Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 128 (2018) 3253–3272

stochastic processes and their applications

www.elsevier.com/locate/spa

Ergodicity of scalar stochastic differential equations with Hölder continuous coefficients

Luu Hoang Duc[a,b,*], Tat Dat Tran[a], Jürgen Jost[a,c]

[a] *Max-Planck-Institute for Mathematics in the Sciences, Inselstrasse 22, D-04103 Leipzig, Germany*
[b] *Institute of Mathematics, Viet Nam Academy of Science and Technology, 18 Hoang Quoc Viet Road, 10307 Ha Noi, Viet Nam*
[c] *Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM 87501, USA*

Received 25 August 2016; received in revised form 7 August 2017; accepted 30 October 2017
Available online 8 November 2017

Abstract

It is well-known that for a one dimensional stochastic differential equation driven by Brownian noise, with coefficient functions satisfying the assumptions of the Yamada–Watanabe theorem (Yamada and Watanabe, 1971, [31,32]) and the Feller test for explosions (Feller, 1951, 1954), there exists a unique stationary distribution with respect to the Markov semigroup of transition probabilities. We consider systems on a restricted domain D of the phase space $\mathbb{R}$ and study the rate of convergence to the stationary distribution. Using a geometrical approach that uses the so called *free energy function* on the density function space, we prove that the density functions, which are solutions of the Fokker–Planck equation, converge to the stationary density function exponentially under the Kullback–Leibler divergence, thus also in the total variation norm. The results show that there is a relation between the Bakry–Émery curvature dimension condition and the dissipativity condition of the transformed system under the Fisher–Lamperti transformation. Several applications are discussed, including the Cox–Ingersoll–Ross model and the Ait-Sahalia model in finance and the Wright–Fisher model in population genetics.
© 2017 Elsevier B.V. All rights reserved.

Keywords: Stationary distributions; Invariant measure; Fokker–Planck equation; Kullback–Leibler divergence; Cox–Ingersoll–Ross model; Ait-Sahalia model; Wright–Fisher model

* Corresponding author.
E-mail addresses: duc.luu@mis.mpg.de, lhduc@math.ac.vn (L.H. Duc), trandat@mis.mpg.de (T.D. Tran), jost@mis.mpg.de (J. Jost).

https://doi.org/10.1016/j.spa.2017.10.014
0304-4149/© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Stochastic differential equations with Hölder continuous coefficients arise as models in many sciences. An instance is the Wright–Fisher diffusion model for the genetic drift of alleles with mutation [15]; in the one dimensional case (that is, when there are only two alleles present), it is of the form

$$dX_t = [\theta_1 - (\theta_1 + \theta_2)X_t]dt + \sqrt{X_t(1 - X_t)}dW_t, \quad X_t \in [0, 1],$$

where we assume that the mutation rates $\theta_1, \theta_2 > 0$. Another example is the Cox–Ingersoll–Ross [5] model for the short term interest rate

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_t \geq 0,$$

where $k, \theta, \sigma > 0$ satisfy $k\theta \geq \frac{\sigma^2}{2}$.

For such nonlinear systems, the diffusion coefficient is only Hölder continuous on a restricted domain [11], therefore the existence and uniqueness of a solution cannot be proved using classical arguments like contraction mappings, but one rather needs to invoke the Yamada and Watanabe theorem [31,32,24]. Moreover, the solution in general does not depend differentiably on the initial values in the state space, and thus the linearization method for studying the stability problem fails to apply here.

In this paper, we study the Markov semigroup generated by such a system; this is a strong Feller process [16,17] provided that the Feller test of explosion succeeds. In particular, assume that D is a Polish space (complete metric and separable). A probability measure μ on D is called *stationary (invariant)* with respect to the (Markov) diffusion process X_t with semigroup $(T_t)_{t\geq 0}$ on $C_0^\infty(D)$ and generator G if

$$\int_D T_t f(y)\mu(dy) = \int_D f(y)\mu(dy), \quad \forall t \geq 0,\ f \in C_0^\infty(D),$$

or equivalently (due to [19] Theorem 2.3)

$$\int_D Gf(y)\mu(dy) = 0, \quad \forall f \in C_0^\infty(D).$$

It is called *reversible* if

$$\int_D g(y)T_t f(y)\mu(dy) = \int_D f(y)T_t g(y)\mu(dy), \quad \forall t \geq 0,\ f, g \in C_0^\infty(D),$$

or equivalently

$$\int_D g(y)Gf(y)\mu(dy) = \int_D f(y)Gg(y)\mu(dy), \quad \forall f, g \in C_0^\infty(D).$$

Also, it is well-known that the density function of the transition probability satisfies the Fokker–Planck equation. The existence of stationary measures with respect to the Markov semigroup is deduced from the Krylov–Bogoliubov theorem [9]. However, when the process is reversible (with appropriate boundary conditions), we can prove directly that there exists a unique stationary measure for the Markov semigroup which can be written as a Gibbs measure $\mu_\infty(dy) = u_\infty(y)dy = \frac{e^{-\psi(y)}}{Z}dy$.

Another important issue is the rate of convergence to the stationary distribution. It is well known from the Harris theorem [23] that if there exists a Lyapunov type function, an initial distribution under the Markov semigroup will converge exponentially to the stationary

distribution. Such a Lyapunov function is constructed in Hairer [22] for the system with additive noise, however it is not explicitly given in general.

Our approach here is to present an explicit construction of a Lyapunov type function for the Markov semigroup on the space of probability measures. Specifically, we consider only those systems in which there exists a so-called *free energy functional* defined as the sum of the *relative entropy* and the *potential energy* (see, e.g. [25]) along the flow of densities $\{u(t, x, \cdot)\}_{t\geq 0}$. We then prove that the free energy functional decreases in time t along the flow of densities.

Our main result is the convergence rate of the densities under the flow to the stationary density using the geometrical method by Bakry–Émery [3,21]. By considering the backward generator of the Markov semigroup $(L^*, \mathcal{D}(L^*))$, the *carré du champ operator*, the *iterated carré du champ operator* of L^* and the *Bakry–Émery curvature-dimension condition*, we derive a good estimate for the exponential rate of convergence of the density function $u(t, x, \cdot)$ to $u_\infty(\cdot)$ with respect to the Kullback–Leibler divergence.

It is interesting to note that for one dimensional systems, there is a relation between the Bakry–Émery curvature dimension condition and the dissipativity condition of the transformed system under the so called *Fisher–Lamperti transformation* which is supposed to be continuous. This transformation helps to prove the existence of a global random attractor (see [2,6,7,27]) for the Cox–Ingersoll–Ross model (34), the Wright–Fisher model (37) and the Ait-Sahalia model (39).

2. Existence and uniqueness of the stationary distribution

We consider the following one dimensional stochastic system

$$dX_t = a(X_t)dt + b(X_t)dW_t \tag{1}$$

where $D \in \Big\{[d_-, d_+], (-\infty, d_+], [d_-, \infty), \mathbb{R}\Big\}$, and $a, b : D \to \mathbb{R}$ are C^3-functions on the interior of a domain $D \subset \mathbb{R}$ (D can contain ∞), which satisfy the following assumptions.

