GIBBSIAN DYNAMICS AND THE GENERALIZED LANGEVIN EQUATION

DAVID P. HERZOG[1], JONATHAN C. MATTINGLY[2] AND HUNG D. NGUYEN[3]

ABSTRACT. We study the statistically invariant structures of the nonlinear generalized Langevin equation (GLE) with a power-law memory kernel. For a broad class of memory kernels, including those in the subdiffusive regime, we construct solutions of the GLE using a Gibbsian framework, which does not rely on existing Markovian approximations. Moreover, we provide conditions on the decay of the memory to ensure uniqueness of statistically steady states, generalizing previous known results for the GLE under particular kernels as a sum of exponentials.

1. INTRODUCTION

1.1. Overview.

We study the generalized Langevin equation

$$\begin{aligned} dx(t) &= v(t)\,dt, \\ dv(t) &= -v(t)\,dt - U'(x(t))\,dt - \int_{-\infty}^{t} K(t-r)v(r)\,dr\,dt + \sqrt{2}\,dW(t) + F(t)\,dt, \end{aligned} \tag{1.1}$$

describing the motion of a particle with position $x(t) \in \mathbb{R}$ and velocity $v(t) \in \mathbb{R}$ in a potential U. The particle is subject to a viscous friction force $-v(t)\,dt$ and a convolution term involving the *convolution kernel* K, modeling a thermal drag force with memory effects. By the *fluctuation-dissipation* relation, both of these forces are respectively balanced by stochastic processes $W(t)$ and $F(t)$, where $W(t)$ is a standard one-dimensional Brownian motion and $F(t)$ is a mean-zero stationary Gaussian process with covariance given by

$$\mathbb{E}[F(t_1)F(t_2)] = K(|t_1 - t_2|), \quad \text{for all } t_1, t_2 \in \mathbb{R}. \tag{1.2}$$

In the absence of memory effects, that is setting $K \equiv 0$ and $F \equiv 0$ in (1.1) above, large-time properties of the resulting Markovian system are well-understood, in the sense that under general conditions on the potential U, it is known that the system admits a unique ergodic invariant measure $\pi(x, v)$ which is exponentially attractive and whose formula is given by

$$\pi(dx, dv) \propto \exp(-H(x, v))\,dx\,dv, \tag{1.3}$$

where

$$H(x, v) = \frac{v^2}{2} + U(x)$$

denotes the Hamiltonian of the system. For example, see [4, 5, 14, 21, 29, 34] and the references within. When $K \equiv 0$ and $F \equiv 0$, one can equally speak of stationary solutions of (1.1) as they are in one-to-one correspondence with the invariant measures, namely the fixed points of the Markov semigroup generated by (1.1) without the memory terms. On the other hand, in the presence of memory in (1.1), comparatively much less is known about both the existence and uniqueness of statistically stationary states under general conditions on K. The goal of this paper is to make progress on bridging this gap between the standard Langevin equation ($K \equiv 0, F \equiv 0$) and its generalized counterpart (1.1) with memory.

In general, there is no Markovian dynamics on $\mathbb{R}^2$ associated with (1.1); and hence, no directly analogous concept of an invariant measure. Thus, we are left to study the stationary solutions of (1.1) on $\mathbb{R}^2$ as this concept remains well-defined. There are, however, special cases where one can still define a convenient Markov process associated to (1.1) on an extended, but still finite dimensional, statespace. When $K(t)$ can be written as a finite sum of exponentials; that is,

$$K(t) = \sum_{k=1}^{n} c_k e^{-\lambda_k t}, \tag{1.4}$$

for some constants $c_k \in \mathbb{R}, \lambda_k > 0$, one can augment the resulting system (1.1) by a finite number of auxiliary variables to produce a Markov process on a higher, but finite-dimensional space. This corresponding finite-dimensional system was studied rigorously in [28, 29]. There, under general hypotheses on U, it was shown that the system is uniquely ergodic and the marginal invariant distribution of the pair (x, v) is precisely π as in (1.3) [12, 28, 29]. Because the sum above is finite, however, it cannot describe a kernel with power-law decay, i.e., a kernel $K(t)$ satisfying

$$K(t) \sim t^{-\alpha} \quad \text{as} \quad t \to \infty, \tag{1.5}$$

for some $\alpha > 0$. Subsequently, this approach was extended to handle such memory kernels by writing K as an infinite-sum of exponentials ($n \to \infty, c_k = c_k(\lambda_k) > 0$ in (1.4)) [11]. See Remark 2.12 below. The resulting dynamics is an infinite-dimensional Markov process on an Sobolev-like space, and it was shown that there exists an explicit invariant probability measure whose (x, v)–marginal agrees with (1.3). This is true for memory kernels in this specific form regardless of the memory decay rate $\alpha > 0$ as in (1.5). However, to establish uniqueness of this measure, the restriction $\alpha \in (1, \infty)$ was imposed leaving out the important *subdiffusive regime* of $\alpha \in (0, 1)$ (see the discussion in Section 1.2 below). One of our goals here is to push through this threshold.

We will study (1.1) both when the Gaussian forcing satisfies the structural assumption in (1.4) (with $n = \infty$) and when it does not. For general stationary Gaussian forcing F, there is not necessarily a Markovian dynamics associated to (1.1).[1] Hence, we lack a natural notion of an Markov invariant measure and study the stationary solutions of (1.1) instead. We give general conditions guaranteeing that there is at most one stationary solution. Although there is no Markov formulation of the stochastic dynamics, there is however a natural skew-flow on the *infinite past* $C((-\infty, 0], \mathbb{R}^2)$ of the trajectories of (x, v) fibered over the Gaussian forcing F.

When (1.4) holds with $n < \infty$ or $n = \infty$, then there is a natural Markovian formulation of the stochastic dynamics [11, 28, 29]. We will study a different Markovian formulation than used in those works. The assumption in (1.4) implies that $F(t)$ can be constructed as a functional of a (possibly) infinite collection of independent Brownian Motions on the time interval $(-\infty, t]$. We formulate a Markovian dynamics which takes as its state space the trajectories of (x, v) on the *infinite past* $C((-\infty, 0], \mathbb{R}^2)$ and the infinite past of the collection of independent Brownian Motions used to construct F. We show that $\alpha > 1/2$ this dynamics has at most one invariant measure; or equivalently, at most one stationary solution, cf. Theorem 4.3.

Remark 1.1. Gibbsian Dynamics: It is possible to enlarge the statespace of any dynamics to make it Markovian. In the extreme, by making the statespace the entire trajectory $\{\big(x(t), v(t), F(t)\big) : t \in (-\infty, \infty)\}$, the dynamics is simply the shift map $\theta_t \colon (x, v, F) \mapsto \big(x(\,\cdot\, + t), v(\,\cdot\, + t), F(\,\cdot\, + t)\big)$. At this level of generality, the fact that the dynamics is Markovian provides little useful structure. However, our setting has more structure.

In the continuous time Markov setting, the distribution of infinitesimal increments is a function the current state of the process. In the Gibbsian setting, as envisioned in [19, 36], the distribution of infinitesimal increments is a function of the entire past. We will return to this setting in Section 3.3. The term Gibbssian comes from the dynamics being dictated, not by a compatible family of Markov

[1] One can always consider as the state space the path space of a process on the time interval $(-\infty, \infty)$. The dynamics is then the deterministic shift of the trajectories. Lifting of the deterministic process to pathspace is not the type of stochastic Markov dynamics we seek.

measures (depending only on the boundary data in space-time), but rather a compatible family of Gibbs measures (in the general sense of [8]).

1.2. **Physical motivation.** It is important to note some of the physical reasons for considering memory kernels K in general, and in the power-law regime in particular. The standard Langevin equation is commonly used to describe microparticle motion embedded in Newtonian fluids, which amounts to the implicit assumption that there is no time correlation between the foreign microparticles and the thermally fluctuating fluid molecules. Following Newton's Second Law [29], the two-dimensional Langevin equation has the form (1.1) with $K \equiv 0$ and $F \equiv 0$. On the other hand, for viscoelastic fluids, elasticity induces time correlation between foreign particles and fluid molecules, leading to memory effects. Thus the standard Langevin equation is not sufficient to describe the motion of the particles suspended in the fluid. In order to capture such phenomena, the generalized Langevin equation (1.1) with general K was introduced in [16, 25, 26] and later popularized in [18].

It is known that the unconstrained GLE (i.e. $U \equiv 0$ in (1.1)) exhibits *anomalous diffusion*; that is, the mean-squared displacement $\mathbb{E}x(t)^2$ may not be asymptotically proportional to t as $t \to \infty$. In fact, it was shown in [6, 22] that when $K \in L^1(\mathbb{R})$, the unconstrained GLE is *asymptotically diffusive*, i.e., $\mathbb{E}x(t)^2 \sim t$ as $t \to \infty$. Otherwise, if $K(t) \sim t^{-\alpha}$, $\alpha \in (0,1)$, then the unconstrained GLE is *asymptotically subdiffusive*, i.e. $\mathbb{E}x(t)^2 \sim t^\alpha$ and when $\alpha = 1$, there is a transition phase between *diffusion* and *subdiffusion*, i.e., $K(t) \sim t^{-1}$ implies $\mathbb{E}x(t)^2 \sim t/\log(t)$ as $t \to \infty$. For viscoelastic fluids, the subdiffusive regime is observed in experiments [9, 17, 23, 24, 30, 31, 32], which is why we are primarily interested in the scenario where K has a power-law decay rate $\alpha \in (0, 1]$.

The rest of the paper is organized as follows. In Section 2, we introduce assumptions and briefly state the well-posedness result for (1.1). In Section 3, we see that the dynamics (1.1) induces a skew-flow on the *skew* path space. Section 4 discusses the associated stationary solution(s) for this dynamics. Furthermore, we prove our main result on the uniqueness of the associated stationary measures in this section. The argument proving uniqueness, in particular, makes use of some auxiliary results collected and proved in Section 5. In Section 6, we establish the existence of a stationary measure when the kernel can be written as an infinite sum of exponentials. In Appendix A, we establish the well-posedness result in detail. In Appendix B, we prove a technical result which allows us to bound the expected value of the maximum of $F(t)$ over finite intervals of time. This result is employed in the proof of well-posedness.

2. Assumptions and well-posedness

2.1. **Well-posedness.** We begin by clarifying what we mean by a solution of (1.1). Throughout, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ where the set Ω is endowed with a probability measure $\mathbb{P}$ and a filtration of sigma-algebras $\{\mathcal{F}_t : t \in \mathbb{R}\}$.

Definition 2.1 (Solution on $(-\infty, \infty)$)**.** A *(weak) solution* to (1.1) on the time interval $(-\infty, \infty)$ is a probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ on which a triple of stochastic processes (ξ, F, W) is defined so that the following conditions are satisfied:

1. (1) $\xi(t) = (x(t), v(t))$, $F(t)$ and $W(t)$ are all stochastic processes adapted to the filtration $\{\mathcal{F}_t\}$.
2. (2) $F(t, \omega)$ is a stationary Gaussian process with mean zero and covariance K in the sense of (1.2) and $W(t, \omega)$ is a standard, two-sided Brownian Motion both with respect to $\{\mathcal{F}_t\}$ such that F and W are independent.
3. (3) With probability one, the triple (ξ, F, B) solves (1.1); that is, with probability one, for all $t_0, t_1 \in \mathbb{R}$ with $t_0 < t_1$ we have

$$
\begin{aligned}
x(t_1) - x(t_0) &= \int_{t_0}^{t_1} v(t)\,dt,\\
v(t_1) - v(t_0) &= -\int_{t_0}^{t_1} \Big[v(t) + U'(x(t)) + \int_{-\infty}^{t} K(t-r)v(r)\,dr\Big]\,dt \\
&\qquad + \sqrt{2}\big(W(t_1) - W(t_0)\big) + \int_{t_0}^{t_1} F(t)\,dt.
\end{aligned} \tag{2.1}
$$

Definition 2.2 (Solution with an initial past)**.** A (weak) solution to (1.1) on the time interval (T_0, T_1) with $T_0 \in \mathbb{R}$ and $T_1 \in \mathbb{R} \cup \{\infty\}$ with *initial past* $\xi_0 = (x_0, v_0) \in C((-\infty, T_0], \mathbb{R}^2)$ satisfies the same conditions as in Definition 2.1 but the stochastic processes need only be defined on the time interval (T_0, T_1) with the exception of $\xi = (x, v)$ which is defined on $(-\infty, T_1)$ with $\xi(t) = \xi_0(t)$ for $t \in (-\infty, T_0]$. Additionally, (2.1) need only hold for $t_0, t_1 \in (T_0, T_1)$.

