \frac{\gamma}{m}\right)^2 \bar{v}(t)^2 + 4\lambda^4 \bar{x}(t)^2 \right. \\ &\quad \left. + \frac{1}{m^2} \left( \sum_{k \geq 1} \sqrt{c_k} |\bar{z}_k(t)| \right)^2 + \frac{1}{m^2} |\Phi'(x(t)) - \Phi'(\bar{x}(t))|^2 \right] \\ &= I_1(t) + I_2(t) + I_3(t) + I_4(t). \end{aligned}$$

For  $I_1(t) + I_2(t)$ , apply (5.14) to find

$$I_1(t) + I_2(t) \leq \frac{2m^2}{\gamma} \left[ \left(3\lambda - \frac{\gamma}{m}\right)^2 C_2^2 + 4\lambda^4 C_1^2 \right] e^{-2\lambda t} = C_3 e^{-2\lambda t}, \tag{5.17}$$

where  $C_3 := \frac{2m^2}{\gamma} \left[ \left(3\lambda - \frac{\gamma}{m}\right)^2 C_2^2 + 4\lambda^4 C_1^2 \right]$ . For  $I_3(t)$ , employ (5.15) to see that

$$\begin{aligned} I_3(t) &\leq \frac{2}{\gamma} \left( \sum_{k \geq 1} \sqrt{c_k} e^{-\lambda k t} (|\bar{z}_k)_0| + C_2 \sqrt{c_k} \right)^2 \\ &\leq \frac{4}{\gamma} \left( \sum_{k \geq 1} \sqrt{c_k} e^{-\lambda k t} (|\bar{z}_k)_0| \right)^2 + \frac{4C_2^2}{\gamma} \left( \sum_{k \geq 1} c_k e^{-\lambda k t} \right)^2 \\ &\leq \frac{4\bar{X}_0^2}{\gamma} \sum_{k \geq 1} \frac{c_k e^{-2\lambda k t}}{k^{-2s}} + \frac{4C_2^2}{\gamma} K(t)^2, \end{aligned}$$

where the last inequality follows by Cauchy–Schwarz inequality since  $K(t) = \sum_{k \geq 1} c_k e^{-\lambda_k t}$  by definition. Lastly, to estimate  $I_4(t)$ , assumptions 8 and (5.14) together imply that

$$I_4(t) \leq \frac{4}{\gamma} \bar{x}(t)^2 (f(\bar{x}(t))^2 + \Phi(x(t))^{2q}) \leq \frac{4C_1^2 e^{-2\lambda t}}{\gamma} \left( \sup_{|y| \leq C_1} f(y)^2 + \Phi(x(t))^{2q} \right).$$

Now on the event  $E_R$ , we note that

$$\sup_{t \geq 0} e^{-\lambda t} \Phi(x(t))^{2q} < \left( m\Theta(X_0) + \frac{2qma}{\lambda} + mR \right)^{2q} =: C_4.$$

Hence

$$\begin{aligned} I_4(t) &\leq \frac{4C_1^2 e^{-\lambda t}}{\gamma} \left( e^{-\lambda t} \sup_{|y| \leq C_1} f(y)^2 + \sup_{t \geq 0} e^{-\lambda t} \Phi(x(t))^{2q} \right) \\ &\leq \frac{4C_1^2}{\gamma} \left( \sup_{|y| \leq C_1} f(y)^2 + C_4 \right) e^{-\lambda t} \\ &= C_5 e^{-\lambda t}. \end{aligned}$$

Combining these bounds for  $I_1, I_2, I_3, I_4$  shows that on  $E_R$ ,

$$\begin{aligned} \int_0^\infty u_0(t)^2 dt &\leq \int_0^\infty I_1(t) + I_2(t) + I_3(t) + I_4(t) dt \\ &\leq \int_0^\infty C_3 e^{-2\lambda t} + \frac{4\bar{X}_0^2}{\gamma} \sum_{k \geq 1} \frac{c_k e^{-2\lambda_k t}}{k^{-2s}} + \frac{4C_2^2}{\gamma} K(t)^2 + C_5 e^{-\lambda t} dt \\ &= \frac{4\bar{X}_0^2}{\gamma} \sum_{k \geq 1} \frac{c_k}{2\lambda_k k^{-2s}} + \int_0^\infty C_3 e^{-2\lambda t} + \frac{4C_2^2}{\gamma} K(t)^2 + C_5 e^{-\lambda t} dt. \end{aligned} \tag{5.18}$$