Hypothesis A. The functions a, b satisfy

- there exists a positive increasing concave function $L : [0, \infty) \to [0, \infty)$ with $L(0) = 0$, such that $|a(x) - a(y)| \leq L(|x - y|)$ for all $x, y \in D$ and $\int_{0^+}^{p} L(r)^{-1}dr = \infty$ for every $p > 0$;
- there exists a positive increasing function $\rho : [0, \infty) \to [0, \infty)$ with $\rho(0) = 0$, such that $|b(x) - b(y)| \leq \rho(|x - y|)$ for all $x, y \in D$, and $\int_{0^+}^{p} \rho(r)^{-2}dr = \infty$ for every $p > 0$.
- a, b satisfy the linear growth condition: there exists a constant $K > 0$ such that $|a(x)|^2 + |b(x)|^2 \leq K(1 + |x|^2)$ for all $x \in D$.
- boundary conditions: $a(d_-) > 0 > a(d_+)$; $b(d_-) = b(d_+) = 0$, $b(x) > 0$ for $x \in (d_-, d_+)$.

Theorem 1. *Under Hypothesis A for the coefficient functions a, b there exists a unique solution of (1) with initial condition $X_0 = x \in D$.*

Proof. First, we extend (if necessary) the coefficient functions of (1) from the domain D into the whole real line by considering the system

$$dX_t = \bar{a}(X_t)dt + \bar{b}(X_t)dW_t \tag{2}$$

where $\bar{a}(x) \equiv a(x)$, $\bar{b}(x) \equiv b(x)$ for all $x \in D$ and $\bar{a}(x) = a(d_-)(> 0)$, $\bar{b}(x) = b(d_-)(= 0)$ for all $x < d_-$, $\bar{a}(x) = a(d_+)(< 0)$, $\bar{b}(x) = b(d_+)(= 0)$ for all $x > d_+$. It is easy to check that $\bar{a}, \bar{b}$

then satisfy Hypothesis A. Hence by using the same arguments as in the proof of the Yamada and Watanabe theorem (see [31,32] or [24, Theorem 3.2, p. 168]), there exists a unique solution $\bar{X}_t$ of (2) with initial condition X_0. For that, we need to note that, while the proof of [24, Theorem 3.2, p. 168] is for bounded $\bar{a}$ and $\bar{b}$, it also has a localized version for unbounded $\bar{a}, \bar{b}$ using stopping times. The linear growth condition of $\bar{a}, \bar{b}$ is then sufficient to conclude that the solution of (2) is no where explosive (see [24, Theorem 2.4, p. 163]).

Next, we need to prove that all solutions of (2) stay in the domain D almost surely, i.e. $\mathbb{P}(\bar{X}_t(\cdot, x) \in D, \forall t \in [0, T]) = 1$ for any $x \in D$. It is enough to prove that in case the left boundary of D is d_- then $\mathbb{P}(\bar{X}_t(\cdot, x) \geq d_-, \forall t \in [0, T]) = 1$. To do that, we follow the Deelstra and Delbaen technique [11] by introducing for any $\epsilon > 0$ stopping times

$$\begin{aligned}
\tau_1 &= \inf\{t \in [0, T] | \bar{X}_t < d_- - \epsilon\} \\
\sigma_1^1 &= \inf\{t \in [0, T] | \bar{X}_t < d_- - 2\epsilon\} \\
\sigma_1^2 &= \inf\{t > \tau_1 | \bar{X}_t = d_-\} \\
\sigma_1 &= \sigma_1^1 \wedge \sigma_1^2 \wedge T.
\end{aligned}$$

Trivially $\tau_1 < \sigma_1 \leq T$. Define the set $A := \{\inf_{u \in [0,T]} \bar{X}_u < d_- - 2\epsilon\}$. If $\sigma_1^1 \leq \sigma_1^2$ then $\sigma_1 = \sigma_1^1$, hence $\bar{X}_{\sigma_1} - \bar{X}_{\tau_1} = -\epsilon$ due to the continuity of $\bar{X}$ over time. But on the other hand, $\bar{b}(u) = b(d_-) = 0$ and $\bar{a}(u) = a(d_-) > 0$ for every $u \in [\tau_1, \sigma_1]$, thus by using Ito's formula for the stopping times,

$$\bar{X}_{\sigma_1} - \bar{X}_{\tau_1} = \int_{\tau_1}^{\sigma_1} \bar{a}(u) du + \int_{\tau_1}^{\sigma_1} \bar{b}(u) dW_u > 0,$$

which is a contradiction. We conclude that on this set, $\sigma_1^2 < \sigma_1^1 \leq T$ almost surely. By definition, we conclude that $\bar{X}_{\sigma_1} = d_-$ on A.

Similarly, define some more stopping times:

$$\begin{aligned}
\tau_2 &= \inf\{\sigma_1 < t \in [0, T] | \bar{X}_t < d_- - \epsilon\} \wedge T \\
\sigma_2^1 &= \inf\{\sigma_1 < t \in [0, T] | \bar{X}_t < d_- - 2\epsilon\} \\
\sigma_2^2 &= \inf\{t > \tau_2 | \bar{X}_t = d_-\} \\
\sigma_2 &= \sigma_2^1 \wedge \sigma_2^2 \wedge T.
\end{aligned}$$

Analogously on A, $\tau_1 < \sigma_1 < \tau_2 < \sigma_2 \leq T$ and $\bar{X}_{\tau_2} = d_- - \epsilon$. Therefore, we can repeat this argument and conclude that on A there exists a strictly increasing sequence of stopping times $\tau_1 < \sigma_1 < \cdots < \tau_n < \sigma_n < \cdots \leq T$. This sequence converges to a limit $\tau \leq T$, hence all the subsequences (τ_n) and (σ_n) have to converge to τ. However, $\bar{X}_{\tau_n} = d_- - \epsilon$ while $\bar{X}_{\sigma_n} = d_-$ on A. By the continuity of $\bar{X}$ over time, it follows that A has measure zero, or equivalently $\mathbb{P}(\inf_{u \in [0,T]} \bar{X}(u) < d_- - 2\epsilon) = 0$. Since it is true for any $\epsilon > 0$, we have proved that $\mathbb{P}(\inf_{u \in [0,T]} \bar{X}(u) < d_-) = 0$.

Finally, by the definition of $\bar{a}, \bar{b}$, the solution $\bar{X}$ is actually also a solution of (1), thus we conclude the existence and uniqueness of (1) on the domain D for the time interval $[0, T]$. As T varies, we actually can extend the solution of (2) to the whole time interval $[0, \infty)$. □

It is well-known that the density function of the Markov semigroup with transition probability $\mathcal{P}_t(x, O) := \mathbb{P}\{X_t(\cdot, x) \in O | X_0 = x\}$ satisfies the Fokker–Planck equation

$$\begin{aligned}\partial_t u(t, x, y) &= Lu(t, x, y)\\ &= -\partial_y[a(y)u(t, x, y)] + \frac{1}{2}\partial_{yy}[b^2(y)u(t, x, y)]\\ &= \partial_y\Big(-a(y)u(t, x, y) + \frac{1}{2}\partial_y\big(b^2(y)u(t, x, y)\big)\Big).\end{aligned} \tag{3}$$

In this section, we only consider systems that satisfy special conditions which ensure the existence of a free energy function. To do that, we first consider the definition.