Remark 2.3. In this paper, we will prove strong existence of solutions on $[T_0, \infty)$ given an initial past $\xi_0 = (x_0, v_0)$ belonging to an appropriate subclass of $C((-\infty, T_0]; \mathbb{R}^2)$. Moreover, we will also establish weak uniqueness, which together is stronger than weak existence and weak uniqueness.

Throughout, we will employ the following assumption on the potential U in (1.1).

Assumption 2.4. *The potential $U\colon \mathbb{R} \to \mathbb{R}$ is such that $U \in C^3(\mathbb{R})$, $\int_{\mathbb{R}} |U'(x)|e^{-U(x)}dx < \infty$ and the global estimate holds*

$$
b(U(x) + 1) \ge |x|^{1+\delta} \quad \text{for all} \quad x \in \mathbb{R},
$$

for some constants $b > 0$ and $\delta \in (0, 1)$.

Remark 2.5. The first two conditions on U are not directly used in this paper. They were previously used in [11, Theorem 7] to construct an explicit invariant measure for the Markov system (6.3) below. We then will use this result to construct a stationary measure for the dynamics (1.1) in Section 6.

We also use the following condition on the memory kernel.

Assumption 2.6. *$K \in C^1([0, \infty); [0, \infty))$ and there exists $\widetilde{K} \in C([0, \infty))$ for which*

$$
\sup_{s \ge 0} \frac{K(t+s)}{K(s)} = \widetilde{K}(t) \quad \text{for all} \quad t \ge 0.
$$

In order to state our main existence and uniqueness result, for $t \in \mathbb{R}$ let

$$
\mathcal{C}(-\infty, t] := \Big\{(x, v) \in C((-\infty, t], \mathbb{R}^2) \, : \, \int_{-\infty}^{t} K(t-r)|v(r)|dr < \infty\Big\}. \tag{2.2}
$$

Proposition 2.7. *Suppose that Assumption 2.4 and Assumption 2.6 are satisfied. Then there exists a subset $\mathcal{K} \subset C((-\infty, \infty); \mathbb{R})$ so that $\mathbb{P}(F \in \mathcal{K}) = 1$ and for every $t_0 \in \mathbb{R}$, $F \in \mathcal{K}$ and every initial condition $\xi_0 = (x_0, v_0) \in \mathcal{C}(-\infty, t_0]$, there exists a unique solution $\xi = (x, v)$ with initial past ξ_0 on the time interval $[t_0, \infty)$ such that $\xi \in \mathcal{C}(-\infty, t]$ for all $t \ge t_0$. Furthermore, we have the energy estimate*

$$
\begin{aligned}
&\mathbb{E} \sup_{t_0 \le r \le t} H(x(r), v(r)) \\
&\qquad \le \Big[H(x_0(t_0), v_0(t_0)) + \Big(\int_{-\infty}^{t_0} K(t_0 - r)|v_0(r)|dr\Big)^2 + \mathbb{E} \sup_{t_0 \le r \le t} F(r)^2 + 1\Big] e^{c(t_0, t)},
\end{aligned} \tag{2.3}
$$

where we recall that $H(x, v) = \frac{1}{2}v^2 + U(x)$.

The proof of Propoosition 2.7 is given later in Appendix A.

Remark 2.8. For a general centered stationary Gaussian process $F(t)$, it is not immediately obvious that for all $t_0 < t$

$$\mathbb{E} \sup_{t_0 \le r \le t} F(r)^2 < \infty. \tag{2.4}$$

In Appendix B, we will make use of the condition that $K \in C^1$, cf. Assumption 2.6, to show that this is indeed the case for the process $F(t)$.

2.2. **Structural assumptions on the noise.** At times, we will further assume that memory kernel K has the following specific form previously employed in [11].

Assumption 2.9. *There exists continuously differentiable functions $J_\ell \colon [0,\infty) \to [0,\infty), \ell \ge 1$, so that the stationary Gaussian forcing $F(t)$ can be represented as*

$$F(t) = \sum_{\ell \ge 1} \int_{-\infty}^{t} J_\ell(t-s) dB^{(\ell)}(s), \tag{2.5}$$

where $\{B^{(\ell)} : \ell \ge 1\}$ is a collection of mutually independent standard two-sided Brownian motions. Furthermore,

$$t \mapsto \sum_{\ell \ge 1} \int_0^\infty J_\ell(t+r) J_\ell(t) dt,$$

is continuously differentiable.

Remark 2.10. Assumption 2.9 together with the fluctuation-dissipation relation (1.2) immediately imply that the memory kernel $K(t)$ is continuously differentiable and of the form

$$K(r) = \sum_{\ell \ge 1} K_\ell(t) \quad \text{where} \quad K_\ell(t) = \int_0^\infty J_\ell(t+r) J_\ell(t) dt.$$

We will also need some structure on the decay of the kernel at infinity.

Assumption 2.11. *There exist constants $t_* > 0$, $C > 0$ and $\alpha > 1/2$ such that*

$$K(t) \le C t^{-\alpha} \quad \textit{for all} \quad t \ge t_*.$$

Remark 2.12. When F is of the form (2.5), an example of particular interest is when J_ℓ, $\ell \ge 1$, is given by

$$J_\ell(t) = \sqrt{2 c_\ell \lambda_\ell} e^{-\lambda_\ell t},$$

where

$$c_\ell = \frac{1}{\ell^{1+\alpha\beta}} \qquad \text{and} \qquad \lambda_\ell = \frac{1}{\ell^\beta}, \tag{2.6}$$

for some constants $\alpha > 0, \beta > 1$. In this case,

$$K(t) = \sum_{\ell \ge 1} c_k e^{-\lambda_\ell t}, \tag{2.7}$$

and one can show that [1, Example 3.2]

$$K(t) \sim t^{-\alpha}, \quad t \to \infty.$$

Hence, K is a power-law memory kernel which clearly satisfies Assumptions 2.6 and 2.11.

Remark 2.13. Note that if we first suppose that K is of the form (2.7), Doob's Theorem [7] and the fluctuation-dissipation relation (1.2) together imply that F must be of the form

$$F(t) = \sum_{\ell \ge 1} \sqrt{2 \lambda_\ell c_\ell} \int_{-\infty}^{t} e^{-\lambda_\ell (t-r)} dB^{(\ell)}(r), \tag{2.8}$$

where in the above, $\{B^{(\ell)}\}_{\ell\geq 1}$ are two-sided, independent standard Brownian motions.

When Assumption 2.9 holds, we arrive at the following form for the GLE

$$\begin{aligned} d\,x(t) &= v(t)\,dt, \\ d\,v(t) &= -v(t)\,dt - U'(x(t))\,dt - \sum_{\ell\geq 1}\int_{-\infty}^{t} K_\ell(t-r)v(r)dr\,dt \\ &\qquad + \sum_{\ell\geq 1}\int_{-\infty}^{t} J_\ell(t-r)dB^{(\ell)}(r)\,dt + \sqrt{2}\,dW(t), \end{aligned} \tag{2.9}$$

where W is a standard, two-sided, real-valued Brownian motion independent of the collection $\{B^{(\ell)}\}_{\ell\geq 1}$ and K_ℓ is as in Remark 2.10.

3. Solutions on Pathspace

Since we often work on the phase space $C(\mathbb{R};\mathbb{R}^2)$ and its subspaces, we use the topology on $C(\mathbb{R};\mathbb{R}^2)$ defined in the follow sense: A sequence $\{g_n\}\in C(\mathbb{R};\mathbb{R}^2)$ is said to *converge* to $g\in C(\mathbb{R};\mathbb{R}^2)$ if the convergence holds in the sup norm on any bounded time interval. That is, for all fixed $T>0$,

$$\sup_{t\in[-T,T]} |g_n(t)-g(t)|\to 0, \quad \text{as } n\to\infty.$$

The closed sets in $C(\mathbb{R};\mathbb{R}^2)$ are then defined with respect to the above mode of convergence, hence inducing the corresponding topology of open sets as well as the Borel sigma algebra of subsets of $C(\mathbb{R};\mathbb{R}^2)$.

Having fixed the topology in $C(\mathbb{R};\mathbb{R}^2)$, in Section 3.1 below, we discuss the notion of skew-flows generated by (1.1).

3.1. Skew-flow and kernel.

Associated to the dynamics (1.1) is a skew product flow. That is, given a realization of F and W on the time interval $(-\infty,\infty)$, we define the family maps

$$\varphi^{F,W}_{t_0,t}:\mathcal{C}(-\infty,t_0]\to\mathcal{C}(-\infty,t],\ \ t_0\leq t, \tag{3.1}$$

as the extension of an initial past $\xi_0\in\mathcal{C}(-\infty,t_0]$ to a function in $\mathcal{C}(-\infty,t]$ by appending to the front of ξ_0 the solution (1.1) on the time interval $[t_0,t]$ with initial past ξ_0 and random forcing W and F. When ξ_0 is deterministic, $\varphi^{F,W}_{t_0,t}\xi_0$ is a random path adapted to

$$\mathcal{F}_{t_0,t}=\sigma(F(r),W(r)-W(t_0):r\in[t_0,t]) \quad\text{with}\quad (\varphi^{F,W}_{t_0,t}\xi_0)(r)=\xi_0(r) \quad\text{for}\quad r\leq t_0.$$

Observe that if θ_t denotes the shift map in the space of trajectories, $\theta_t\varphi^{F,W}_{0,t}\xi_0\colon\mathcal{C}(-\infty,0]\to\mathcal{C}(-\infty,0]$, so that the skew-flow S_t defined by

$$S_t\colon(\xi_0,F,W)\mapsto(\theta_t\varphi^{F,W}_{0,t}\xi_0,\theta_tF,\theta_tW-W(t)), \tag{3.2}$$

is a random semi-flow on the space $\mathcal{C}(-\infty,0]\times C((-\infty,\infty);\mathbb{R}^2)$. In particular $S_{s+t}=S_sS_t$.

Next we define the skew transition kernel R^F_t on $\mathcal{C}(-\infty,0]$ by taking the law of $\theta_t\varphi^{F,W}_{0,t}\xi_0$ conditioned on ξ_0 and F; namely,

$$R^F_t(\xi_0,A):=\mathbb{P}(\theta_t\varphi^{F,W}_{0,t}\xi_0\in A\,|\,\xi_0,F),$$

for $(\xi_0,F)\in\mathcal{S}_{\text{skew}}:=\mathcal{C}(-\infty,0]\times C((-\infty,\infty);\mathbb{R})$, and $A\subset\mathcal{C}(-\infty,0]$ Borel. Observe we have the following skew structure stemming from (3.2)

$$R^F_tR^{\theta_tF}_s=R^F_{t+s},$$

or more explicitly,

$$R^F_{t+s}(\xi_0, A) = \int_{\mathcal{C}(-\infty,0]} R^F_t(\xi_0, d\zeta) R^{\theta_t F}_s(\zeta, A).$$

3.2. **Markovian kernel.** Looking at (2.9), we see that when Assumption 2.9 is enforced, we can consider a solution to be a triple of stochastic processes (ξ, W, B) where ξ and W are as before but $B = \{B^{(\ell)}\}_{k\geq 1}$ is a countable collection of standard two-sided independent Brownian Motions independent of W. We can then define a map $\psi^W_t : \mathcal{S}(-\infty, 0] \to \mathcal{S}(-\infty, t]$ where

$$\mathcal{S}(-\infty, t] := \mathcal{C}(-\infty, t] \times C((-\infty, t], \mathbb{R})^{\mathbb{N}}, \tag{3.3}$$

and $\psi^W_t(\xi_0, B_0)$ is equal to the pair (ξ, B) obtained by continuing the Brownian motions B_0 over the interval $[0, t]$ and then extending ξ over the same interval by flowing (2.9) using F as in (2.5) with the Brownian Motions in B. We again have a skew-flow defined for $(\xi, B) \in \mathcal{S}(-\infty, 0]$ and $W \in C((-\infty, \infty); \mathbb{R})$ by

$$(\xi, B, W) \mapsto (\theta_t \psi^W_t(\xi, B), \theta_t W)\,.$$

In contrast to Section 3.1, where the stochastic processes over which the skew-flow is fibered, namely (W, F), this skew-flow is fibered over a process, namely W, whose future increments are independent of its past increments. Thus, we can obtain a Markov kernel by averaging over the randomness in W. We cannot average over the randomness in B as the increment added to ξ over the time interval $[0, t]$ depends on the entire history of B back to time $-\infty$.