We invoke assumption 3 again to see that  $\sum_{k \geq 1} \frac{c_k}{2\lambda_k k^{-2s}} < \infty$ . Furthermore, in view of (2.6),  $K(t)^2 \sim t^{-2\alpha}$  as  $t \rightarrow \infty$ , implying  $K(t)^2$  is integrable. We thus infer from (5.18) a constant  $C_6 = C_6(X_0, \tilde{X}_0, \gamma, m, \alpha, \beta) > 0$  such that on  $E_R$ ,

$$\int_0^\infty u_0(t)^2 dt \leq C_6. \tag{5.19}$$

Finally, we choose  $\kappa > C_6$  in the definition of  $\tau = \tau(\kappa)$  forcing  $E_R \subset \{\tau = \infty\}$ . We therefore, conclude that  $\mathbb{P}\{\tau_\kappa = \infty\} > 0$ . The proof is thus complete.  $\square$

We now finish this section by giving the proofs of proposition 21 and corollary 23.

**Proof of proposition 21.** We have

$$\begin{aligned} \mathcal{L}\Theta &= -\frac{\gamma}{m} v^2 - \frac{1}{m} \sum_{k=1}^N \lambda_k z_k^2 - \sum_{k > N} \lambda_k k^{-2s} z_k^2 - \frac{1}{m} \sum_{k > N} \sqrt{c_k} z_k v \\ &\quad + \sum_{k > N} \sqrt{c_k} k^{-2s} z_k v + \frac{\gamma}{m^2} + \frac{1}{m} \sum_{k=1}^N \lambda_k + \sum_{k > N} \lambda_k k^{-2s}. \end{aligned}$$

Young’s inequality combined with Cauchy–Schwarz inequality then gives

$$\frac{1}{m} \sum_{k>N} \sqrt{c_k} z_k v \leq \frac{\gamma}{4m} v^2 + \frac{1}{\gamma m} \sum_{k>N} \frac{c_k}{k^{-2s} \lambda_k} \sum_{k>N} k^{-2s} \lambda_k z_k^2$$

and

$$\sum_{k>N} k^{-2s} \sqrt{c_k} z_k v \leq \frac{\gamma}{4m} v^2 + \frac{m}{\gamma} \sum_{k>N} \frac{k^{-2s} c_k}{\lambda_k} \sum_{k>N} k^{-2s} \lambda_k z_k^2.$$

Combining the previous two inequalities with the first we obtain

$$\mathcal{L}\Theta \leq -\frac{\gamma}{2m} v^2 - \frac{1}{m} \sum_{k=1}^N \lambda_k z_k^2 - a_1 \sum_{k>N} k^{-2s} \lambda_k z_k^2 + a, \tag{5.20}$$

where

$$a := \frac{\gamma}{m^2} + \frac{1}{m} \sum_{k=1}^N \lambda_k + \sum_{k>N} \lambda_k k^{-2s}, \quad a_1 := 1 - \frac{1}{\gamma m} \sum_{k>N} \frac{c_k}{k^{-2s} \lambda_k} - \frac{m}{\gamma} \sum_{k>N} \frac{k^{-2s} c_k}{\lambda_k}. \tag{5.21}$$

We invoke Condition (D) of assumption 3 again to see that

$$\begin{aligned} \sum_{k \geq 1} \lambda_k k^{-2s} &= \sum_{k \geq 1} \frac{1}{k^{\beta+2s}} < \infty, & \sum_{k \geq 1} \frac{c_k}{k^{-2s} \lambda_k} &= \sum_{k \geq 1} \frac{1}{k^{1+(\alpha-1)\beta-2s}} < \infty, \\ \text{and} \quad \sum_{k \geq 1} \frac{k^{-2s} c_k}{\lambda_k} &= \sum_{k \geq 1} \frac{1}{k^{1+(\alpha-1)\beta+2s}} < \infty, \end{aligned}$$

which implies that  $a < \infty$  and that  $N$  can be chosen large enough such that  $0 < a_1 < \infty$ . We therefore conclude  $\mathcal{L}\Theta \leq a$ , which is the desired inequality.  $\square$