Definition 2. The family of densities $\{u(t, \cdot)\}_{t\geq 0}$ on a $\sigma-$finite measure space (D, μ) is said to satisfy the condition $I(A, \psi)$ if it solves a diffusion equation of the form

$$\partial_t u(t, y) = \partial_y\Big(A(y)u(t, y)\partial_y\big[\log u(t, y) + \psi(y)\big]\Big), \tag{4}$$

where $A(y) = \frac{1}{2}b^2(y)$. This family of densities $\{u(t, \cdot)\}_{t\geq 0}$ is said to satisfy the condition $II(A, \psi)$ if in addition to $I(A, \psi)$, one has

$$\int_D e^{-\psi(y)}\mu(dy) < \infty. \tag{5}$$

We rewrite a system that satisfies condition $I(A, \psi)$ in the one dimensional case as

$$\partial_t u(t, x, y) = \partial_y\Big(A(y)u(t, x, y)\partial_y\big(\log u(t, x, y) + \psi(y)\big)\Big).$$

We introduce the potential energy function $\psi(y)$

$$\psi(y) := \log A(y) - G(y) = \log\frac{1}{2}b(y)^2 - G(y) \tag{6}$$

where $G(y)$ is the indefinite integral of $\frac{a(y)}{A(y)}$, i.e. $G(y) = \int^y \frac{a(y)}{A(y)}dy$. It is easy to check that

$$A(y)u(t, x, y)\partial_y\big(\log u(t, x, y) + \psi(y)\big) = -a(y)u(t, x, y) + \frac{1}{2}\partial_y\big(b^2(y)u(t, x, y)\big).$$

Thus Eq. (3) satisfies the condition $I(A, \psi)$. We will then write the Kolmogorov (forward/backward) operators in this form, i.e.

$$Lu(y) = \partial_y\Big(A(y)\partial_y u(y)\Big) + \partial_y\Big(A(y)u(y)\partial_y\psi(y)\Big),$$

and

$$L^*u(x) = \partial_x\Big(A(x)\partial_x u(x)\Big) - \partial_x\psi(x)A(x)\partial_x u(x).$$

We need the following assumption.

Hypothesis B. The potential energy function ψ satisfies the condition $II(A, \psi)$ and

$$A(y)e^{-\psi(y)}\Big|_{\partial D} = \lim_{y\to\partial D} e^{G(y)} = 0 \tag{7}$$

If ψ satisfies the Hypothesis B, then we have immediately.

Lemma 3. *The measure*

$$\mu_\infty(dy) = u_\infty(y)dy = \frac{e^{-\psi(y)}}{Z}dy,$$

(where $Z = \int_D e^{-\psi(y)}dy < \infty$ due to condition (5), is the normalization coefficient), is reversible with respect to L^*, i.e.

$$\int_D fL^*g d\mu_\infty = \int_D gL^*f d\mu_\infty, \quad \forall f, g \in C^2(D). \tag{8}$$

Proof. Integrating by parts, we have

$$\begin{aligned}
(8) &\Leftrightarrow \int_D f[(Ag_y)_y - A\psi_y g_y]e^{-\psi}dy = \int_D g[(Af_y)_y - A\psi_y f_y]e^{-\psi}dy \\
&\Leftrightarrow \int_D \Big[f(Ag_y)_y - g(Af_y)_y\Big]e^{-\psi}dy = \int_D A(fg_y - gf_y)\psi_y e^{-\psi}dy \\
&\Leftrightarrow \int_D \Big[A(fg_{yy} - gf_{yy}) + A_y(fg_y - gf_y)\Big]e^{-\psi}dy = \int_D A(fg_y - gf_y)d(-e^{-\psi}) \\
&\Leftrightarrow \int_D \Big[A(fg_{yy} - gf_{yy}) + A_y(fg_y - gf_y)\Big]e^{-\psi}dy \\
&\quad = -(fg_y - gf_y)Ae^{-\psi}\Big|_{\partial D} + \int_D d\Big[A(fg_y - gf_y)\Big]e^{-\psi}.
\end{aligned}$$

From assumption (7), the first term on the right hand side of the last equation is zero, which then implies that

$$\begin{aligned}
(8) &\Leftrightarrow \int_D \Big[A(fg_{yy} - gf_{yy}) + A_y(fg_y - gf_y)\Big]e^{-\psi}dy \\
&\quad = \int_D \Big[A(fg_y - gf_y)_y + A_y(fg_y - gf_y)\Big]e^{-\psi}dy \\
&\Leftrightarrow \int_D A(fg_{yy} - gf_{yy})e^{-\psi}dy = \int_D A(fg_y - gf_y)_y e^{-\psi}dy.
\end{aligned}$$

The last equation holds for every $f, g \in C^2(D)$. This implies the proof. □

Theorem 4. *Assume that a, b satisfy Hypotheses A, B. Then there exists a unique stationary measure for the Markov semigroup $\mathcal{P}_t(x, O) := \mathbb{P}\{X_t(\cdot, x) \in O | X_0 = x\}$. Moreover, the stationary distribution is of the form*

$$\mu_\infty(dy) = u_\infty(y)dy = \frac{e^{-\psi(y)}}{Z}dy \tag{9}$$

Proof. Since the original system is one dimensional, we can prove by using Lemma 3 that the invariant measure, if it exists, is reversible. We can then follow similar arguments as in [29] (Theorem 3.7) as well as in [30] to prove the existence and uniqueness of the invariant measure. □

Remark 5. In general, when the probability measure μ_∞ is not reversible, we need to prove that the Markov semigroup $\mathcal{P}_t(x, \cdot)$ is strongly Feller and tight. The existence is then a consequence of the Krylov–Bogoliubov theorem. Moreover, we need to prove that $\mathcal{P}_t$ is irreducible and thus regular by applying the Khasminskii theorem, and then apply the Doob theorem to conclude that $\lim_{t\to\infty}\mathcal{P}_t(x, O) = \mu_\infty(O)$ which in particular proves the uniqueness of μ_∞.

For more information about this method, see Prato and Zabczyk [9].

We now would like to find conditions so that the process will never reach the boundary points with probability one. To do so, we first recall some concepts from potential theory:

Definition 6. Let $(X_t)_{t\geq 0}$ be a reversible process in D with generator L^* and reversible measure $u_\infty(x)dx$. The corresponding Dirichlet form is

$$D(f) := -\int_D f(x)Lf(x)u_\infty(x)dx. \tag{10}$$

For $c > 0$, we also define the quadratic form by

$$D_c(f) = D(f) + c\int_D f(x)^2 u_\infty(x)dx. \tag{11}$$

Definition 7. We define the capacity of an open set $O \subseteq D$ by

$$cap_c(O) := \inf_{f\in \mathbf{F}_O^c} D_c(f), \qquad \mathbf{F}_O^c := \left\{f \in \mathcal{D}_c \,:\, f(x) \geq 1,\ \forall x \in O\right\} \tag{12}$$

For an arbitrary set $B \subseteq D$ the capacity of B is defined as

$$cap_c(B) := \inf_{O \text{ open}\,:\, O\supseteq B} cap_c(O) \tag{13}$$

Theorem 8. *If in addition to Hypothesis A, the coefficients satisfy*

$$-a'(x) + b''(x)b(x) + b'^2(x) \geq c > 0, \forall x \in D, \tag{14}$$

then the solution is in D^o with probability one.

Proof. First, observe that an eigenfunction f of L^* for the eigenvalue c is given by

$$f_\varepsilon(x) = \frac{u_\infty(d_- + \varepsilon)}{u_\infty(x)}\frac{G(x)}{G(d_- + \varepsilon)},$$

where $G(x) = \int_x^{d_+} \frac{u_\infty(y)}{b^2(y)}dy$. Similar to the arguments in [4], we can prove by direct computation that, under this additional condition (14),

$$cap_c(\partial D) = \lim_{\varepsilon\to 0} D_c(f_\varepsilon) = 0.$$

Then, we apply a classical result in potential theory (see for example [18], Theorem 4.3.1, page 103) to show that the points with capacity zero are exceptional, which means that the boundary can never be reached with probability one. □

3. Rate of convergence to the stationary distribution

In this section we present a direct construction of a Lyapunov-type function for the Markov semigroup on the space of probability measures. It is worth noting here that the existence of a Lyapunov function for a Markov semigroup on the phase space is assumed in Harris' theorem (see [23] or Hairer [22]), but such a function in general is not explicitly given.