With these considerations, we define the Markov kernel P_t on $\mathcal{S}(-\infty, 0]$ by

$$P_t\big((\xi_0, B_0), A\big) = \mathbb{P}(\theta_t \psi^W_t(\xi_0, B_0) \in A \,|\, \xi_0, B_0), \tag{3.4}$$

for $(\xi_0, B_0) \in \mathcal{S}(-\infty, 0]$ and $A \subset \mathcal{S}(-\infty, 0]$ Borel.

3.3. **Gibbsian Process.** In the simplest setting, a Gibbsian SDE is given by

$$d\xi(t) = f(\theta_t \xi_{(-\infty,t]})dt + g(\theta_t \xi_{(-\infty,t]})dW(t), \tag{3.5}$$

where θ_t is the shift, $\xi(t) \in \mathbb{R}^d$, $\xi_{(-\infty,t]} \in C((-\infty, t]; \mathbb{R}^d)$, $\theta_t \xi_{(-\infty,t]} \in C((-\infty, 0]; \mathbb{R}^d)$ and $f, g \colon C((-\infty, 0]; \mathbb{R}^d) \to \mathbb{R}^d$ are the generators of the process. While the resulting Itô process [15] is not a Markov proces on $\mathbb{R}^d$, it generates a well defined dynamics and a collection of compatible distributions on the future conditioned on the past. If $F \equiv 0$, then (1.1) is in the form of (3.5) and the dynamics is determined by f and g. This process can be seen as a Markov process on $C((-\infty, 0]; \mathbb{R}^d)$ as in Sections 3.1 and 3.2 and Remark 1.1. The ideas of this note can be generalized to the solution of (3.5) under appropriate assumptions as in [2, 13, 19, 20, 35]. Here we have the extra complication of the forcing F. Hence in general one must view the process (1.1) under the consideration as a Gibbsian skew-flow in the spirit of Section 3.1. In the restricted setting of Section 3.2, we arrive at a standard Gibbsian prossess if the history of the generating Brownian Motions are included in the statespace.

3.4. **Solutions on the time interval** $[0, \infty)$**.** Proposition 2.7 gives a finite-time existence and uniqueness result for initial pasts in $\mathcal{C}(-\infty, 0]$. Thus solutions do not blow up in finite time, but it is possible that they may tend to ∞ as $t \to \infty$. Hence this fact induces a well-defined mapping

$$\varphi^{F,W}_\infty \colon \mathcal{C}(-\infty, 0] \to C([0, \infty); \mathbb{R}^2),$$

but it is still possible that

$$\mathbb{P}(|\varphi^{F,W}_\infty(\xi_0)(s)| \to \infty \text{ as } s \to \infty) > 0.$$

In the next section, we will consider the large-time behavior of the system, in particular the existence and uniqueness of stationary solutions.

Because the mapping $\varphi_\infty^{F,W}$ makes sense, we can define a family of kernels $Q^F_{[0,\infty)}$ on the *infinite future* by

$$Q^F_{[0,\infty)}(\xi_0, A) = \mathbb{P}(\pi_{[0,\infty)}\varphi_\infty^{F,W}(\xi_0) \in A \,|\, F, \xi_0), \tag{3.6}$$

for initial pasts $\xi_0 \in \mathcal{C}(-\infty, 0]$ and Borel sets $A \subset C([0,\infty) : \mathbb{R}^2)$. Here, $\pi_{[0,\infty)}$ denotes the projection of the trajectory onto the time interval $[0,\infty)$. While R_t^F captures the effect of starting from an inial past at time $-t$ and flowing forward to time 0, Q^F captures the distribution on the infinite future starting from ξ_0 at time 0.

4. Stationary Solutions and Invariant Measures

Recall that the stochastic process (ξ, F) on the time interval $(-\infty,\infty)$ is *stationary* if for any finite collection of times $t_1, \cdots, t_n \in (-\infty,\infty)$ the distribution of the random vector

$$\big((\xi(t_1+s), F(t_1+s)), \dots, (\xi(t_n+s), F(t_n+s))\big),$$

is independent of $s \in \mathbb{R}$. Recalling that θ_t denotes the shift mapping on the space of trajectories, stationarity is equivalent in our setting to the distribution of the path $\theta_t(\xi, F)$ being independent of t.

4.1. For the skew kernel R_t^F.

A stationary solution (ξ, F) on the time interval $(-\infty,\infty)$ always generates a skew-invariant measure μ^F on Borel subsets of $\mathcal{C}(-\infty, 0]$, which is invariant for the kernel R_t^F and indexed by a realization $F \in C((-\infty,\infty);\mathbb{R})$:

$$\mu^F R_t^F = \mu^{\theta_t F},$$

where we define the measure $\mu^F R_t^F$ for $A \subset \mathcal{C}(-\infty, 0]$ Borel by

$$\mu^F R_t^F(A) = \int_{\mathcal{C}(-\infty,0]} R_t^F(\xi, A)\mu^F(d\xi)\,.$$

Let $\mathrm{Law}(\xi, F)$ denote the law of stationary solution (ξ, F) on the space $\mathcal{S}_{\text{skew}}$. The disintegration of $\mathrm{Law}(\xi, F)$ relative to $\mathrm{Law}(F)$, restricted to $\mathcal{C}(-\infty, 0]$, is the desired skew-invariant μ^F. Here, the skew invariance follows from the stationarity of (ξ, F). On the other hand, given a skew-invariant measure μ^F on $\mathcal{C}(-\infty, 0]$, let $\tilde{\mu}^F$ be the extension of μ^F to the time interval $(-\infty,\infty)$ using the dynamics $\varphi_{0,t}^{F,W}$. Then $\tilde{\mu}^F(d\xi)\mathrm{Law}(F)(df)$ is the law of the desired stationary process (ξ, F).

4.2. For the Markov kernel P_t.

When Assumption 2.9 holds, recall that (ξ, B) evolves as a Markov process on the state space

$$\mathcal{S}(-\infty, 0] := \mathcal{C}(-\infty, 0] \times C((-\infty, 0], \mathbb{R})^{\mathbb{N}},$$

under the Markov kernel P_t defined in (3.4). In this setting, there is a one-to-one correspondence between stationary solutions on the time interval $(-\infty,\infty)$ and invariant probability measures μ for P_t. Given an invariant probability measure μ for P_t on $\mathcal{S}(-\infty, 0]$, we can create a stationay measure $\tilde{\mu}$ on the interval $(-\infty,\infty)$ by flowing the dynamics forward from μ by the map ψ_t^W defined in Section 3.2 from a random initial past distributed according to μ and then taking the measure obtained on $\mathcal{S}(-\infty,-\infty)$ by averaging over the realization of W.

Conversely, given a stationary solution on $\mathcal{S}(-\infty,\infty)$ then we can simply restrict the distribution to a measure on $\mathcal{S}(-\infty, 0]$. The resulting measure will be invariant for the Markov Kernel P_t.

4.3. **Existence and uniqueness of stationary measures.** Recalling the space $\mathcal{C}(-\infty,t]$ defined in (2.2), for $\varrho>0$ we introduce the following subset of *moderate growth*:

$$\mathcal{C}_\varrho(-\infty,t]=\left\{(x,v)\in\mathcal{C}(-\infty,t]:\sup_{r\le t}\frac{|x(r)|}{1+|r|^\varrho}<\infty\right\}, \tag{4.1}$$

and define

$$\mathcal{C}_\varrho(-\infty,\infty)=\bigcup_{n\in\mathbf{Z},n\ge 0}\mathcal{C}_\varrho(-\infty,n].$$

Our main result concerning the existence of an invariant measure for the Markov kernel P_t is the following theorem whose proof is deferred to Section 6.

Theorem 4.1. *Suppose that U satisfies Assumption 2.4 and that Assumption 2.9 is satisfied with the choice of J_ℓ as in Remark 2.12. Then there exists an invariant measure μ_* for P_t defined in* (3.4)*. Moreover, for every $\varrho>0$,*

$$\mu_*(\mathcal{C}_\varrho(-\infty,0])=1. \tag{4.2}$$

Remark 4.2. The proof of Theorem 4.1 relies on constructing an explicit invariant measure for an infinite-dimensional auxiliary Markovian system. A good Lyapunov-type estimate for the equation (1.1) which would ensure the abstract existence of such a measure in more generality is currently unavailable. It is thus left as an open problem to determine whether (4.2) always holds for any invariant measure μ.

The following is our uniqueness result which pairs with the existence result given in Theorem 4.1. However, it is worth noting that the uniqueness result applies in many settings where we do not know that there exists a stationary measure.

Theorem 4.3. *Suppose that U satisfies Assumption 2.4 and that the memory kernel K satisfies Assumptions 2.6 and 2.11. For every $\varrho<\alpha-1/2$, the skew dynamics S_t admits at most one stationary solution (ξ,F) such that $\mathrm{Law}(\xi\,|\,F)$ is supported in $\mathcal{C}_\varrho(-\infty,\infty)$.*

The following corollary is an immediate result of the Theorem 4.3 when we are in the Markovian setting discussed in Section 3.2 and Section 4.2.

Corollary 4.4. *When Assumption 2.9 holds in addition to the assumptions of Theorem 4.3, there exists at most one invariant measure supported on*

$$\mathcal{S}_\varrho(-\infty,0]:=\{(\xi,B)\in\mathcal{S}(-\infty,0]:\xi\in\mathcal{C}_\varrho(-\infty,0]\},$$

for the Markov semigroup on that space discussed in Section 3.2.

The proof of Theorem 4.3 makes use of a coupling argument employed in [2, 10, 13, 19, 20, 35, 36] to show that starting from two distinct initial history paths, the time averages of their solutions in the future must converge to the same place, hence yielding uniqueness of a given stationary measure. Two of the main ingredients in the coupling argument are the following two results to be proved in the next section.

Proposition 4.5. *Under the hypotheses of Theorem 4.3, for any stationary solution (ξ,F) of* (1.1)*, the marginal of $\mathrm{Law}(\xi\,|\,F)$ at any fixed time t is equivalent to Lebesgue measure on $\mathbb{R}^2$.*

Proposition 4.6. *Under the hypotheses of Theorem 4.3, let ξ_0 and $\widetilde{\xi}_0$ be two initial pasts in $\mathcal{C}_\varrho(-\infty,0]$ such that $\xi_0(0)=\widetilde{\xi}_0(0)$. Then for almost every realization of F, the measures $Q^F_{[0,\infty)}(\xi_0,\,\cdot\,)$ and $Q^F_{[0,\infty)}(\widetilde{\xi}_0,\,\cdot\,)$ are equivalent.*

Given these two results, we can now conclude Theorem 4.3.

Proof of Theorem 4.3. We first fix some notation. Given a set $A \subset C(\mathbb{R};\mathbb{R}^2)$, a measure ν on Borel subsets of $C(\mathbb{R};\mathbb{R}^2)$ and a time interval $\mathrm{T} \subset \mathbb{R}$, we denote by $\pi_{\mathrm{T}} A$ and $\pi_{\mathrm{T}}\nu$ to be respectively the projection of A and ν on T. In other words, letting $\pi_{\mathrm{T}}\xi$ be the projection of a trajectory ξ onto the time interval T, we set

$$\pi_{\mathrm{T}} A = \{\pi_{\mathrm{T}}\xi : \xi \in A\},$$

and for any Borel set $B \subset C\big(\mathrm{T};\mathbb{R}^2\big)$,

$$\pi_{\mathrm{T}}\nu(B) := \nu(\{\xi \in C(\mathbb{R};\mathbb{R}^2) : \pi_{\mathrm{T}}\xi \in B\}).$$

Let (ξ_1, F_1) and (ξ_2, F_2) be two stationary solutions of (1.1). Without lost of generality, we may assume that $\mathrm{Law}(\xi_1, F_1)$ and $\mathrm{Law}(\xi_2, F_2)$ are ergodic by ergodic decomposition. As discussed in Section 4.1, we can disintegrate $\mathrm{Law}(\xi_i, F_i)$ into $\mathrm{Law}(\xi_i \,|\, F)$ relative to $\mathrm{Law}(F)$ since $\mathrm{Law}(F) = \mathrm{Law}(F_1) = \mathrm{Law}(F_2)$. Letting $\nu_i = \mathrm{Law}(\xi_i \,|\, F)$, $i = 1, 2$, we aim to prove $\nu_1 = \nu_2$ assuming ν_1 and ν_2 are supported in $\mathcal{C}_\varrho(-\infty,\infty)$.