**Proof of corollary 23.** Fix  $\eta > 0$  and apply Ito’s Formula to  $e^{-\eta t} \Theta(X(t))$  to find

$$d(e^{-\eta t} \Theta(X(t))) = -\eta e^{-\eta t} \Theta(X(t)) dt + e^{-\eta t} \mathcal{L}\Theta(X(t)) dt + dM_\eta(t) \tag{5.22}$$

where the martingale  $M_\eta$  satisfies

$$\begin{aligned} dM_\eta(t) &= e^{-\eta t} \frac{\sqrt{2\gamma}}{m} v(t) dW_0(t) + \frac{e^{-\eta t}}{m} \sum_{k=1}^N \sqrt{2\lambda_k} z_k(t) dW_k(t) \\ &\quad + e^{-\eta t} \sum_{k>N} \sqrt{2\lambda_k} k^{-2s} z_k(t) dW_k(t). \end{aligned}$$

Note also that the quadratic variation process  $\langle M_\eta \rangle$  has

$$d\langle M_\eta \rangle(t) = \frac{2\gamma e^{-2\eta t}}{m^2} v(t)^2 dt + \frac{2e^{-2\eta t}}{m^2} \sum_{k=1}^N \lambda_k z_k(t)^2 dt + 2e^{-2\eta t} \sum_{k>N} \lambda_k k^{-4s} z_k(t)^2 dt. \tag{5.23}$$

We recall from (5.20) that

$$\mathcal{L}\Theta(X(t)) \leq -\frac{\gamma}{2m}v(t)^2 - \frac{1}{m} \sum_{k=1}^N \lambda_k z_k(t)^2 - a_1 \sum_{k>N} k^{-2s} \lambda_k z_k(t)^2 + a, \quad (5.24)$$

where  $a, a_1$  are defined in (5.21). Combining (5.22)–(5.24), for every  $\varepsilon > 0$  we obtain the estimate

$$\begin{aligned} d(e^{-\eta t}\Theta(X(t))) &\leq ae^{-\eta t}dt + dM_\eta(t) - \frac{\varepsilon}{2}d\langle M_\eta \rangle(t) \\ &\quad - e^{-2\eta t} \left[ \frac{\gamma}{2m}v(t)^2 + \frac{1}{m} \sum_{k=1}^N \lambda_k z_k(t)^2 + a_1 \sum_{k>N} k^{-2s} \lambda_k z_k(t)^2 \right. \\ &\quad \left. - \frac{\varepsilon}{2} \left( \frac{\gamma}{m^2}v(t)^2 + \frac{1}{m^2} \sum_{k=1}^N \lambda_k z_k(t)^2 + \sum_{k>N} \lambda_k k^{-4s} z_k(t)^2 \right) \right] dt. \end{aligned}$$

By choosing  $\varepsilon = \varepsilon(\gamma, m, \alpha, \beta) > 0$  smaller if necessary, the bracket term on the above RHS is nonpositive. Hence

$$d(e^{-\eta t}\Theta(X(t))) \leq ae^{-\eta t}dt + dM_\eta(t) - \frac{\varepsilon}{2}d\langle M_\eta \rangle(t).$$

Integrating with respect to  $t$  we find

$$e^{-\eta t}\Theta(X(t)) - \Theta(X_0) \leq \int_0^\infty ae^{-\eta r}dr + M_\eta(t) - \frac{\varepsilon}{2}\langle M_\eta \rangle(t) = \frac{a}{\eta} + M_\eta(t) - \frac{\varepsilon}{2}\langle M_\eta \rangle(t).$$