Definition 9. For a family of densities $\{u(t,\cdot)\}_{t\geq 0}$ on a $\sigma-$finite measure space (D,μ) with condition $I(A,\psi)$,

- the potential energy functional is defined by

$$\Psi(u(t,\cdot)) := \int_D u(t,y)\psi(y)\mu(dy); \tag{15}$$

- the (negative) entropy functional is defined by

$$S_\mu(f) = \int_D f \log f d\mu; \tag{16}$$

- the free energy functional is defined by

$$F(u(t,\cdot)) := \int_D u(t,y)\Big(\log u(t,y) + \psi(y)\Big)\mu(dy) = S_\mu(u(t,\cdot)) + \Psi(u(t,\cdot)). \tag{17}$$

Definition 10. Let f_1, f_2 be densities on a $\sigma-$finite measure space (D, μ). The relative entropy (Kullback–Leibler divergence) of f_1 with respect to f_2 is

$$D_{\mathrm{KL}}(f_1 \parallel f_2) := \begin{cases} \displaystyle\int_D f_1(y) \log \frac{f_1(y)}{f_2(y)} \mu(dy), & \text{if } \mathrm{supp}(f_1) \subset \mathrm{supp}(f_2) \\ \infty, \ \text{otherwise} \end{cases}$$

Before presenting our main theorem, we need the following assumption.

Hypothesis C. There exists a $\rho > 0$ such that

$$\frac{1}{2} b b_{yy} - a_y + \frac{a b_y}{b} \geq \rho > 0. \tag{18}$$

Theorem 11. *Assume that a, b satisfy Hypotheses A, B, C. Then the density function $u(t,x,\cdot)$ converges to $u_\infty(\cdot)$ exponentially w.r.t. the Kullback–Leibler divergence, i.e.*

$$D_{\mathrm{KL}}(u(t,x,\cdot)\|u_\infty(\cdot)) \leq e^{-\rho t} D_{\mathrm{KL}}(u(0,x,\cdot)\|u_\infty(\cdot)) \tag{19}$$

Proof. The proof follows the approach in Tran et al. [30] and is divided into several steps.

Step 1: First, we can prove that $F(u(t,x,\cdot))$ decreases in time t along the flow of densities. In fact,

$$\partial_t F(u(t,x,\cdot)) = -\int_D I(t,x,y) dy, \tag{20}$$

where

$$I(t,x,y) = \frac{1}{2} b^2(y) u(t,x,y) \Big[\partial_y \psi(y) + \frac{1}{u(t,x,y)} \partial_y u(t,x,y)\Big]^2 \geq 0.$$

Step 2: Next, assume that there exists a unique stationary distribution $\mu_\infty(dy) = u_\infty(y)dy$. We focus on the rate of the convergence of u to u_∞. Putting

$$h(t,x,y) := \frac{u(t,x,y)}{u_\infty(y)},$$

we shall investigate the rate of the convergence of h to 1.

The stationary density is the Gibbs density function

$$u_\infty(y) = \frac{e^{-\psi(y)}}{Z},$$

which is independent of x. Thus

$$\log u_\infty(y) + \psi(y) = -\log Z.$$

Since Z is independent of y, this implies

$$\begin{aligned}\partial_y(\log u(t,x,y)+\psi(y)) &= \partial_y\Big(\log\frac{u(t,x,y)}{u_\infty(y)}\Big)+\partial_y(\log u_\infty(y)+\psi(y))\\ &= \partial_y(\log h(t,x,y)).\end{aligned}$$

Therefore, using the same computations as in [30], we can derive the following partial differential equation for h from that of u.

$$\partial_t h(t,x,y)=\partial_y\Big[A(y)\partial_y h(t,x,y)\Big]-\partial_y\psi(y)A(y)\partial_y h(t,x,y)=L^*h(t,x,y), \tag{21}$$

where

$$A(y):=\frac{1}{2}b(y)^2. \tag{22}$$

Also, the difference of the current and the final free energy is equal to the relative entropy (Kullback–Leibler divergence) between the corresponding densities and also equal to the (negative) entropy of their ratio with respect to the stationary probability measure:

$$F(u(t,x,\cdot))-F(u_\infty(\cdot))=D_{\mathrm{KL}}(u(t,x,\cdot)\parallel u_\infty(\cdot))=S_{\mu_\infty}(h(t,x,\cdot))\geq 0. \tag{23}$$

Moreover, the rate of change of the free energy functional is equal to the negative of the entropy production

$$\frac{d}{dt}S_{\mu_\infty}(h(t,x,\cdot))=\partial_t F(u(t,x,\cdot))=-J_{\mu_\infty}(h):=-\int_D\frac{A(y)[\partial_y h(t,x,y)]^2}{h(t,x,y)}\mu_\infty(dy). \tag{24}$$

Step 3: Now we check the convergence rate of the density function to the stationary distribution. For that we need to apply the geometric theory of Markov operators. From now on, for simplicity, we will write

$$q_y:=\partial_y q(y),\ q_{yy}:=\partial_{yy}q(y),\ q_{yyy}:=\partial_{yyy}q(y).$$

Following [21], we consider an operator $(L^*, D(L^*))$ defined on a measure space (D,μ) of the form

$$\begin{aligned}L^*f &:= \partial_y\Big[A\partial_y f\Big]-\partial_y\psi A\partial_y f\\ &= \Big[Af_y\Big]_y-\psi_y Af_y\\ &= Af_{yy}+A_y f_y-\psi_y Af_y, \forall f\in\mathcal{A}=L^2(\Omega,\mu)\cap D(L^*).\end{aligned}$$

where

$$\psi_y=\frac{A_y}{A}-\frac{a}{A}. \tag{25}$$

Then, we define the carré du champ operator of L^* by

$$\Gamma(f,g)=\frac{1}{2}\Big(L^*(fg)-fL^*g-gL^*f\Big),\quad \forall f,g\in\mathcal{A} \tag{26}$$

and the iterated carré du champ operator of L^* by

$$\Gamma_2(f,g)=\frac{1}{2}\Big(L^*\Gamma(f,g)-\Gamma(f,L^*g)-\Gamma(g,L^*f)\Big),\quad \forall f,g\in\mathcal{A}. \tag{27}$$

We will also denote $\Gamma(f, f) = \Gamma_1(f)$ and $\Gamma_2(f, f) = \Gamma_2(f)$ for short. We note that for the above operator L^*, we always have

$$\Gamma(f, g) = A f_y g_y.$$

A direct computation then shows that

$$\Gamma_1(h) = \frac{1}{2}\Big[L^*(h^2) - 2hL^*(h)\Big] = Ah_y^2 \tag{28}$$

$$\begin{aligned}\Gamma_2(h) &= \frac{1}{2}\Big[L^*(\Gamma_1(h)) - 2\Gamma(h, L^*(h))\Big] \\ &= \Big(\frac{A_y^2}{4A} - \frac{A_{yy}}{2} + A\psi_{yy} + \frac{\psi_y A_y}{2}\Big)\Gamma_1(h) + (Ah_{yy} + \frac{1}{2}A_y h_y)^2.\end{aligned} \tag{29}$$