Fixing an arbitrary bounded function $\phi : C([0,\infty);\mathbb{R}^2) \to \mathbb{R}$ which only depends on some compact set of time, Birkhoff's Ergodic Theorem implies that there exists a set

$$A_i \subset \mathcal{C}_\varrho(-\infty,\infty),$$

such that $\nu_i(A_i) = 1$ and for every $\xi \in A_i$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(\pi_{[0,\infty)}\theta_t\xi)dt = \int \phi(\pi_{[0,\infty)}\xi)\nu_i(d\xi) =: \phi_i. \tag{4.3}$$

It suffices to prove that $\phi_1 = \phi_2$. To this end, for each $\zeta \in C\big((-\infty,0];\mathbb{R}^2\big)$, we set

$$B_i(\zeta) = \{\pi_{[0,\infty)}z : z \in A_i, \pi_{(-\infty,0]}z = \zeta\}.$$

Since ν_i is supported in A_i, it is clear that

$$\pi_{[0,\infty)}\nu_i(\pi_{[0,\infty)}A_i) = \nu(\{z \in A_i : \pi_{[0,\infty)}z \in \pi_{[0,\infty)}A_i\}) = \nu_i(A_i) = 1.$$

On the other hand, recalling $Q^F_{[0,\infty)}$ is the future law as in (3.6), observe that

$$\begin{aligned} 1 = \pi_{[0,\infty)}\nu_i(\pi_{[0,\infty)}A_i) &= \int_{\pi_{(-\infty,0]}A_i} Q^F_{[0,\infty)}\big(\,\zeta, \pi_{[0,\infty)}A_i\,\big)\pi_{(-\infty,0]}\nu_i(d\zeta), \\ &= \int_{\pi_{(-\infty,0]}A_i} Q^F_{[0,\infty)}\big(\,\zeta, \{\pi_{[0,\infty)}z : z \in A, \pi_{(-\infty,0]}z = \zeta\}\,\big)\pi_{(-\infty,0]}\nu_i(d\zeta), \\ &= \int_{\pi_{(-\infty,0]}A_i} Q^F_{[0,\infty)}\big(\,\zeta, B_i(\zeta)\,\big)\pi_{(-\infty,0]}\nu_i(d\zeta), \end{aligned}$$

We then conclude that for almost every $\zeta \in \pi_{(-\infty,0]}A_i$ with respect to $\pi_{(-\infty,0]}\nu_i$, we see that

$$Q^F_{[0,\infty)}\big(\,\zeta, B_i(\zeta)\,\big) = 1. \tag{4.4}$$

In view of Proposition 4.5, we know that $\pi_0\nu_1$ and $\pi_0\nu_2$ are both equivalent to Lebesgue measure in $\mathbb{R}^2$. So that $\pi_0A_1 \cap \pi_0A_2 \neq \emptyset$. Together with (4.4), it follows that there exist ζ_1 and ζ_2 such that $\zeta_1(0) = \zeta_2(0)$ and $Q^F_{[0,\infty)}\big(\,\zeta_i, B_i(\zeta_i)\,\big) = 1$ for $i = 1, 2$. As Proposition 4.6 implies that $Q^F_{[0,\infty)}\big(\,\zeta_1,\,\cdot\,\big)$ is equivalent to $Q^F_{[0,\infty)}\big(\,\zeta_2,\,\cdot\,\big)$, we also know that

$$Q^F_{[0,\infty)}\big(\,\zeta_1, B_2(\zeta_2)\big) = 1 = Q^F_{[0,\infty)}\big(\,\zeta_2, B_1(\zeta_1)\big),$$

and hence

$$Q^F_{[0,\infty)}\big(\,\zeta_i, B_1(\zeta_1) \cap B_2(\zeta_2)\,\big) = 1, \quad \text{for } i = 1, 2.$$

In particular, this implies that $B_1(\zeta_1) \cap B_2(\zeta_2) \neq \emptyset$. By the definition of $B_i(\zeta_i)$, there exist $z_i \in A_i$, $i = 1, 2$ such that $\pi_{(-\infty,0]} z_i = \zeta_i$ and

$$\pi_{[0,\infty)} z_1 = \pi_{[0,\infty)} z_2 \in B_1(\zeta_1) \cap B_2(\zeta_2),$$

whence for all $t \geq 0$,

$$\pi_{[0,\infty)} \theta_t z_1 = \pi_{[0,\infty)} \theta_t z_2.$$

As a consequence, we have from (4.3) that

$$\bar{\phi}_1 = \lim_{T\to\infty} \frac{1}{T} \int_0^T \phi(\pi_{[0,\infty)} \theta_t z_1) dt = \lim_{T\to\infty} \frac{1}{T} \int_0^T \phi(\pi_{[0,\infty)} \theta_t z_2) dt = \bar{\phi}_2.$$

As ϕ was from a class of functions sufficiently rich to determine the laws of $\xi^{(i)}$, $i = 1, 2$, we conclude the laws are the same since we have proven that $\bar{\phi}^{(1)} = \bar{\phi}^{(2)}$. □

5. Proof of Proposition 4.5 and Proposition 4.6

In order to setup the proof of Proposition 4.6, observe that we may express equation (1.1) in a convenient form using integration-by-parts on the convolution term. Indeed, by Assumption 2.11, there exist constants $C, t_* > 0$ and $\alpha > 1/2$ such that $K(t) \leq Ct^{-\alpha}$ as $t \geq t_*$. Since K is continuously differentiable, L'Hospital's rule implies that for any $\epsilon > 0$, $K'(t)t^{\alpha+1-\epsilon} \to 0$ as $t \to \infty$. Now, given that $\xi_0 = (x_0, v_0) \in \mathcal{C}_\varrho(-\infty, 0]$ where $\varrho < \alpha - 1/2$, using integration-by-parts we may thus rewrite (1.1) as

$$\begin{aligned} dx(t) &= v(t)dt, \\ m\, dv(t) &= -\gamma v(t)\, dt - U'(x(t))\, dt - K(0)x(t)\, dt + \int_{-\infty}^t K'(t-r)x(r)\, dr\, dt \\ &\quad + \sqrt{2\gamma}\, dW(t) + F(t)\, dt. \end{aligned} \tag{5.1}$$

Proof of Proposition 4.6. Suppose $\xi_0, \tilde{\xi}_0 \in \mathcal{C}_\varrho(-\infty, 0]$. Let $\overline{\xi}_0 = \xi_0 - \tilde{\xi}_0$ and observe that (5.1) with $m = \gamma = 1$ and initial condition $\widetilde{\xi}_0$ can be expressed as

$$\begin{aligned} d\widetilde{x}(t) &= \widetilde{v}(t)\, dt, \\ d\widetilde{v}(t) &= -\widetilde{v}(t)\, dt - U'(\widetilde{x}(t))\, dt - K(0)\widetilde{x}(t)dt + \int_{-\infty}^0 K'(t-r)x_0(r)drdt \\ &\quad + \int_0^t K'(t-r)\widetilde{x}(r)drdt + F(t)\, dt \\ &\quad + \sqrt{2}dW(t) - \int_{-\infty}^0 K'(t-r)\overline{x}_0(r)drdt. \end{aligned} \tag{5.2}$$

If the following Novikov condition is satisfied

$$\mathbb{E} \exp\Big\{\frac{1}{2}\int_0^\infty \Big(\int_{-\infty}^0 K'(t-r)\overline{x}_0(r)dr\Big)^2 dt\Big\} < \infty,$$

then Girsanov's theorem would imply the desired measure equivalence on future paths. Since $\overline{x}_0$ is deterministic, it suffices to show that the above integral is finite. To this end, we note that since $\xi_0, \widetilde{\xi}_0 \in C_\varrho((-\infty, 0], \mathbb{R}^2)$, $\overline{x}_0(\cdot)$ satisfies the growth bound

$$\|\overline{x}_0\|_\varrho := \sup_{r\leq 0} \frac{|\overline{x}_0(r)|}{1 + |r|^\varrho} < \infty.$$

Using this fact, we estimate as follows:

$$\begin{aligned}\int_0^\infty\Big(\int_{-\infty}^0 K'(t-r)\overline{x}_0(r)dr\Big)^2dt &= \int_0^\infty\Big(\int_{-\infty}^0 K'(t-r)(1+|r|^\varrho)\frac{\overline{x}_0(r)}{1+|r|^\varrho}\,dr\Big)^2dt\\ &\le \|\overline{x}_0\|_\varrho^2\int_0^\infty\Big(\int_{-\infty}^0 K'(t-r)(1+|r|^\varrho)\,dr\Big)^2dt\\ &= \|\overline{x}_0\|_\varrho^2\int_0^\infty\Big(\int_0^\infty K'(t+r)(1+r^\varrho)\,dr\Big)^2dt.\end{aligned}$$

For $\epsilon>0$ to be chosen later, recalling by L'Hospital's rule applied to $K(t)/t^{\epsilon-\alpha}$, by Assumption 2.11, we saw that $K'(t)/t^{\epsilon-\alpha-1}\to 0$ as $t\to\infty$. Hence, there exist $C>0$ and $t_0>1$ such that $|K'(t)|\le Ct^{\epsilon-\alpha-1}$ for all $t\ge t_0$. It then follows that

$$\begin{aligned}&\int_0^\infty\Big(\int_0^\infty K'(t+r)(1+r^\varrho)\,dr\Big)^2dt\\ &\quad=\int_0^{t_0}\Big(\int_0^{t_0}K'(t+r)(1+r^\varrho)\,dr+\int_{t_0}^\infty K'(t+r)(1+r^\varrho)\,dr\Big)^2dt\\ &\qquad+\int_{t_0}^\infty\Big(\int_0^\infty K'(t+r)(1+r^\varrho)\,dr\Big)^2dt\\ &\quad\le C_1+C_2\int_0^{t_0}\Big(\int_{t_0}^\infty\frac{1+r^\varrho}{r^{1+\alpha-\epsilon}}\,dr\Big)^2dt+C_3\int_{t_0}^\infty\Big(\int_0^\infty\frac{1+r^\varrho}{(t+r)^{1+\alpha-\epsilon}}\,dr\Big)^2dt.\end{aligned}$$

Choosing $0<\epsilon<\alpha-\varrho-1/2$, notice that the first integral on the right hand side of the last line above is finite since $\alpha-\epsilon>\varrho$. For the final integral above, recalling that $t_0>1$ and making the substitution $u=r/t$ produces

$$\begin{aligned}\int_{t_0}^\infty\Big(\int_0^\infty\frac{1+r^\varrho}{(t+r)^{1+\alpha-\epsilon}}\,dr\Big)^2dt &= \int_{t_0}^\infty\frac{1}{t^{2(\alpha-\epsilon)}}\Big(\int_0^\infty\frac{1+t^\varrho u^\varrho}{(1+u)^{1+\alpha}}\,du\Big)^2dt\\ &\le\int_{t_0}^\infty\frac{1}{t^{2(\alpha-\epsilon-\varrho)}}dt\Big(\int_0^\infty\frac{1+u^\varrho}{(1+u)^{1+\alpha}}\,du\Big)^2<\infty,\end{aligned}$$

since $\alpha-\epsilon-\varrho>1/2$. This finishes the proof.

□

We now turn to the proof of Proposition 4.5. In order to show equivalence in measures, we aim to compare (1.1) with a standard, memoryless Langevin equation. However, because of the memory terms and the nonlinearity U', we do not do so directly. Instead, we will consider a truncated version of (1.1), which will be useful in verifying Novikov's condition. More precisely, let $\theta_n\in C^\infty(\mathbb{R},[0,1])$ satisfy $\theta_n(x)=1$ for all $|x|\le n$ and $\theta_n(x)=0$ for $|x|\ge n+1$, and consider the following system with initial path ξ_0

$$\begin{aligned}d\,x(t)&=v(t)dt,\\ d\,v(t)&=-v(t)dt-U'(x(t))dt+\sqrt{2}dW(t)\\ &\qquad+\theta_n\big(|x(t)|+|v(t)|+|F(t)|\big)\Big(-\int_{-\infty}^t K(t-r)v(r)dr+F(t)\Big)dt.\end{aligned}\tag{5.3}$$

In the following auxiliary result, we show that the solution of (5.3) converges to that of (1.1) as n tends to infinity.

Lemma 5.1. *Given an initial condition $\xi_0\in\mathcal{C}(-\infty,0]$ as in (2.2), let ξ^n and ξ respectively be the solutions of (5.3) and (1.1) (with $m=\gamma=1$ in (1.1)) with the same initial history ξ_0. Then, for* *all* $t \geq 0$,

$$\lim_{n\to\infty} \mathbb{E} \sup_{0\le r\le t} \Big\{ |x^n(r) - x(r)| + |v^n(r) - v(r)| \Big\} = 0. \tag{5.4}$$

The proof of Lemma 5.1 follows a standard comparison argument that will be deferred to the end of this section. Assuming this result, we now establish Proposition 4.5.