Since  $\Theta(X(t)) \geq \Phi(X(t))/m$  by the definition of  $\Theta(X)$ , we infer that

$$\frac{e^{-\eta t}\Phi(x(t))}{m} - \Theta(X_0) - \frac{a}{\eta} \leq M_\eta(t) - \frac{\varepsilon}{2}\langle M_\eta \rangle(t).$$

Invoking the exponential martingale inequality we obtain

$$\begin{aligned} \mathbb{P}\left\{ \sup_{t \geq 0} \frac{e^{-\eta t}\Phi(x(t))}{m} - \Theta(X_0) - \frac{a}{\eta} \geq r \right\} &\leq \mathbb{P}\left\{ \sup_{t \geq 0} \left[ M_\eta(t) - \frac{\varepsilon}{2}\langle M_\eta \rangle(t) \right] \geq r \right\} \\ &\leq e^{-\varepsilon r}, \end{aligned}$$

thus completing the proof. □

### 6. Discussion

We have rigorously studied the GLE in a potential well  $\Phi$  with a power-law decay memory  $K(t)$ , i.e.  $K(t) \sim t^{-\alpha}$ ,  $\alpha > 0$  as  $t \rightarrow \infty$ . Using a Mori–Zwanzig approach, when the memory  $K$  can be written as an infinite sum of exponentials, we represent the non-linear GLE as an infinite-dimensional Markovian system. With nominal conditions on the potential  $\Phi$ , we show that for every  $\alpha > 0$ , this Markovian system is well-posed in suitable spaces and admits an invariant measure. Moreover, using an asymptotic coupling argument, the system is shown

to be uniquely ergodic when  $\alpha > 1$ . The problem of unique ergodicity remains open when  $\alpha \in (0, 1]$ .

A related research topic that is of direct interest is to establish the convergence rate to stationarity. Due to the memory's power-law decay, one might conjecture that this system does not approach the invariant measure with an exponential rate (commonly called geometric ergodicity). That is, it is conceivable that there is a unique invariant measure, but the approach is algebraic instead. Current methods for proving an algebraic rate of convergence to stationarity rely on finding a type of Lyapunov function that is currently unknown for this system.

Lastly, we would like to touch on the term 'ergodicity breaking,' which has appeared in the physics literature in connection with models of anomalous subdiffusion [27, 48]. In particular, there are claims that solutions to the generalized Langevin equation in a quadratic potential can break ergodicity in the sense that a popular expression for a particle time-average does not match a stationary population's ensemble average [36, 37]. It is important to point out though that the time average used in these papers, sometimes called the pathwise mean-squared displacement,

$$\frac{1}{T - \Delta} \int_0^{T-\Delta} (x(s + \Delta) - x(s))^2 ds \quad (6.1)$$

is a 'sliding window' average of squared-displacements and not equivalent to the time-average formula (1.3) used in the mathematical literature on ergodic theory. Notably, the authors of [36] and [37] are able to show that the inequivalence of (6.1) with ensemble averages occurs even for the Ornstein–Uhlenbeck SDE (Brownian motion in a quadratic potential), which is well-known to be a (geometrically) ergodic process in the sense of Birkoff. Therefore, the results and conjectures discussed under the heading of 'ergodicity breaking' in the physics literature do not necessarily align with the results and conjectures we present here. Having said that, the MSD formula (6.1) is an essential tool in the particle tracking literature and we believe it is an interesting and unsolved question as to why this time-average fails to match ensemble averages as one might expect for an ergodic process.

## Acknowledgment

The authors would like to thank Jonathan Mattingly and Pete Kramer for fruitful discussions on the topic of this paper. All authors are grateful for support from the mathematics departments at Tulane University and Iowa State University as well as for support through grants NSF DMS-1644290 (SAM and HDN), NSF DMS-1612898 and DMS-1855504 (DPH), and NSF DMS-1816551, DMS-1522616, and Simons Foundation grant 515990 (NEGH).