From (22), (25) and Hypothesis C, we can estimate directly the first coefficient in (29)

$$\frac{A_y^2}{4A} - \frac{A_{yy}}{2} + A\psi_{yy} + \frac{\psi_y A_y}{2} = \frac{1}{2}bb_{yy} - a_y + \frac{ab_y}{b} \geq \rho > 0.$$

Hence, L satisfies the curvature-dimension condition $CD(\rho, n)$ (see [21]), where $\rho > 0$ and $n = \infty$, i.e. for all $f \in \mathcal{A}$

$$\Gamma_2(f) \geq \rho\Gamma_1(f) + \frac{1}{n}(L^* f)^2, \quad \mu_\infty - a.e. \tag{30}$$

Using similar arguments as in [30], since L^* satisfies $CD(\rho, \infty)$, it follows from [3, Proposition 5.7.1, p. 268] that μ_∞ satisfies the logarithmic Sobolev inequality $LS(\frac{1}{\rho})$, i.e. for all densities f we have

$$Ent_{\mu_\infty}(f) = \int_D f \log f d\mu_\infty \leq \frac{1}{2\rho}\int_D \frac{Af'^2}{f} d\mu_\infty = \frac{1}{2\rho}\int_D \frac{\Gamma(f)}{f} d\mu_\infty.$$

In particular,

$$S_{\mu_\infty}(h(t, x, \cdot)) \leq e^{-\rho t} S_{\mu_\infty}(h(0, x, \cdot)). \quad \square$$

Remark 12. Using the Csiszár–Kullback–Pinsker inequality (see e.g. [8])

$$\|\mu - \nu\|_{TV} \leq \sqrt{\frac{1}{2} D_{\mathrm{KL}}(\mu \parallel \nu)} \tag{31}$$

where

$$\|\mu - \nu\|_{TV} = \sup\{|\mu(O) - \nu(O)| : O \text{ is an event to which probabilities are assigned}\}$$

is the total variation metric (or statistical distance) between two probability measure μ and ν, we combine (19) and (31) to conclude that $P(t, x, dy) = u(t, x, y)dy$ converges to $\mu_\infty(dy) = u_\infty(y)dy$ with exponential rate ρ in the total variation metric acting on the space of probability measures.

On the other hand, if we define the semigroup $S_t\mu(A) := \int_D P_t(A, y)\mu(dy)$ on the space of measures, then

$$\begin{aligned}\|S_t\mu - \mu_\infty\|_{TV} &= \sup_A \Big|\int_D P(t, y, A)\mu(dy) - \mu_\infty(A)\Big| \\ &\leq \sup_A \int_D |P(t, y, A) - \mu_\infty(A)|\mu(dy) \leq \|P(t, y, \cdot) - \mu_\infty\|_{TV}.\end{aligned}$$

Therefore, $S_t\mu \to \mu_\infty$ in total variation norm as $t \to \infty$.

4. Existence of the global random attractor

In this section, we consider the situation in which the Fisher–Lamperti transformation can be applied, meaning that there exists a continuous transformation which transforms (1) into a system with additive noise (see e.g. [26]). We find that under the Hypotheses A–C, there actually exists a global random attractor for the random dynamical system generated by (1). This result is somewhat stronger than the existence of a stationary measure because it implies the convergence in the pathwise sense rather than the convergence in distribution only.

To do that, we first recall the basic notions of random dynamical systems. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. On this probability space we consider a measurable flow θ

$$\theta : \mathbb{R} \times \Omega \to \Omega$$

for which $\theta_t(\cdot) : \Omega \to \Omega$ is $\mathbb{P}$-preserving, i.e. $\mathbb{P}(\theta_t^{-1}(A)) = \mathbb{P}(A)$ for every $A \in \mathcal{F}$, $t \in \mathbb{R}$, and $(\theta_t)_{t\in\mathbb{R}}$ satisfies the group property, i.e. $\theta_{t+s} = \theta_t \circ \theta_s$ for all $t, s \in \mathbb{R}$. A general model for noise is the quadruple $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\in\mathbb{R}})$ which is called a *metric dynamical system*. For system (1), θ can be constructed as the Wiener shift w.r.t. the corresponding probability space of the Wiener process [2].

Definition 13. Let X be a Banach space. Then we define a random dynamical system (RDS) as a measurable mapping

$$\varphi : \mathbb{R}^+ \times \Omega \times X \to X$$

satisfying the cocycle property

$$\begin{aligned} \varphi(t+s, \omega, x) &= \varphi(t, \theta_s\omega, \cdot) \circ \varphi(s, \omega, x) \qquad \text{for all } t,\ s \in \mathbb{R}_+,\ \omega \in \Omega,\ x \in X, \\ \varphi(0, \omega) &= \mathrm{id}_X. \end{aligned}$$

An RDS φ is called *continuous* if each mapping $(t, x) \mapsto \varphi(t, \omega, x)$ is continuous.

Below, we shall need the following Gronwall-type lemma.

Lemma 14. *Given $\varepsilon \in (0, 1]$, let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be continuous. Assume that there exists a continuous nondecreasing super-linear function $k : \mathbb{R}^+ \to \mathbb{R}^+$ with $k(t) = 0$ iff $t = 0$ such that*

$$f(t) \le \varepsilon + \int_0^t k(f(s))ds, \quad \forall t \ge 0.$$

Then

$$f(t) \le \varepsilon G^{-1}(G(1) + t), \quad \forall t \ge 0,$$

where

$$G(t) = \int_0^t \frac{1}{k(s)} ds.$$

Proof. Put $u(t) = \frac{f(t)}{\varepsilon}$. Then it follows from the assumption that

$$\begin{aligned} u(t) &\le 1 + \int_0^t \frac{k(f(s))}{\varepsilon} ds \\ &\le 1 + \int_0^t k(u(s))ds \quad \text{(due to the super-linearity of } k\text{)}. \end{aligned}$$

Put $v(t) = 1 + \int_0^t k(u(s))ds$. It implies that $u(t) \le v(t)$ and $v(t)$ is nondecreasing. Moreover, because k is nondecreasing we have $v'(t) = k(u(t)) \le k(v(t))$ and $k(v(t)) \ge k(v(0)) = k(1) > 0$. Therefore, $G(v(t)) - G(v(0)) = \int_0^t \frac{dv(s)}{k(v(s))} \le t$. Because $G'(t) = \frac{1}{k(t)} > 0$, both G and G^{-1} are increasing. This implies the proof. □

Theorem 15. *Assume that there exists the inverse map $Q^{-1}(x)$ of the indefinite integral $Q(x)$ of $\frac{1}{b(x)}$, i.e. $Q(x) = \int^x \frac{1}{b(y)}dy$ such that both $Q(x)$ and $Q^{-1}(x)$ are continuous on D and $Q(D)$. Assume that there exists a continuous nondecreasing super-linear function $k : \mathbb{R}^+ \to \mathbb{R}^+$ with $k(0) = 0$ such that $c(y) := \frac{a(x)}{b(x)} - \frac{1}{2}b_x(x)$ with $x = Q^{-1}(y)$ satisfies*

$$|c(z) - c(y)| \le k(|z - y|), \ \forall z, y \in D.$$

Then the solution of (1) generates a continuous random dynamical system $\varphi : \mathbb{R}^+ \times \Omega \times D \to D$ defined by $\varphi(t, \omega)X_0 = X_t(\omega, X_0)$.