Proof of Proposition 4.5. Let $Q^F_t(\xi_0,\cdot)$ be the law at time t of $\varphi^{F,W}_{0,t}\xi_0$ on $\mathbb{R}^2$. By stationarity,

$$\pi_t \mathrm{Law}(\xi,\cdot|F) = \int Q^F_t(\xi_0,\cdot)\nu_{(-\infty,0]}(d\xi_0\times dF).$$

It therefore suffices to show that $Q^F_t(\xi_0,\cdot)$ is equivalent to Lebesgue measure.

Recalling that $\xi^n = (x^n, v^n)$ denotes the solution of (5.3), let $Q^{n,F}_{[0,t]}(\xi_0,\cdot)$ be the law induced by ξ^n on $C([0,t],\mathbb{R}^2)$ and let $Q^{n,F}_t(\xi_0,\cdot)$ be the marginal of $Q^{n,F}_{[0,t]}(\xi_0,\cdot)$ at time t. We note that

$$\int_{-\infty}^t K(t-r)v^n(r)dr = \int_{-\infty}^0 K(t-r)v_0(r)dr + \int_0^t K(t-r)v^n(r)dr.$$

By Assumption 2.6 and the definition of θ_n, the following estimate holds almost surely

$$\begin{aligned}
\theta_n\Big(|x^n(t)| + |v^n(t)| + |F(t)|\Big)\Big| -\int_{-\infty}^t K(t-r)v(r)dr + F(t)\Big| \\
\le \int_{-\infty}^0 K(t-r)|v_0(r)|dr + n\int_0^t K(t-r)dr + n \\
\le \widetilde{K}(t)\int_{-\infty}^0 K(-r)|v_0(r)|dr + n\int_0^t K(t-r)dr + n,
\end{aligned}$$

implying the following Novikov-type condition is satisfied

$$\mathbb{E}\exp\Big\{ \tfrac{1}{2}\int_0^t \theta_n\big(|x^n(r)|+|v^n(r)|+|F(r)|\big)^2 \Big(-\int_{-\infty}^r K(r-\ell)v(\ell)d\ell + F(r)\Big)^2 dr\Big\} < \infty.$$

As a consequence, $Q^{n,F}_{[0,t]}(\xi_0,\cdot)$ is equivalent to the law $\widetilde{Q}_{[0,t]}(\xi_0(0),\cdot)$ induced by the solution of the following Langevin equation

$$\begin{aligned}
d\,x(t) &= v(t)dt, & x(0) &= x_0(0),\\
d\,v(t) &= -v(t)dt - U'(x(t))dt + \sqrt{2}dW(t), & v(0) &= v_0(0).
\end{aligned}$$

The above system is well-understood. By verifying Hormander's condition, it is clear that $\widetilde{Q}_t(\xi_0(0),\cdot)$ as the marginal law of $\widetilde{Q}_{[0,t]}(\xi_0(0),\cdot)$ at time t is equivalent to Lebesgue measure [29]. It follows immediately that $Q^{n,F}_t(\xi_0,\cdot)$ must be too. By taking n to infinity, in light of Lemma 5.1, $Q^{n,F}_t(\xi_0,\cdot)$ converges to $Q^F_t(\xi_0,\cdot)$, which preserves measure equivalence. The proof is thus complete. □

We finally give the proof of Lemma 5.1 whose proof is somewhat standard.

Proof of Lemma 5.1. We first note that by adapting the energy estimate as in the proof of Proposition 2.7 to (5.3), we have the following uniform bound in n

$$\begin{aligned}
&\mathbb{E}\sup_{r\in[0,t]} H(x_n(r), v_n(r)) \\
&\le \Big(H(x_0(0), v_0(0)) + \Big(\int_{-T}^0 K(-w)|v_0(w)|\,dw\Big)^2 + \mathbb{E}\sup_{r\in[0,t]}|F(r)|^2 + 1\Big)e^{c(t)}. \qquad (5.5)
\end{aligned}$$

Now consider the stopping time τ_n given by

$$\tau_n = \inf\{t \geq 0 : |x(t)| + |v(t)| + |F(t)| \geq n\}.$$

It is clear that $\xi(r) = \xi^n(r)$ for all $0 \leq r \leq \tau_n$. Using Holder's inequality and recalling $\delta \in (0,1)$ as in Assumption 2.4:

$$\begin{aligned}
&\mathbb{E} \sup_{0\leq r\leq t} |x^n(r) - x(r)| + |v^n(r) - v(r)| \\
&\qquad = \mathbb{E}\Big(\sup_{0\leq r\leq t} |x^n(r) - x(r)| + |v^n(r) - v(r)| 1\{\tau_n \leq t\}\Big) \\
&\qquad \leq c\Big(\mathbb{E} \sup_{0\leq r\leq t} |x^n(r)|^{1+\delta} + |v^n(r)|^{1+\delta} + \mathbb{E} \sup_{0\leq r\leq t} |x(r)|^{1+\delta} + |v(r)|^{1+\delta}\Big)^{1/(1+\delta)} \Big(\mathbb{P}(\tau_n \leq t)\Big)^{\delta/(1+\delta)}.
\end{aligned}$$

We invoke the energy estimates (2.3) and (5.5) and recall that the nonlinear potential U dominates $|x|^{1+\delta}$, cf. Assumption 2.4, to see that

$$\begin{aligned}
&\mathbb{E} \sup_{0\leq r\leq t} |x^n(r)|^{1+\delta} + |v^n(r)|^{1+\delta} + \mathbb{E} \sup_{0\leq r\leq t} |x(r)|^{1+\delta} + |v(r)|^{1+\delta} \\
&\qquad \leq c\Big(\mathbb{E} \sup_{0\leq r\leq t} |x^n(r)|^{1+\delta} + |v^n(r)|^2 + \mathbb{E} \sup_{0\leq r\leq t} |x(r)|^{1+\delta} + |v(r)|^2 + 1\Big) \leq e^{c(t,\xi_0,F)},
\end{aligned}$$

where $c(t,\xi_0,F) > 0$ is a constant independent of n. Also, by Chebyshev's inequality and Lemma B.2,

$$\begin{aligned}
\mathbb{P}(\tau_n \leq t) &= \mathbb{P}(\sup_{0\leq r\leq t} |x(r)| + |v(r)| + F(r) \geq n) \\
&\leq \frac{1}{n}\Big(\mathbb{E} \sup_{0\leq r\leq t} |x(r)| + |v(r)| + \sup_{0\leq r\leq t} |F(r)|\Big) \leq \frac{e^{c(t,\xi_0,F)}}{n}.
\end{aligned}$$

Altogether, we arrive at the bound

$$\mathbb{E} \sup_{0\leq r\leq t} |x^n(r) - x(r)| + |v^n(r) - v(r)| \leq \frac{e^{c(t,\xi_0,F)}}{\sqrt{n}},$$

which converges to zero as n tends to infinity. This finishes the proof. □

6. Existence of an invariant measure

In this section, we assume that the memory kernel K is of the form

$$K(t) = \sum_{\ell\geq 1} \int_0^\infty J_\ell(s+t) J_\ell(s)\, ds, \tag{6.1}$$

where the functions J_ℓ, $\ell \geq 1$, are as in Remark 2.12. In this case, we will see here that we can construct an explicit stationary measure for the Markov flow on $\mathcal{C}_\varrho((-\infty,0])$ by pulling back a known invariant measure for an augmented version of (1.1).

Introducing the auxiliary variable $z_k(t)$ given by

$$z_\ell(t) = \sqrt{c_\ell} \int_{-\infty}^t e^{-\lambda_\ell (t-r)} v(r) dr dt + \sqrt{2\lambda_\ell} \int_{-\infty}^t e^{-\lambda_\ell(t-r)} dB^{(\ell)}(r) dt, \tag{6.2}$$

we find that equation (1.1) can be expressed as

$$\begin{aligned}
d\, x(t) &= v(t)\, dt, \\
d\, v(t) &= -v(t)\, dt - U'(x(t))\, dt - \textstyle\sum_{\ell\geq 1} \sqrt{c_\ell} z_\ell(t)\, dt + \sqrt{2}\, dW(t), \\
d\, z_\ell(t) &= -\lambda_\ell z_\ell(t)\, dt + \sqrt{c_\ell} v(t)\, dt + \sqrt{2\lambda_\ell}\, dB^{(\ell)}(t), \quad \ell \geq 1.
\end{aligned} \tag{6.3}$$

In this setting, the relationship between the system above and the original equation (1.1) must account for a specific initial condition in the past. For now, however, we view this system as a Markovian dynamics started from a given initial condition on the phase space $\mathcal{H}_{-s}$ where

$$\mathcal{H}_{-s} := \big\{X = (x, v, z_1, \dots) : \|X\|^2_{-s} := x^2 + v^2 + \textstyle\sum_{\ell\geq 1}\ell^{-2s} z_\ell^2 < \infty\big\}. \tag{6.4}$$

In the above, the real parameter s is such that

$$1 < 2s < \alpha\beta, \tag{6.5}$$

and $\alpha, \beta > 0$ are as in Remark 2.12. Under these hypotheses, the system (6.3) is well-posed on $\mathcal{H}_{-s}$, and the probability measure on $\mathcal{H}_{-s}$ given by

$$\mu \propto \pi \times \prod_{\ell\geq 1} \nu_\ell, \tag{6.6}$$

where π is the Boltzmann-Gibbs measure in (x, v) as in (1.3) and $\{\nu_k\}_{k\geq 1}$ are independent copies of the standard normal distribution $N(0, 1)$ on $\mathbb{R}$, is an invariant probability measure for the Markov process (6.3) [11, Theorem 7].

6.1. **The induced measure on path space.** Consider an arbitrary collection of real numbers

$$t_1 \leq t_2 \leq \cdots \leq t_n,$$

and a collection of Borel sets $A_1, \dots, A_n \subset \mathcal{H}_{-s}$. If $X_{t_1}(\,\cdot\,)$ denotes the solution of (6.3) distributed as μ at time t_1, we define $\widehat{\mu}_{t_1,\dots,t_n}$ on the cylinder set $A_1 \times \cdots \times A_n$ by

$$\widehat{\mu}_{t_1,\dots,t_n}(A_1 \times \cdots \times A_n) = \mathbb{P}\{X_{t_1}(t_1) \in A_1, \dots, X_{t_1}(t_n) \in A_n\}. \tag{6.7}$$

Since μ is invariant for the Markov process (6.3), it can be shown by Kolmogorov's extension theorem (by taking a continuous version of the process X solving (6.3)) that the family

$$\{\widehat{\mu}_{t_1,\dots,t_n} : n \in \mathbb{N},\, t_1 \leq t_2 \leq \ldots \leq t_n\},$$

is consistent, and hence induces a stationary measure, denoted by $\widehat{\mu}$, on Borel subsets of $C(\mathbb{R}, \mathcal{H}_{-s})$ whose finite-dimensional distributions are as in (6.7). Let $\xi_* = (x_*, v_*)$ denote the projection of the corresponding stationary process on $C(\mathbb{R}, \mathcal{H}_{-s})$ onto $C((-\infty, 0], \mathbb{R}^2)$. By definition, it follows that $\psi^W_t(\xi_*, B)$ is stationary on $\mathcal{S}(-\infty, 0]$ given by (3.3). Let μ_* denote its corresponding distribution, which in particular is invariant for the Markov semigroup P_t as in (3.4).

We will next show that μ_* concentrates on a path space with moderate growth, thereby finishing the proof of Theorem 4.1.

Lemma 6.1. *Let μ_* be the probability measure in $\mathcal{S}(-\infty, 0]$ constructed above. Then, for every $\varrho > 0$,*

$$\mu_*(\mathcal{C}_\varrho(-\infty, 0]) = 1,$$

where $\mathcal{C}_\varrho(-\infty, 0]$ is as in (4.1).