## References

- [1] Abate J and Whitt W 1999 Infinite-series representations of Laplace transforms of probability density functions for numerical inversion *J. Oper. Res. Soc. Japan* **42** 268–85
- [2] Adelman S A 1976 Fokker–Planck equations for simple non-Markovian systems *J. Chem. Phys.* **64** 124–30
- [3] Albeverio S, Flandoli F and Sinai Y G 2008 *SPDE in Hydrodynamics: Recent Progress and Prospects: Lectures given at the CIME Summer School held in (Cetraro Italy, August 29–September 3 2005)* (Berlin: Springer)

- [4] Conrad F and Grothaus M 2010 Construction, ergodicity and rate of convergence of N-particle Langevin dynamics with singular potentials *J. Evol. Equ.* **10** 623–62
- [5] Cooke B, Herzog D P, Mattingly J C, McKinley S A and Schmidler S C 2017 Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard–Jones-like repulsive potential *Commun. Math. Sci.* **15** 1987–2025
- [6] Córdoba A, Indei T and Schieber J D 2012 Elimination of inertia from a generalized Langevin equation: applications to microbead rheology modeling and data analysis *J. Rheol.* **56** 185–212
- [7] Da Prato G and Zabczyk J 2014 *Stochastic Equations in Infinite Dimensions* (Cambridge: Cambridge University Press)
- [8] Desposito M A and Viñales A D 2008 Memory effects in the asymptotic diffusive behavior of a classical oscillator described by a generalized Langevin equation *Phys. Rev. E* **77** 031123
- [9] Didier G and Nguyen H 2019 Asymptotic analysis of the mean squared displacement under fractional memory kernels (arXiv:1901.03007)
- [10] Ernst D, Hellmann M, Köhler J and Weiss M 2012 Fractional Brownian motion in crowded fluids *Soft Matter* **8** 4886–9
- [11] Fricks J, Yao L, Elston T C and Forest M G 2009 Time-domain methods for diffusive transport in soft matter *SIAM J. Appl. Math.* **69** 1277–308
- [12] Glatt-Holtz N, Mattingly J C and Richards G 2017 On unique ergodicity in nonlinear stochastic partial differential equations *J. Stat. Phys.* **166** 618–49
- [13] Glatt-Holtz N and Ziane M 2009 Strong pathwise solutions of the stochastic Navier–Stokes system *Adv. Differ. Equ.* **14** 567–600
- [14] Goychuk I 2009 Viscoelastic subdiffusion: from anomalous to normal *Phys. Rev. E* **80** 046125
- [15] Goychuk I 2010 Subdiffusive Brownian ratchets rocked by a periodic force *Chem. Phys.* **375** 450–7
- [16] Goychuk I 2012 Viscoelastic subdiffusion: generalized Langevin equation approach *Adv. Chem. Phys.* **150** 187
- [17] Goychuk I 2015 Anomalous transport of subdiffusing cargos by single kinesin motors: the role of mechano–chemical coupling and anharmonicity of tether *Phys. Biol.* **12** 016013
- [18] Goychuk I 2015 Modeling magnetosensitive ion channels in the viscoelastic environment of living cells *Phys. Rev. E* **92** 042711
- [19] Goychuk I 2018 Sensing magnetic fields with magnetosensitive ion channels *Sensors* **18** 728
- [20] Goychuk I 2018 Viscoelastic subdiffusion in a random gaussian environment *Phys. Chem. Chem. Phys.* **20** 24140–55
- [21] Goychuk I and Kharchenko V 2012 Fractional brownian motors and stochastic resonance *Phys. Rev. E* **85** 051131
- [22] Goychuk I and Kharchenko V 2013 Rocking subdiffusive ratchets: origin, optimization and efficiency *Math. Modelling Nat. Phenom.* **8** 144–58
- [23] Goychuk I, Kharchenko V and Metzler R 2014 How molecular motors work in the crowded environment of living cells: coexistence and efficiency of normal and anomalous transport *PLoS One* **9** e91700
- [24] Goychuk I, Kharchenko V and Metzler R 2014 Molecular motors pulling cargos in the viscoelastic cytosol: how power strokes beat subdiffusion *Phys. Chem. Chem. Phys.* **16** 16524–35
- [25] Grothaus M and Stilgenbauer P 2015 A hypocoercivity related ergodicity method with rate of convergence for singularly distorted degenerate kolmogorov equations and applications **83** 331–79
- [26] Hairer M, Mattingly J C and Scheutzow M 2011 Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations *Probab. Theory Relat. Fields* **149** 223–59
- [27] He Y, Burov S, Metzler R and Barkai E 2008 Random time-scale invariant diffusion and transport coefficients *Phys. Rev. Lett.* **101** 058101
- [28] Herzog D P and Mattingly J C 2017 Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials *Comm. Pure Appl. Math.* **72** 2231–55
- [29] Hill D B, Vasquez P A, Mellnik J, McKinley S A, Vose A, Mu F, Henderson A G, Donaldson S H, Alexis N E and Boucher R C 2014 A biophysical basis for mucus solids concentration as a candidate biomarker for airways disease *PloS One* **9** e87681
- [30] Höfling F and Franosch T 2013 Anomalous transport in the crowded world of biological cells *Rep. Prog. Phys.* **76** 046602
- [31] Hohenegger C 2017 On equipartition of energy and integrals of generalized Langevin equations with generalized rouse kernel *Commun. Math. Sci.* **15** 539–54