Proof. Introduce the transformation $Y_t = Q(X_t)$. Then by Ito's formula it follows from (1) that

$$\begin{aligned} dY_t &= \frac{dQ}{dx}(X_t)dX_t + \frac{1}{2}\frac{d^2Q}{dx^2}(X_t)b^2(X_t)dt \\ &= \frac{1}{b(X_t)}\big[a(X_t)dt + b(X_t)dW_t\big] - \frac{1}{2}b_x(X_t)dt \\ &= c(Y_t)dt + dW_t. \end{aligned} \tag{32}$$

It is easy to see that $Y_t(\omega, Y_0)$ is continuous w.r.t. $t \in \mathbb{R}^+$. On the other hand, by the comparison principle in the one dimensional space D, we have $Y_t(\omega, Y_1) \ge Y_t(\omega, Y_0) \ge 0$ for all $t \ge 0$ a.s. if $Y_1 > Y_0 \ge 0$ a.s. Therefore

$$\begin{aligned} Y_t(\omega, Y_1) - Y_t(\omega, Y_0) &= Y_1 - Y_0 + \int_0^t \big[c(Y_s(\omega, Y_1)) - c(Y_s(\omega, Y_0))\big]ds \\ &\le |Y_1 - Y_0| + \int_0^t k(|Y_s(\omega, Y_1) - Y_s(\omega, Y_0)|)ds. \end{aligned}$$

Then by Lemma 14 stated above and noting that $G^{-1}(G(1) + t)$ is bounded uniformly in $t \in [0, T]$, it follows that $Y_t(\omega, Y_0)$ is a continuous function w.r.t. Y_0 uniformly in $t \in [0, T]$. By using the triangle inequality

$$|Y_{t_1}(\omega, Y_1) - Y_t(\omega, Y_0)| \le |Y_{t_1}(\omega, Y_1) - Y_t(\omega, Y_1)| + |Y_t(\omega, Y_1) - Y_t(\omega, Y_0)|,$$

it follows that $Y_t(\omega, Y_0)$ is continuous w.r.t. (t, Y_0). Therefore $X_t(\omega, X_0) = Q^{-1}(Y_t(\omega, Y_0))$ is continuous w.r.t. (t, X_0).

By applying Arnold [2, Chapter 2], we get from the existence and uniqueness of a solution of (1) that the system generates a random dynamical system $\varphi(t, \omega)X_0 := X_t(\omega, X_0)$ which is continuous w.r.t. (t, X_0). Moreover, if $\psi(t, \omega)$ is the RDS generated by the transformed system (32) then $\psi(t, \omega) = Q \circ \varphi(t, \omega) \circ Q^{-1}$. □

Remark 16. In general we can extend the random dynamical system φ to the whole two sided time interval $\mathbb{R}$ instead of $\mathbb{R}^+$ by defining $\varphi(t, \omega) := \varphi(-t, \theta_t\omega)^{-1}$ for any $t \le 0$, see Arnold [2, Chapter 1] for more details.

Definition 17. A global random attractor of an RDS φ is a random compact set A which is invariant, i.e. $\varphi(t,\omega)A(\omega) = A(\theta_t\omega)$ for all $t \geq 0$, and which attracts all the bounded sets in the pullback sense, i.e.,

$$\lim_{t\to\infty} d(\varphi(t,\theta_{-t}\omega)M|A(\omega)) = 0,$$

for all bounded sets M, where $d(M|A) := \sup_{x\in M} d(x, A)$ is the semi-Hausdorff distance.

Remark 18. (i) The random attractor acts as the support of the invariant measure of the generated random dynamical systems. In fact,

$$\lim_{t\to\infty} \varphi(t,\theta_{-t}\omega)\mu_\infty = \delta_{a(\omega)}, \tag{33}$$

where δ_a is a Dirac measure and the limit in (33) is in the sense of weak topology. We recommend Crauel [6] and Arnold [2, Chapter 1, Proposition 1.8.4, pp. 42] for a presentation of the relation between invariant measures of the random dynamical system and the stationary distribution of the corresponding Markov semigroup.

(ii) Under the assumptions of Theorems 11 and 15, a direct computation shows that

$$\begin{aligned}
\frac{dc(y)}{dy} &= \frac{d}{dx}\Big(\frac{a(x)}{b(x)} - \frac{1}{2}b_x(x)\Big)\frac{dx}{dy} \\
&= \Big(\frac{a_x(x)b(x) - a(x)b_x(x)}{b^2(x)} - \frac{1}{2}b_{xx}(x)\Big)\frac{dx}{dQ(x)} \\
&= \Big(-\frac{1}{2}b_{xx}(x) + \frac{a_x(x)}{b(x)} - \frac{a(x)b_x(x)}{b^2(x)}\Big)b(x) \\
&= -\Big(\frac{1}{2}b(x)b_{xx}(x) - a_x(x) + \frac{a(x)b_x(x)}{b(x)}\Big) \leq -\rho,
\end{aligned}$$

where the last inequality is due to Hypothesis C. Hence,

$$[c(y_1) - c(y_2)](y_1 - y_2) \leq \max_{z\in D} c'(z)(y_1 - y_2)^2 \leq -\rho(y_1 - y_2)^2,$$

which proves the dissipativity of the coefficient function c. Therefore, by applying similar arguments as in [20], we can prove that there exists a global random attractor $\bar{A}$ for the RDS ψ generated by the transformed system (32), i.e. for any $x_0 \in D$

$$\lim_{t\to\infty} Q \circ \varphi(t,\theta_{-t}\omega) \circ Q^{-1}(x_0) = \bar{A}(\omega).$$

However, since Q and Q^{-1} are not assumed to have bounded derivatives, in general it does not follow that $A(\omega) := Q^{-1}(\bar{A}(\omega))$ is the global random attractor of φ, i.e. we do not have

$$\lim_{t\to\infty} \varphi(t,\theta_{-t}\omega) \circ Q^{-1}(x_0) = Q^{-1}(\bar{A}(\omega)).$$

Below we will present several examples in which there does exist a random attractor for the generated RDS.

(iii) We note that in general for higher dimensional systems, the relation between the dissipative constant of the transformed system (32) and the Bakry–Émery (BE) curvature dimension is still an open question.

Example 19 (*Cox–Ingersoll–Ross Interest Rate Model*)**.** We consider the Cox–Ingersoll–Ross [5] (CIR) model for the short term interest rate

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t,\ X_t \geq 0, \tag{34}$$

where $k, \theta, \sigma > 0$ such that $k\theta \geq \frac{\sigma^2}{2}$. Feller [16] proved that such a process is nonnegative. The Markov semigroup is proved to be Feller by Duffie et al. [14]. Given X_0, it is well known that $\frac{4k}{\sigma^2(1-e^{-kt})}X_t$ follows a noncentral χ^2 distribution with degree of freedom $\frac{4k\theta}{\sigma^2}$ and non-centrality parameter $\frac{4k}{\sigma^2(1-e^{-kt})}X_0e^{-kt}$, i.e. the probability density function has the form

$$f(x; x_0) = \frac{2k}{\sigma^2(1-e^{-kt})}e^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}}I_q(2\sqrt{uv}),$$

with $q = \frac{2k\theta}{\sigma^2} - 1$, $u = \frac{2k}{\sigma^2(1-e^{-kt})}x_0e^{-kt}$, $v = \frac{2k}{\sigma^2(1-e^{-kt})}x$ and $I_q(2\sqrt{uv})$ is the modified Bessel function of the first kind of order q. Moreover, $\lim_{t\to\infty}\frac{1}{t}\int_0^t X_s ds = X_\infty$ by the law of large numbers, where X_∞ follows a Gamma distribution with shape parameter $\frac{2k\theta}{\sigma^2}$ and scale parameter $\frac{\sigma^2}{2k}$, i.e. the density function has the form

$$f_\infty(x) = \frac{2k}{\sigma^2\Gamma(\nu)}x^{\nu-1}e^{-\frac{2k}{\sigma^2}x}$$

with $\nu = \frac{2k\theta}{\sigma^2}$ (see e.g. Deelstra and Delbaen [12,13] and Deelstra [10]).