Proof. By Borel-Cantelli, it suffices to prove that

$$\sum_{n\geq 1} \widehat{\mu}\Big\{X(\cdot) = (\xi, z_1, z_2, \dots) \in C(\mathbb{R}, \mathcal{H}_{-s}) \,:\, \sup_{-n\leq r\leq -n+1} |x(r)| > (n+1)^\varrho\Big\} < \infty.$$

By invariance

$$\widehat{\mu}\Big\{X(\cdot) : \sup_{-n\leq r\leq -n+1} |x(r)| > (n+1)^\varrho\Big\} = \widehat{\mu}\Big\{\sup_{0\leq t\leq 1} |x_0(t)| > (n+1)^\varrho\Big\},$$

where $X_0(\cdot) = (\xi_0(\cdot), z_1(\cdot), \dots)$ denotes the solution of (6.3) with initial distribution μ at time 0. To estimate the righthand side above, we apply Itô's formula to the Hamiltonian $H(\xi) = H(x, v) = \frac{1}{2}v^2 + U(x)$, and obtain for $t \geq 0$

$$dH(\xi(t)) = -v(t)^2 dt + 1dt + v(t)dW(t) - \sum_{\ell \geq 1} \sqrt{c_\ell} z_\ell(t) v(t) dt. \tag{6.8}$$

The cross terms involving $z_\ell(t)$ and $v(t)$ can be bounded from above by

$$\sqrt{c_\ell}|z_\ell(t)v(t)| \leq C\ell^{-2s} z_\ell(t)^2 + \frac{c_\ell \ell^{2s}}{C} v(t)^2,$$

where $C > 0$ is large enough such that $C\sum_{\ell \geq 1} c_\ell \ell^{2s} < 1$. Integrating (6.8) on $[0, t]$, $t \leq 1$ using the estimates above then produces

$$H(\xi(t)) \leq H(\xi(0)) + 1 + C\int_0^1 \sum_{\ell \geq 1} \ell^{-2s} z_\ell(s)^2 \, ds + \sup_{0 \leq t \leq 1} \int_0^t v(r)dW(r).$$

Fixing $\varepsilon \in (1/2, s)$ and recalling Assumption 2.4, namely, $U(x)$ dominates $|x|^{1+\delta}$, $\delta \in (0, 1)$, we have the following chain of implications

$$\begin{aligned}
&\Big\{\sup_{0 \leq t \leq 1} |x_0(t)| \geq (n+1)^\varrho\Big\} \subset \Big\{\sup_{0 \leq t \leq 1} U(x_0(t)) \geq c(n+1)^{(1+\delta)\varrho}\Big\} \\
&\subset \Big\{H(\xi(0)) + 1 + C\textstyle\int_0^1 \sum_{\ell \geq 1} \ell^{-2s} z_\ell(s)^2 \, ds + \sup_{0 \leq t \leq 1} \int_0^t v(r)dW(r) \geq c(n+1)^{(1+\delta)\varrho}\Big\}. \\
&\subset \Big\{U(x(0)) \geq c(n+1)^{(1+\delta)\varrho}\Big\} \cup \{\tfrac{1}{2}v(0)^2 + 1 + \sup_{0 \leq t \leq 1} \textstyle\int_0^t v(r)dW(r) \geq c(n+1)^{(1+\delta)\varrho}\Big\} \\
&\quad \bigcup_{\ell \geq 1} \Big\{C\textstyle\int_0^1 \ell^{-2s} z_\ell(s)^2 \, ds \geq c(n+1)^{(1+\delta)\varrho} \ell^{-2\varepsilon}\Big\} = I_x \cup I_v \cup_{\ell \geq 1} I_\ell.
\end{aligned}$$

We are left to estimate each of the above events. For $p > 0$, using Chebychev's inequality, we estimate I_x as follows:

$$\begin{aligned}
\widehat{\mu}(I_x) = \mu\{I_x\} &\leq \frac{c}{(n+1)^{(1+\delta)\varrho p}} \int_{\mathcal{H}_{-s}} U(x)^p \mu(dX) \\
&= \frac{c}{(n+1)^{(1+\delta)\varrho p}} \int_{\mathbb{R}} U(x)^p e^{-U(x)} dx \leq \frac{c}{(n+1)^{(1+\delta)\varrho p}}.
\end{aligned}$$

For I_v, we first employ Burkholder's inequality to see that

$$\mathbb{E} \sup_{0 \leq t \leq 1} \left(\int_0^t v(r) dB_0(r)\right)^{2p} \leq c\int_0^1 \mathbb{E}|v(r)|^{2p} dr.$$

Using the fact that μ is invariant for $X_0(t)$, we estimate I_v

$$\begin{aligned}
\widehat{\mu}(I_v) &\leq \frac{c}{(n+1)^{(1+\delta)\varrho p}} \left[\int_{\mathcal{H}_{-s}} v^{2p} \mu(dX) + \int_0^1 \mathbb{E} \int_{\mathcal{H}_{-s}} v(r)^{2p} \mu(dX) dr\right] \\
&= \frac{c}{(n+1)^{(1+\delta)\varrho p}} \int_{\mathcal{H}_{-s}} 2v^{2p} \mu(dX) \\
&= \frac{c}{(n+1)^{(1+\delta)\varrho p}} \int_{\mathbb{R}} 2v^{2p} e^{-v^2/2} dv \leq \frac{c}{(n+1)^{(1+\delta)\varrho p}}.
\end{aligned}$$

Likewise, for I_ℓ we find that

$$\begin{aligned}\widehat{\mu}(I_\ell) &\le \frac{c\ell^{2(\varepsilon-s)p}}{(n+1)^{(1+\delta)\varrho p}}\int_0^1 \mathbb{E}\int_{\mathcal{H}_{-s}} |z_\ell^2(r)|^p \mu(dX)dr \\ &= \frac{c\ell^{2(\varepsilon-s)p}}{(n+1)^{(1+\delta)\varrho p}}\int_{\mathbb{R}} |z|^{2p} e^{-z^2/2}dz.\end{aligned}$$

We now collect everything and note that $\varepsilon \in (1/2, s)$ to arrive at

$$\widehat{\mu}\Big\{\sup_{0\le t\le 1} |x(t)| \ge (n+1)^{\varrho}\Big\} \le \frac{c}{(n+1)^{(1+\delta)\varrho p}}\big[1+\sum_{\ell\ge 1}\ell^{2(\varepsilon-s)p}\big] \le \frac{c}{(n+1)^{(1+\delta)\varrho p}},$$

which holds for p sufficiently large, e.g., $2(s-\varepsilon)p > 1$. Furthermore, we emphasize that the above constant c is independent of n. It follows that

$$\sum_{n\ge 1}\widehat{\mu}\Big\{X(\cdot) = (\xi, z_1, z_2, \dots) \in C(\mathbb{R}, \mathcal{H}_{-s}) \,:\, \sup_{-n\le r\le -n+1} |x(r)| > (n+1)^{\varrho}\Big\} \le \sum_{n\ge 1}\frac{c}{(n+1)^{(1+\delta)\varrho p}},$$

which is summable as long as p is chosen such that $(1+\delta)\varrho p > 1$. The proof is thus complete. □

Acknowlegement

The authors graciously acknowledge support from the Department of Mathematics at Duke University and the Department of Mathematics at Iowa State University. DPH was supported in part by NSF Grants DMS-1612898 and DMS-1855504. JCM thanks the NSF for its partial suport through the grant DMS-1613337. The authors also would like to thank Gustavo Didier for helpful discussions in the development of this work.

Appendix A. Well-posedness

In this section, we show that equation (1.1) is well-posed as stated in Proposition 2.7. We first construct strong, i.e., pathwise, solutions. Then, the existence and uniqueness of weak solutions simply follow by using classical arguments [27, Chapter 5].

First, fixing $T > 0$ we consider a slightly different approximating equation

$$\begin{aligned}dx(t) &= v(t)\,dt, \\ dv(t) &= -v(t)\,dt - U'(x(t))\,dt - \int_{-T}^t K(t-s)v(s)\,ds\,dt + \sqrt{2}\,dW(t) + F(t)\,dt,\end{aligned} \tag{A.1}$$

where we have truncated the memory term in (1.1) at time $-T$. Following the standard iteration procedure for standard SDEs with globally Lipschitz coefficients [3, 27], we can obtain the well-posedness of relation (A.1) assuming that U' is globally Lipschitz:

Lemma A.1. *Fix $T > 0$, $K \in C(\mathbb{R})$ and suppose that U' is globally Lipschitz. Then for all $\xi_0 = (x_0, v_0) \in \mathcal{C}(-\infty, 0]$, there exists a unique continuous adapted solution $\xi^T(t) = (x^T(t), v^T(t))$ of equation A.1 for all times $t \ge 0$ with $\xi^T(0) = (x_0(0), v_0(0))$.*

In order to remove the globally Lipschitz hypothesis in Lemma A.1, we use an energy estimate to show absence of explosion under the assumption that $U' \in C^1(\mathbb{R})$ with $U' \to \infty$ as $|x| \to \infty$.

Lemma A.2. *Fix $T > 0$, $K \in C(\mathbb{R})$ and suppose Assumption 2.6 holds. Furthermore, suppose that U' in equation (A.1) satisfies $U' \to \infty$ as $|x| \to \infty$. Then for all $\xi_0 = (x_0, v_0) \in \mathcal{C}(-\infty, 0]$, there exists a unique continuous solution $\xi^T(t) = (x^T(t), v^T(t))$ of equation (A.1) for all times $t \ge 0$ with $\xi^T(0) = (x_0(0), v_0(0))$.*

Proof. Recalling θ_n as in (5.3), let $H_n(x,v) = \frac{1}{2}v^2 + U(x)\theta_n(x)$. Define $U_n : \mathbb{R} \to \mathbb{R}$ by $U_n(x) = U(x)\theta_n(x)$ and note that the system (A.1) with U' replaced by U_n' has unique solutions $(x_n(t), v_n(t))$ as in Lemma A.1 with $(x_n(0), v_n(0)) = \xi(0) \in \mathbb{R}^2$. Furthermore, these solutions agree with the solutions of equation (A.1) for all times $t < \sigma_n := \inf\{t \geq 0 : H(x_n(t), v_n(t)) \geq n\}$ where H is the Hamiltonian. Now, fix $t > 0$ and note that Itô's formula implies

$$\begin{aligned}\mathbb{E} \sup_{r\in[0,t]} H_n(x_n(r), v_n(r)) &\leq H(\xi(0)) + \mathbb{E} \sup_{r\in[0,t]} \int_0^r \left\{ |v_n(u)| \int_{-T}^u K(u-w)|v_n(w)|\, dw \right\} du \\ &\quad + \sqrt{2}\mathbb{E} \sup_{r\in[0,t]} \left| \int_0^r v_n(u) dW(u) \right| + \mathbb{E} \sup_{r\in[0,t]} \left| \int_0^r v_n(u) F(u)\, du \right| \\ &=: H(\xi(0)) + (I)_t + (II)_t + (III)_t.\end{aligned}$$

For the term $(I)_t$, we note that Assumption 2.6 gives

$$\begin{aligned}&\int_0^r \left\{ |v_n(u)| \int_{-T}^u K(u-w)|v_n(w)|\, dw \right\} du \\ &= \int_0^r \left\{ |v_n(u)| \int_{-T}^0 \frac{K(u-w)}{K(-w)} K(-w)|v_n(w)|\, dw \right\} du + \int_0^r \left\{ |v_n(u)| \int_0^u K(u-w)|v_n(w)|\, dw \right\} du \\ &\leq \int_0^r |v_n(u)| \widetilde{K}(u) \int_{-T}^0 K(-w)|v_0(w)|\, dw\, du + \int_0^r \sup_{w\in[0,u]} |v_n(w)|^2 \int_0^u K(u-w)\, dw\, du.\end{aligned}$$

Hence we can estimate $(I)_t$ as

$$(I)_t \leq \int_0^t c_1(r) \mathbb{E} \sup_{s\in[0,r]} |v_n(s)|^2\, dr + c_2(t) \bigg(\int_{-T}^0 K(-w)|v_0(w)|\, dw \bigg)^2,$$

for some continuous functions c_i on $[0,t]$. For the term $(II)_t$, note that Doob's Maximal Inequality implies

$$(II)_t = \sqrt{2}\mathbb{E} \sup_{r\in[0,t]} \left| \int_0^r v_n(u) dW(u) \right| \leq c\bigg(1 + \int_0^t \mathbb{E} \sup_{u\in[0,r]} |v_n(u)|^2 dr \bigg).$$

Concerning $(III)_t$, we use Young's inequality for products to obtain

$$(III)_t \leq \tfrac{1}{2} \int_0^t \mathbb{E} \sup_{u\in[0,r]} |v_n(u)|^2 dr + \tfrac{1}{2} t \mathbb{E} \sup_{r\in[0,t]} |F(r)|^2.$$

We collect the estimates above to arrive at the bound

$$\begin{aligned}\mathbb{E} \sup_{r\in[0,t]} H_n(x_n(r), v_n(r)) &\leq H_n(\xi(0)) + c_1(t) \int_0^t \mathbb{E} \sup_{u\in[0,r]} H_n(x_n(u), v_n(u)) dr \\ &\quad + c_2(t) \bigg(\int_{-T}^0 K(-w)|v_0(w)|\, dw \bigg)^2 + \tfrac{1}{2} t \mathbb{E} \sup_{r\in[0,t]} |F(r)|^2 + c_3(t),\end{aligned}$$

whence using Grönwall's inequality and the Monotone Convergence Theorem

$$\begin{aligned}&\mathbb{E} \sup_{r\in[0,t]} H(x_n(r), v_n(r)) \\ &\leq \bigg(H(x_0(0), v_0(0)) + \bigg(\int_{-T}^0 K(-w)|v_0(w)|\, dw \bigg)^2 + \mathbb{E} \sup_{r\in[0,t]} |F(r)|^2 + 1 \bigg) e^{c(t)},\end{aligned} \tag{A.2}$$

Turning back to σ_n, we note that

$$\mathbb{E}\sup_{r\in[0,t]} H_n(x_n(r),v_n(r)) \geq \mathbb{E}\Big[\sup_{r\in[0,t]} H_n(x_n(r),v_n(r))\cdot 1\{\sigma_{n-1}<t\}\Big] \geq (n-1)\mathbb{P}(\sigma_{n-1}<t), \tag{A.3}$$

which together with (A.2) yields $\mathbb{P}(\sigma_n<t)\leq \frac{1}{n}c(t)$. By taking n to infinity, we immediately obtain $\mathbb{P}(\sigma_\infty<t)=0$ for any $t\geq 0$. Hence $\mathbb{P}(\sigma_\infty=\infty)=1$, finishing the proof. □

Our next goal is to allow the memory to depend on the infinite past by carefully passing T to infinity in (A.1).