- [32] Hohenegger C and McKinley S A 2017 Fluid–particle dynamics for passive tracers advected by a thermally fluctuating viscoelastic medium *J. Comput. Phys.* **340** 688–711
- [33] Hohenegger C and McKinley S A 2018 Reconstructing complex fluid properties from the behavior of fluctuating immersed particles *SIAM J. Appl. Math.* **78** 2200–26
- [34] Indei T, Schieber J D, Córdoba A and Pilyugina E 2012 Treating inertia in passive microbead rheology *Phys. Rev. E* **85** 021504
- [35] Jacod J 2006 *Calcul Stochastique et Problèmes de Martingales* vol 714 (Berlin: Springer)
- [36] Jeon J H and Metzler R 2010 Fractional brownian motion and motion governed by the fractional langevin equation in confined geometries *Phys. Rev. E* **81** 021103
- [37] Jeon J H and Metzler R 2012 Inequivalence of time and ensemble averages in ergodic systems: exponential versus power-law relaxation in confinement *Phys. Rev. E* **85** 021147
- [38] Karatzas I and Shreve S 2012 *Brownian Motion and Stochastic Calculus* vol 113 (Berlin: Springer)
- [39] Kharchenko V and Goychuk I 2012 Flashing subdiffusive ratchets in viscoelastic media *New J. Phys.* **14** 043042
- [40] Kharchenko V and Goychuk I 2013 Subdiffusive rocking ratchets in viscoelastic media: transport optimization and thermodynamic efficiency in overdamped regime *Phys. Rev. E* **87** 052119
- [41] Kou S C 2008 Stochastic modeling in nanoscale biophysics: subdiffusion within proteins *Ann. Appl. Stat.* **2** 501–35
- [42] Kou S C and Xie X S 2004 Generalized Langevin equation with fractional Gaussian noise: subdiffusion within a single protein molecule *Phys. Rev. Lett.* **93** 180603
- [43] Kubo R 1966 The fluctuation-dissipation theorem *Rep. Prog. Phys.* **29** 255
- [44] Kulik A and Scheutzow M 2015 Generalized couplings and convergence of transition probabilities *Probab. Theory Relat. Fields* **171** 1–44
- [45] Kuo H H and Lee Y J 2011 Integration by parts formula and the Stein lemma on abstract Wiener space *Commun. Stoch. Anal.* **5** 10
- [46] Kupferman R 2004 Fractional kinetics in kac–zwanzig heat bath models *J. Stat. Phys.* **114** 291–326
- [47] Lampo T J, Stylianidou S, Backlund M P, Wiggins P A and Spakowitz A J 2017 Cytoplasmic RNA-protein particles exhibit non-Gaussian subdiffusive behavior *Biophys. J.* **112** 532–42
- [48] Lubelski A, Sokolov I M and Klafter J 2008 Nonergodicity mimics inhomogeneity in single particle tracking *Phys. Rev. Lett.* **100** 250602
- [49] Lysy M, Pillai N S, Hill D B, Forest M G, Mellnik J W, Vasquez P A and McKinley S A 2016 Model comparison and assessment for single particle tracking in biological fluids *J. Am. Stat. Assoc.* **111** 1413–26
- [50] Mason T G and Weitz D A 1995 Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids *Phys. Rev. Lett.* **74** 1250
- [51] Mattingly J C 2002 Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics *Commun. Math. Phys.* **230** 421–62
- [52] Mattingly J C, Stuart A M and Higham D J 2002 Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise *Stoch. Process. Appl.* **101** 185–232
- [53] McKinley S A and Nguyen H D 2018 Anomalous diffusion and the generalized Langevin equation *SIAM J. Math. Anal.* **50** 5119–60
- [54] McKinley S A, Yao L and Forest M G 2009 Transient anomalous diffusion of tracer particles in soft matter *J. Rheol.* **53** 1487–506
- [55] Meroz Y, Sokolov I M and Klafter J 2013 Test for determining a subdiffusive model in ergodic systems from single trajectories *Phys. Rev. Lett.* **110** 090601
- [56] Metzler R, Jeon J H, Cherstvy A G and Barkai E 2014 Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking *Phys. Chem. Chem. Phys.* **16** 24128–64
- [57] Morgado R, Oliveira F A, Batrouni G G and Hansen A 2002 Relation between anomalous and normal diffusion in systems with memory *Phys. Rev. Lett.* **89** 100601
- [58] Neuman K C and Nagy A 2008 Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy *Nat. Methods* **5** 491
- [59] Newby J, Schiller J L, Wessler T, Edelstein J, Forest M G and Lai S K 2017 A blueprint for robust crosslinking of mobile species in biogels with weakly adhesive molecular anchors *Nat. Commun.* **8** 833
- [60] Ottobre M and Pavliotis G A 2011 Asymptotic analysis for the generalized langevin equation *Nonlinearity* **24** 1629