In this case $b(y) = \sigma\sqrt{y}$ which is $\frac{1}{2}$-Hölder continuous in the domain $D := [0, \infty)$, therefore the existence and uniqueness of a solution of (34) can be derived from [11]. Note that the condition (14) in the CIR case is $c(x) \equiv k > 0$ so $X_t > 0$ for all $t \geq 0$ with probability one.

The potential energy function can be explicitly computed as

$$\begin{aligned}\psi(y) &= \log\Big[\frac{1}{2}b^2(y)\Big] - \int^y \frac{2a(z)}{b^2(z)}dz\\ &= \log\Big[\frac{\sigma^2}{2}y\Big] - \int^y \frac{2k(\theta - z)}{\sigma^2 z}dz\\ &= \log\Big[\frac{\sigma^2}{2}\Big] + \Big(1 - \frac{2k\theta}{\sigma^2}\Big)\log y + \frac{2k}{\sigma^2}y.\end{aligned}$$

Hence a direct computation to check conditions (5), (7) and (18) gives us

$$\begin{aligned}\text{i, } &\int_0^\infty e^{-\psi(y)}dy = \frac{2}{\sigma^2}\int_0^\infty y^{(\frac{2k\theta}{\sigma^2}-1)}e^{-\frac{2k}{\sigma^2}y}dy = \frac{2}{\sigma^2}\Big(\frac{2k}{\sigma^2}\Big)^{-\frac{2k\theta}{\sigma^2}}\int_0^\infty z^{(\frac{2k\theta}{\sigma^2}-1)}e^{-z}dz\\ &\qquad = \frac{2}{\sigma^2}\Big(\frac{2k}{\sigma^2}\Big)^{-\frac{2k\theta}{\sigma^2}}\Gamma\Big(\frac{2k\theta}{\sigma^2}\Big),\\ \text{ii, } &A(y)e^{-\psi(y)}|_{\partial D} = y^{\frac{2k\theta}{\sigma^2}}e^{-\frac{2k}{\sigma^2}y}|_{\{0,\infty\}} = 0,\\ \text{iii, } &\frac{1}{2}bb_{yy} - a_y + \frac{ab_y}{b} = \frac{k}{2} + \frac{1}{2y}\Big(k\theta - \frac{\sigma^2}{4}\Big) \geq \frac{k}{2},\end{aligned}$$

where $\Gamma(\cdot)$ is the Gamma function. Hence there exists a unique stationary measure for the corresponding Markov semigroup of system (34), with rate of convergence $\frac{k}{2}$.

We can also prove that there exists a global random attractor. Indeed, the transformation $Y_t = Q(X_t) = \frac{2\sqrt{X_t}}{\sigma}$ gives us

$$dY_t = c(Y_t)dt + dW_t = \Big[\frac{4k\theta - \sigma^2}{2\sigma^2}\frac{1}{Y_t} - \frac{k}{2}Y_t\Big]dt + dW_t. \tag{35}$$

By the comparison principle in one dimensional space D, we have $Y_t(\omega, Y_1) \geq Y_t(\omega, Y_0) \geq 0$ for all $t \geq 0$ a.s. if $Y_1 \geq Y_0$ a.s. Therefore

$$Y_t(\omega, Y_1) - Y_t(\omega, Y_0) \leq Y_1 - Y_0 - \frac{k}{2}\int_0^t \Big[Y_s(\omega, Y_1) - Y_s(\omega, Y_0)\Big]ds,$$

thus

$$\Big|Y_t(\omega, Y_1) - Y_t(\omega, Y_0)\Big| \leq |Y_1 - Y_0|e^{-\frac{k}{2}t}. \tag{36}$$

Also, it is easy to check that the drift coefficient c in (35) satisfies the dissipativity condition, i.e.

$$(y_1 - y_2)(c(y_1) - c(y_2)) = \Big[-\frac{4k\theta - \sigma^2}{2\sigma^2}\frac{1}{y_1 y_2} - \frac{k}{2}\Big](y_1 - y_2)^2 \leq -\frac{k}{2}(y_1 - y_2)^2.$$

Therefore, as seen in Garrido-Atienza et al. [20], there exists a pullback random attractor (and also a forward random attractor because of (36)) $Y^*(\omega)$ which is invariant under the random dynamical system generated by (35). We are going to prove that $X^*(\omega) = \frac{\sigma^2}{4}Y^*(\omega)^2$ is also a random attractor of the random dynamical system generated by (34).

Indeed, taking the expectation in both sides of (34) gives

$$dEX_t = k(\theta - EX_t)dt,$$

thus

$$\frac{\sigma^2}{4}EY_t^2 = EX_t = \theta + [EX_0 - \theta]e^{-kt}.$$

By the Borel–Cantelli lemma, it is easy to prove that for all $Y_0 \geq 0$

$$\lim_{t\to\infty} e^{-\frac{k}{2}t}Y_t(\omega, Y_0) = 0 \text{ a.s.},$$

and in particular $e^{-\frac{k}{2}t}Y^*(\theta_t\omega) \to 0$ as $t \to \infty$ a.s. Hence

$$\begin{aligned}|X_t - X^*(\theta_t\omega)| &= \frac{\sigma^2}{4}\Big|Y_t - Y^*(\theta_t\omega)\Big|\Big|(Y_t - Y^*(\theta_t\omega)) + 2Y^*(\theta_t\omega)\Big| \\ &\leq \frac{\sigma^2}{4}|Y_0 - Y^*(\omega)|^2e^{-kt} + \frac{\sigma^2}{2}|Y_0 - Y^*(\omega)|e^{-\frac{k}{2}t}Y^*(\theta_t\omega),\end{aligned}$$

which tends to zero as $t \to \infty$, proving the attraction of the random attractor $X^*(\omega)$.

Example 20 (*Wright–Fisher Model*)**.** We consider the one dimensional Wright–Fisher model [30]

$$dX_t = [\theta_1 - (\theta_1 + \theta_2)X_t]dt + \sqrt{X_t(1 - X_t)}dW_t,\ X_t \in [0, 1], \tag{37}$$

where $\theta_1, \theta_2 > \frac{1}{2}$. For the existence and uniqueness of a solution of (37), see [30], in this case $b(y) = \sqrt{y(1-y)}$ is $\frac{1}{2}$-Hölder continuous in the domain $D = [0, 1]$. The same method to prove the existence and uniqueness of a stationary distribution is mentioned and proved in [30]. For this system, the condition (14) becomes $c(x) \equiv \theta_1 + \theta_2 - 1 > 0$, which is sufficient to conclude that the solution never reaches the boundary with probability one.