Lemma A.3. *Let $T>0$, $\xi_0=(x_0,v_0)\in\mathcal{C}(-\infty,0]$ and suppose K satisfies Assumption 2.6. Suppose $U\in C^1(\mathbb{R})$ is such that $U(x)\to\infty$ as $|x|\to\infty$, and let $\xi^T(t)=(x^T(t),v^T(t))$ denote the solution of equation (A.1) with $\xi^T(0)=\xi_0(0)$. Then for any $t>0$, the solution ξ^T converges as $T\to\infty$ to ξ in $C([0,t],\mathbb{R}^2)$. Furthermore, ξ is the unique pathwise solution of (1.1) with $\xi(0)=\xi_0(0)$.*

Proof. Let $t>0$. Uniqueness of solutions and the fact that the presumed limit solves (1.1) both follow almost immediately once we show that an appropriate approximating sequence is Cauchy in $C([0,t];\mathbb{R}^2)$. To be more precise, for $T_1\geq T_2>0$, let $\xi_n^{T_1}=(x_n^{T_1},v_n^{T_1})$ and $\xi_n^{T_1}=(x_n^{T_2},v_n^{T_2})$ respectively be the solutions of (A.1) with $U'(x)$ being replaced by $U_n'(x)$ where $U_n(x)=U(x)\theta_n(x)$ as in the proof of Lemma A.2. For simplicity, let $\overline{\xi}_n=\xi_n^{T_1}-\xi_n^{T_2}=(\overline{x}_n,\overline{v}_n)$ and observe that

$$\begin{aligned}|\overline{x}_n(t)|+|\overline{v}_n(t)| &\leq 2\int_0^t|\overline{v}_n(r)|dr+\int_0^t\big|U_n(x_n^{T_1}(r))-U_n(x_n^{T_2}(r)\big|dr\\ &\quad+\int_0^t\int_{-T_1}^{-T_2}K(r-u)|v_0(u)|dudr+\int_0^t\int_0^r K(r-u)|\overline{v}_n(u)|dudr.\end{aligned}$$

Note that by Assumption 2.6,

$$\begin{aligned}\int_0^t\int_{-T_1}^{-T_2}K(r-u)|v_0(u)|dudr &= \int_0^t\int_{-T_1}^{-T_2}\frac{K(r-u)}{K(u)}K(u)|v_0(u)|dudr\\ &\leq \int_0^t\widetilde{K}(r)dr\cdot\int_{-T_1}^{-T_2}K(u)|v_0(u)|du.\end{aligned}$$

Using the fact that U_n is Lipschitz we then obtain

$$\sup_{0\leq r\leq t}|\overline{x}_n(r)|+|\overline{v}_n(r)|\leq c(t,n)\int_0^t\sup_{0\leq u\leq r}|\overline{x}_n(u)|+|\overline{v}_n(u)|dr+c(t)\int_{-T_1}^{-T_2}K(u)|v_0(u)|du.$$

Thus Grönwall's inequality gives

$$\sup_{0\leq r\leq t}|\overline{x}_n(r)|+|\overline{v}_n(r)|\leq e^{c(n,t)}\int_{-T_1}^{-T_2}K(u)|v_0(u)|du. \tag{A.4}$$

Next, let $\sigma_n^{T_1}$ and $\sigma_n^{T_2}$ respectively denote the stopping times associated with $\xi_n^{T_1}(t)$ and $\xi_n^{T_2}(t)$ as in the proof of Lemma A.2. Setting $\overline{\xi}(t) = \xi^{T_1}(t) - \xi^{T_2}(t)$ we find that

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)|\Big] &\le \mathbb{E}\Big[1\{\sigma_n^{T_1}\wedge\sigma_n^{T_2}\ge t\}\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)|\Big]\\ &\quad+\mathbb{E}\Big[1\{\sigma_n^{T_1}\le t\}\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)|\Big]\\ &\quad+\mathbb{E}\Big[1\{\sigma_n^{T_2}\le t\}\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)|\Big]\\ &= (I)_t+(II)_t+(III)_t.\end{aligned}$$

In view of (A.4), we have

$$(I)_t \le \mathbb{E}\sup_{0\le r\le t}|\overline{x}_n(r)|+|\overline{v}_n(r)| \le e^{c(n,t)}\int_{-T_1}^{-T_2}K(u)|v_0(u)|du.$$

Concerning $(II)_t$, we use Holder's inequality and Assumption 2.4 to infer the bound

$$\begin{aligned}(II)_t &\le \bigg(\mathbb{E}\Big[\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)|\Big]^{1+\delta}\bigg)^{\frac{1}{1+\delta}}\Big(\mathbb{P}\Big(\sigma_n^{T_1}\le t\Big)\Big)^{\frac{\delta}{1+\delta}}\\ &\le c\Big(1+\mathbb{E}\sup_{0\le r\le t}|\overline{x}(r)|^{1+\delta}+|\overline{v}(r)|^2\Big)^{\frac{1}{1+\delta}}\Big(\mathbb{P}\Big(\sigma_n^{T_1}\le t\Big)\Big)^{\frac{\delta}{1+\delta}}\\ &\le c\Big(1+\mathbb{E}\Big[\sup_{0\le r\le t}H\big(x^{T_1}(r),v^{T_1}(r)\big)+H\big(x^{T_2}(r),v^{T_2}(r)\big)\Big]\Big)^{\frac{1}{1+\delta}}\Big(\mathbb{P}\Big(\sigma_n^{T^1}\le t\Big)\Big)^{\frac{\delta}{1+\delta}}\\ &\le c\Big(1+\mathbb{E}\Big[\sup_{0\le r\le t}H\big(x^{T_1}(r),v^{T_1}(r)\big)+H\big(x^{T_2}(r),v^{T_2}(r)\big)\Big]\Big)^{\frac{1}{1+\delta}}\cdot\frac{c(t)}{n^{\delta/(1+\delta)}}\le\frac{c(t)}{n^{\delta/(1+\delta)}}.\end{aligned}$$

In the above estimate, we employed (A.3) together with (A.2). Likewise,

$$(III)_t \le \frac{c(t)}{n^{\delta/(1+\delta)}}.$$

Altogether, we arrive at the bound

$$\mathbb{E}\sup_{0\le r\le t}|\overline{x}(r)|+|\overline{v}(r)| \le e^{c(n,t)}\int_{-T_1}^{-T_2}K(u)|v_0(u)|du+\frac{1}{n^{\delta/(1+\delta)}}c(t).$$

Thanks to the assumption that $\xi_0\in\mathcal{C}(-\infty,0]$, it is now clear that $\{\xi^T\}$ is a Cauchy sequence in $C([0,t];\mathbb{R}^2)$ by first taking n sufficiently large and then sending T_1 and T_2 to infinity. As a consequence, there exists a solution ξ for (1.1) with the initial condition $\xi_0\in\mathcal{C}$.

Turning to the uniqueness of ξ, it suffices to show that if $\widetilde{\xi}$ solves (1.1) with the same initial path ξ_0, then ξ and $\tilde{\xi}$ must agree a.s. in $[0,t]$. To see this, consider the stopping times σ_n and $\widetilde{\sigma}_n$ associated with ξ and $\tilde{\xi}$ respectively. Similarly to the above existence part, denoting $\widehat{\xi}=\xi-\tilde{\xi}$, we observe that for $0\le t\le\sigma_n\wedge\widetilde{\sigma}_n$, ξ and $\tilde{\xi}$ both solve equation (1.1) with U' being replaced by $U_n'(x)$. So that, for $0\le t\le\sigma_n\wedge\widetilde{\sigma}_n$, $\widehat{\xi}$ satisfies $\widehat{\xi}(0)=0$ and

$$\begin{aligned}&\frac{d}{dt}\widehat{x}(t)=\widehat{v}(t),\\ &\frac{d}{dt}\widehat{x}(t)=-\widehat{v}(t)-\big[U_n'(x(t))-U_n'(\widetilde{x}(t))\big]-\int_0^t K(t-r)\widehat{v}(r)dr.\end{aligned}$$

Since the nonlinear term is Lipschitz, by Gronwall's inequality, we immediately obtain

$$\mathbb{E}\Big[1\{\sigma_n \wedge \widetilde{\sigma}_n \ge t\} \sup_{0\le r\le t} |\widehat{x}(r)| + |\widehat{v}(r)|\Big] = 0.$$

On the other hand, similar to the estimate of $(II)_t$ above, we also have the bound

$$\begin{aligned}&\mathbb{E}\Big[\Big(1\{\sigma_n \le t\} + 1\{\widetilde{\sigma}_n \le t\}\Big) \sup_{0\le r\le t} |\widehat{x}(r)| + |\widehat{v}(r)|\Big]\\ &\le c\bigg(1 + \mathbb{E}\Big[\sup_{0\le r\le t} H\big(x(r), v(r)\big) + H\big(\widetilde{x}(r), \widetilde{v}(r)\big)\Big]\bigg)^{1/(1+\delta)} \cdot \frac{1}{n^{\delta/(1+\delta)}} c(t) \le \frac{1}{n^{\delta/(1+\delta)}} c(t).\end{aligned}$$

By taking n large, we observe that $\mathbb{E}\sup_{0\le r\le t} |\widehat{x}(r)| + |\widehat{v}(r)|$ is arbitrarily small, forcing

$$\mathbb{E}\sup_{0\le r\le t} |\widehat{x}(r)| + |\widehat{v}(r)| = 0,$$

holds true. The proof is thus complete. □

Given the strong solutions constructed above, we are now ready to give the proof of Proposition 2.7. The argument is relatively short and can be found in previous works (see, for example, [27]).

Proof of Proposition 2.7. The existence of weak solution is clear since we already constructed strong solutions as in Lemma A.3. It remains to show weak uniqueness.

Suppose (ξ, F, W) and $(\tilde{\xi}, \tilde{F}, \tilde{W})$ are two weak solutions as in Definition 2.2 on the interval $[t_0, t]$ with the same initial condition ξ_0. By the uniqueness of strong solutions, we may consider ξ and $\tilde{\xi}$ as the unique path-wise solutions given (F, W) and $(\tilde{F}, \tilde{W})$, respectively. To see that ξ and $\tilde{\xi}$ have the same law, we recall the construction of ξ starting from system (A.1) with U' being Lipschitz. Then, it is clear that the processes ξ^T and $\tilde{\xi}^T$ as in Lemma A.1 must agree in distribution [27, Lemma 5.3.1]. In view of Lemma A.2, this property also holds for general U satisfying Assumption 2.4. Finally, since ξ^T and $\tilde{\xi}^T$ respectively converge to ξ and $\tilde{\xi}$ on $C([t_0, t]; \mathbb{R}^2)$ as $T \to \infty$, cf. proof of Lemma A.3, we immediately establish the equality in law for ξ and $\tilde{\xi}$, thereby concluding the uniqueness of weak solutions. □

Appendix B. Bound on the expected maximum of $F(t)^2$

In this section, we will show that under the condition that the autocorrelation K is continuously differentiable, the corresponding stationary process $F(t)$ must satisfy the supremum bound (2.4). Thanks to stationarity, it suffices to prove (2.4) holds for the time interval $[0, T]$, namely, for all $T \ge 0$,

$$\mathbb{E}\sup_{0\le t\le T} F(t)^2 < \infty.$$

For convenience, we first recap several notions from the technique of generic chaining in [33, Chapter 2]. Consider the time interval $[0, T]$ and the distance

$$d(s, t) \overset{def}{=} \sqrt{\mathbb{E}|F(t) - F(s)|^2}.$$

It is well-known that d is a metric in $[0, T]$. For a set $A \subset [0, T]$, we denote by $\triangle(A)$ the diameter of A with respect to metric d, that is

$$\triangle(A) \overset{def}{=} \inf_{s,t\in A} d(s, t).$$

Next, we provide the definition of an *admissible sequence*.