- [61] Pavliotis G A 2014 *Stochastic Processes and Applications: Diffusion Processes, the Fokker–Planck and Langevin Equations* vol 60 (Berlin: Springer)
- [62] Pottier N 2003 Aging properties of an anomalously diffusing particule *Physica A* **317** 371–82
- [63] Revuz D and Yor M 2013 *Continuous Martingales and Brownian Motion* vol 293 (Berlin: Springer)
- [64] Sandev T and Tomovski Ž 2019 *Fractional Equations and Models: Theory and Applications* vol 61 (Berlin: Springer)
- [65] Tolić-Nørrelykke I M, Munteanu E L, Thon G, Oddershede L and Berg-Sørensen K 2004 Anomalous diffusion in living yeast cells *Phys. Rev. Lett.* **93** 078102
- [66] Viñales A D and Desposito M A 2006 Anomalous diffusion: exact solution of the generalized langevin equation for harmonically bounded particle *Phys. Rev. E* **73** 016111
- [67] Wang B, Kuo J, Bae S C and Granick S 2012 When Brownian diffusion is not Gaussian *Nat. Mater.* **11** 481
- [68] Weber S C, Spakowitz A J and Theriot J A 2010 Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm *Phys. Rev. Lett.* **104** 238102
- [69] Weber S C, Thompson M A, Moerner W E, Spakowitz A J and Theriot J A 2012 Analytical tools to distinguish the effects of localization error, confinement, and medium elasticity on the velocity autocorrelation function *Biophys. J.* **102** 2443–50
- [70] Wei M T, Zaorski A, Yalcin H C, Wang J, Hallow M, Ghadiali S N, Chiou A and Ou-Yang H D 2008 A comparative study of living cell micromechanical properties by oscillatory optical tweezers *Opt. Express* **16** 8594–603
- [71] Weinan E, Mattingly J C and Sinai Y 2001 Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation *Commun. Math. Phys.* **224** 83–106
- [72] Weiss M, Elsner M, Kartberg F and Nilsson T 2004 Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells *Biophys. J.* **87** 3518–24
- [73] Zhang H and Liu K K 2008 Optical tweezers for single cells *J. R. Soc. Interface* **5** 671–90
- [74] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford: Oxford University Press)