Definition B.1. An admissible sequence is an increasing sequence $\{A_n\}_{n\geq 0}$ of partitions of $[0,T]$ such that $A_0=[0,T]$ and for all $n\geq 1$, $\mathrm{card}(A_n)$ is at most $N_n=2^{2^n}$.

Here, increasing sequence means every set of A_{n+1} is contained in some set of A_n.

Given an admissible sequence A_n and a time $t\in[0,T]$, we denote by $A_n(t)$ the element in A_n that contains t and define $\gamma_2(T,d)$ given by

$$\gamma_2(T,d)\overset{def}{=}\inf\sup_{t\in[0,T]}\sum_{n\geq 0}2^{\frac{n}{2}}\triangle(A_n(t)),$$

where the infimum is taken over all admissible sequences. We now state the following result asserting that under the conditions imposed on $F(t)$, $\mathbb{E}\sup_{0\leq t\leq T}F(t)^2$ is always finite.

Lemma B.2. *Let $F(t)$ be a mean-zero Gaussian stationary process whose covariance function K is in $C^1(\mathbb{R})$. Then, for all $T\geq 0$, there exists a positive constant $c(T)$ such that*

$$\mathbb{E}\sup_{0\leq t\leq T}F(t)^2\leq c(T). \tag{B.1}$$

Proof. We first observe that

$$\begin{aligned}\sup_{0\leq t\leq T}F(t)^2=\sup_{0\leq t\leq T}\big(F(t)-F(0)+F(0)\big)^2&\leq 2\sup_{0\leq t\leq T}\big(F(t)-F(0)\big)^2+2F(0)^2\\ &\leq 2\sup_{0\leq t,s\leq T}\big(F(t)-F(s)\big)^2+2F(0)^2,\end{aligned}$$

whence

$$\mathbb{E}\sup_{0\leq t\leq T}F(t)^2\leq 2\mathbb{E}\sup_{0\leq t,s\leq T}\big(F(t)-F(s)\big)^2+2K(0).$$

It therefore suffices to establish an upper bound for $\mathbb{E}\sup_{0\leq t,s\leq T}\big(F(t)-F(s)\big)^2$.

Now, since $F(t)$ is a mean-zero Gaussian process, $F(t)$ satisfies [33, inequality (1.4)], that is for all $r>0$

$$\mathbb{P}(|F(s)-F(t)|\geq r)\leq 2\exp\Big(-\frac{r^2}{2d(s,t)^2}\Big). \tag{B.2}$$

Indeed, by Markov's inequality,

$$\mathbb{P}(|F(s)-F(t)|\geq r)=\mathbb{P}\Big(\exp\Big(-\frac{r^2}{2|F(s)-F(t)|^2}\Big)\geq\frac{1}{\sqrt{e}}\Big)\leq 2\mathbb{E}\exp\Big(-\frac{r^2}{2|F(s)-F(t)|^2}\Big).$$

Observe that $f(x)=e^{-r/x}$ is concave down on $(0,\infty)$. So that, Jensen's inequality implies

$$\mathbb{E}\exp\Big(-\frac{r^2}{2|F(s)-F(t)|^2}\Big)\leq\exp\Big(-\frac{r^2}{2\mathbb{E}|F(s)-F(t)|^2}\Big),$$

which proves (B.2). Now, in light of [33, inequality (2.49)], there exists a positive constant C independent of T such that

$$\mathbb{E}\sup_{0\leq t,s\leq T}\big(F(t)-F(s)\big)^2\leq C\gamma_2(T,d).$$

It remains to show that $\gamma_2(T,d)$ is finite. To this end, consider the the following sequence $\{\tilde{A}_n\}_{n=0}$ given by

$$\tilde{A}_0=[0,T]\quad\text{and}\quad\tilde{A}_n=\Big[0,\frac{T}{N_n}\Big)\cup\Big[\frac{T}{N_n},\frac{2T}{N_n}\Big)\dots\Big[\frac{(N_n-1)T}{N_n},T\Big],\quad n\geq 1.$$

It is straightforward to check that $\tilde{A}_n$ is an admissible sequence. For each $t \in [0,T]$, by definition of $\triangle$, we note that

$$\triangle(\tilde{A}_n(t)) = \sup_{s,r\in\tilde{A}_n(t)} d(s,t) = \sup_{s,r\in\tilde{A}_n(t)} \sqrt{\mathbb{E}(F(s)-F(r))^2}$$
$$= \sup_{s,r\in\tilde{A}_n(t)} \sqrt{2(K(0)-K(|s-r|)}.$$

By the choice of $\tilde{A}_n$, for all $r,s\in\tilde{A}_N(t)$, $|r-s|\le T/N_n$. So that,

$$\sup_{s,r\in\tilde{A}_n(t)} \sqrt{2(K(0)-K(|s-r|)} = \sup_{0\le s\le T/N_n} \sqrt{2(K(0)-K(s))}.$$

Since $K\in C^1(\mathbb{R})$, by the Mean-Value Theorem, for $s\in[0,T/N_n]$

$$|K(0)-K(s)| \le \max_{r\in[0,T]} |K'(r)|\cdot s \le \max_{r\in[0,T]} |K'(r)|\cdot \frac{T}{N_n},$$

implying

$$\triangle(\tilde{A}_n(t)) \le \sqrt{\frac{2T}{N_n}\max_{r\in[0,T]}|K'(r)|}.$$

Turning back to $\gamma_2(T,d)$, we note that

$$\gamma_2(T,d) \le \sup_{t\in[0,T]}\sum_{n\ge0} 2^{n/2}\triangle(\tilde{A}_n(t)) \le \sum_{n\ge0} 2^{n/2}\sqrt{\frac{2T}{N_n}\max_{r\in[0,T]}|K'(r)|}$$
$$= \sqrt{T\max_{r\in[0,T]}|K'(r)|}\sum_{n\ge0}\frac{2^{\frac{n+1}{2}}}{\sqrt{N_n}}$$
$$= \sqrt{T\max_{r\in[0,T]}|K'(r)|}\Big(\sqrt{2}+\sum_{n\ge1}\frac{2^{\frac{n+1}{2}}}{2^{2^{n-1}}}\Big),$$

which is clearly finite. Altogether, we arrive at the bound

$$\mathbb{E}\sup_{0\le t\le T} F(t)^2 \le C\sqrt{T\max_{r\in[0,T]}|K'(r)|} + 2K(0),$$

thereby establishing (B.1) and completing the proof. □

References

[1] Joseph Abate and Ward Whitt. Infinite-series representations of Laplace transforms of probability density functions for numerical inversion. *Journal of the Operations Research Society of Japan*, 42(3):268–285, 1999.

[2] Yuri Bakhtin and Jonathan C. Mattingly. Stationary solutions of stochastic differential equations with memory and stochastic partial differential equations. *Communications in Contemporary Mathematics*, 7(05):553–582, 2005.

[3] Richard F. Bass. *Diffusions and elliptic operators*. Springer Science & Business Media, 1998.

[4] Florian Conrad and Martin Grothaus. Construction, ergodicity and rate of convergence of N-particle Langevin dynamics with singular potentials. *Journal of Evolution Equations*, 10(3):623–662, 2010.

[5] Ben Cooke, David P. Herzog, Jonathan C. Mattingly, Scott A. McKinley, and Scott C. Schmidler. Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard–Jones-like repulsive potential. *Communications in Mathematical Sciences*, 15(7):1987–2025, 2017.

[6] Gustavo Didier and Hung Nguyen. Asymptotic analysis of the mean squared displacement under fractional memory kernels. *SIAM Journal on Mathematical Analysis*, 52(4):3818–3842, 2020.

[7] Joseph L. Doob. The Brownian movement and stochastic equations. *Annals of Mathematics*, pages 351–369, 1942.

[8] Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2011.

[9] S. K. Ghosh, A. G. Cherstvy, D. S. Grebenkov, and R. Metzler. Anomalous, non-Gaussian tracer diffusion in crowded two-dimensional environments. *New Journal of Physics*, 18(1):013027, 2016.

[10] Nathan Glatt-Holtz, Jonathan C. Mattingly, and Geordie Richards. On unique ergodicity in nonlinear stochastic partial differential equations. *J. Stat. Phys.*, 166(3-4):618–649, 2017.

[11] Nathan E. Glatt-Holtz, David P. Herzog, Scott McKinley, and Hung Nguyen. The generalized Langevin equation with a power-law memory in a nonlinear potential well. *Nonlinearity*, 33(6):2820–2852, 2020.

[12] Igor Goychuk. Viscoelastic subdiffusion: generalized Langevin equation approach. *Advances in Chemical Physics*, 150:187, 2012.

[13] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):223–259, 2011.

[14] David P. Herzog and Jonathan C. Mattingly. Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials. *Communications on Pure and Applied Mathematics*, 72(10):2231–2255, 2019.

[15] Kiyosi Itô and Makiko Nisio. On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.*, 4(3):1–75, 1964.

[16] Rep Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255, 1966.

[17] A. J. Levine and T. C. Lubensky. One-and two-particle microrheology. *Physical Review Letters*, 85(8):1774–1777, 2000.

[18] Thomas G. Mason and D. A. Weitz. Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids. *Physical Review Letters*, 74(7):1250, 1995.

[19] Jonathan C. Mattingly. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Comm. Math. Phys.*, 230(3):421–462, 2002.

[20] Jonathan C. Mattingly. On recent progress for the stochastic Navier Stokes equations. In *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. XI, 52. Univ. Nantes, Nantes, 2003.

[21] Jonathan C. Mattingly, Andrew M. Stuart, and Desmond J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Processes and their Applications*, 101(2):185–232, 2002.

[22] Scott A. McKinley and Hung Nguyen. Anomalous diffusion and the generalized Langevin equation. *SIAM Journal on Mathematical Analysis*, 50(5):5119–5160, 2018.

[23] Y. Meroz and I. M. Sokolov. A toolbox for determining subdiffusive mechanisms. *Physics Reports*, 573:1–29, 2015.

[24] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77, 2000.

[25] Hazime Mori. A continued-fraction representation of the time-correlation functions. *Progress of Theoretical Physics*, 34(3):399–416, 1965.

[26] Hazime Mori. Transport, collective motion, and Brownian motion. *Progress of theoretical physics*, 33(3):423–455, 1965.

[27] Bernt Øksendal. *Stochastic differential equations*. Springer, 2003.

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

[29] Grigorios A Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, volume 60. Springer, 2014.

[30] M.J. Saxton. Anomalous diffusion due to obstacles: a Monte Carlo study. *Biophysical Journal*, 66(2):394–401, 1994.

[31] M.J. Saxton. Anomalous diffusion due to binding: a Monte Carlo study. *Biophysical Journal*, 70(3):1250–1262, 1996.

[32] Igor M. Sokolov. Statistics and the single molecule. *Physics*, 1:8, 2008.

[33] Michel Talagrand. *Upper and lower bounds for stochastic processes: modern methods and classical problems*, volume 60. Springer Science & Business Media, 2014.

[34] Cédric Villani. Hypocoercivity. *arXiv preprint math/0609050*, 2006.

[35] E Weinan and Di Liu. Gibbsian dynamics and invariant measures for stochastic dissipative pdes. *Journal of Statistical Physics*, 108(5-6):1125–1156, 2002.

[36] E Weinan, Jonathan C. Mattingly, and Yakov Sinai. Gibbsian Dynamics and Ergodicity for the Stochastically Forced Navier–Stokes Equation. *Communications in Mathematical Physics*, 224(1):83–106, 2001.

[1] Department of Mathematics, Iowa State University, Iowa, USA

[2] Department of Mathematics, Duke University, North Carolina, USA

[3] Department of Mathematics, University of California, Los Angeles, California, USA