The potential energy function can be explicitly computed as

$$\begin{aligned}\psi(y) &= \log\Big[\frac{1}{2}b^2(y)\Big] - \int^y \frac{2a(z)}{b^2(z)}dz \\ &= \log\Big[\frac{y(1-y)}{2}\Big] - \int^y \frac{2(\theta_1 - (\theta_1+\theta_2)z)}{z(1-z)}dz \\ &= \log\Big[\frac{1}{2}y^{1-2\theta_1}(1-y)^{1-2\theta_2}\Big].\end{aligned}$$

A direct computation to check conditions (5), (7) and (18) gives us

$$\begin{aligned}&\text{i,}\quad \int_0^1 e^{-\psi(y)}dy = 2\int_0^1 y^{2\theta_1-1}(1-y)^{2\theta_2-1}dy < \infty,\\ &\text{ii,}\quad A(y)e^{-\psi(y)}|_{\partial D} = y^{2\theta_1}(1-y)^{2\theta_2}|_{\{0,1\}} = 0,\\ &\text{iii,}\quad \frac{1}{2}bb_{yy} - a_y + \frac{ab_y}{b} = \frac{(\theta_1 - \frac{1}{4})(1-y) + (\theta_2 - \frac{1}{4})y}{2y(1-y)} \geq 2\sqrt{(\theta_1 - \frac{1}{4})(\theta_2 - \frac{1}{4})}\\ &\qquad\qquad =: \rho > 0.\end{aligned}$$

Hence there exists a unique stationary measure for the corresponding Markov semigroup of system (37), with rate of convergence ρ.

To prove the existence of a global random attractor, observe that the transformation $Y_t = \arcsin 2\sqrt{X_t(1-X_t)}$ gives us

$$\begin{aligned}dY_t &= \frac{\alpha_1 - (\alpha_1+\alpha_2)X_t}{\sqrt{X_t(1-X_t)}}dt + dW_t = \frac{\alpha_1 - (\alpha_1+\alpha_2)\sin^2\frac{Y_t}{2}}{\frac{1}{2}\sin Y_t}dt + dW_t\\ &= c(Y_t)dt + dW_t,\end{aligned} \tag{38}$$

where $\alpha_i = \theta_i - \frac{1}{4}$. By the comparison principle, it follows that for any $Y_1 \leq Y_2$ then $Y_t(\omega, Y_1) \leq Y_t(\omega, Y_2)$ a.s. Since cot is a decreasing function on $[0, \frac{\pi}{2}]$, it is then easy to check that

$$\begin{aligned}[c(y_2) - c(y_1)](y_2 - y_1) &= \Big[\alpha_1 \cot\frac{y_2}{2} - \alpha_2\tan\frac{y_2}{2} - \alpha_1\cot\frac{y_1}{2} + \alpha_2\tan\frac{y_1}{2}\Big](y_2 - y_1)\\ &\leq -\Big[\tan\frac{y_2}{2} - \tan\frac{y_1}{2}\Big](y_2 - y_1)\\ &\leq -(y_2 - y_1)^2,\end{aligned}$$

i.e. the drift coefficient c in (38) satisfies the dissipativity condition. Again, by [20], it follows that there exists a random attractor (both pullback and forward) $Y^*(\omega)$ for the random dynamical system generated by (38). Moreover,

$$\begin{aligned}d(Y_t(\omega, Y_2) - Y_t(\omega, Y_1)) &= \Big[c(Y_t(\omega, Y_2)) - c(Y_t(\omega, Y_1))\Big]dt\\ &\leq -\Big[\tan\frac{Y_t(\omega, Y_2)}{2} - \tan\frac{Y_t(\omega, Y_1)}{2}\Big]dt\\ &\leq -\frac{1}{2}\Big[Y_t(\omega, Y_2) - Y_t(\omega, Y_1)\Big]dt,\end{aligned}$$

which results in

$$|Y_t(\omega, Y_2) - Y_t(\omega, Y_1)| \leq |Y_2 - Y_1|e^{-\frac{1}{2}t}.$$

Then similar to system (34) there also exists a random attractor $X^*(\omega) = \sin^2 \frac{Y^*(\omega)}{2}$ of the random dynamical system generated by (37), due to the fact that

$$\begin{aligned} |X_t(\omega, X_2) - X_t(\omega, X_1)| &= \Big|\frac{\cos Y_t(\omega, Y_2) - \cos Y_t(\omega, Y_1)}{2}\Big| \\ &= \Big|\sin\frac{[Y_t(\omega, Y_2) - Y_t(\omega, Y_1)]}{2}\Big|\Big|\sin\frac{[Y_t(\omega, Y_2) + Y_t(\omega, Y_1)]}{2}\Big| \\ &\le \Big|\frac{[Y_t(\omega, Y_2) - Y_t(\omega, Y_1)]}{2}\Big|. \end{aligned}$$

Example 21 (*Ait-Sahalia Interest Rate Model*)**.** We now consider the Ait-Sahalia-type model for the stochastic interest rate (see Ait-Sahalia [1] and Szpruch et al. [28])

$$dx(t) = \Big[k(\theta - x(t)) - \alpha x(t)^r + \frac{\beta}{x(t)}\Big]dt + \sigma x(t)^\rho dW(t), \tag{39}$$

where $k, \theta, \alpha, \beta, \sigma > 0$ and $r, \rho > 1$. System (39) does not satisfy Hypothesis A since it contains a singular value 0 in the drift term. However, the coefficient functions of (39) are locally Lipschitz continuous in $(0, \infty)$, thus by using the stopping time technique, Szpruch et al. [28] prove that there exists a unique solution of system (39) which is positive for all $t > 0$ given that $x_0 > 0$. Moreover if $r + 1 > 2\rho$ then the Assumption 2.2 in [28] is satisfied, thus

$$\sup_{0\le t<\infty} E|x|^q < \infty, \qquad E(\sup_{0\le t\le T}|x|^q) < \infty, \qquad \text{for all } q \ge 2.$$

A direct computation shows that

$$\begin{aligned} \int_0^\infty e^{-\psi(y)}dy &= \int_0^\infty \frac{2}{\sigma^2 y^{2\rho}}\exp\Big[\frac{2k\theta y^{1-2\rho}}{\sigma^2(1-2\rho)} + \frac{ky^{2(1-\rho)}}{\sigma^2(\rho-1)} \\ &\quad - \frac{2\alpha y^{r+1-2\rho}}{\sigma^2(r+1-2\rho)} - \frac{\beta}{\sigma^2\rho}y^{-2\rho}\Big]dy \\ &= \frac{2}{\sigma^2}\int_0^\epsilon y^{-2\rho}\exp\Big[y^{-2\rho}\Big(-\frac{\beta}{\sigma^2\rho} + \frac{2k\theta y}{\sigma^2(1-2\rho)} \\ &\quad + \frac{k\theta y^2}{\sigma^2(\rho-1)} - \frac{2\alpha y^{r+1}}{\sigma^2(r+1-2\rho)}\Big)\Big]dy \\ &\quad + \frac{2}{\sigma^2}\int_\epsilon^{\frac{1}{\epsilon}} y^{-2\rho}\exp\Big[\frac{2k\theta y^{1-2\rho}}{\sigma^2(1-2\rho)} + \frac{ky^{2(1-\rho)}}{\sigma^2(\rho-1)} - \frac{2\alpha y^{r+1-2\rho}}{\sigma^2(r+1-2\rho)} \\ &\quad - \frac{\beta}{\sigma^2\rho}y^{-2\rho}\Big]dy \\ &\quad + \frac{2}{\sigma^2}\int_{\frac{1}{\epsilon}}^\infty y^{-2\rho}\exp\Big[y^{r+1-2\rho}\Big(-\frac{2\alpha}{\sigma^2(r+1-2\rho)} - \frac{\beta y^{-1-r}}{\sigma^2\rho} \\ &\quad + \frac{k\theta y^{1-r}}{\sigma^2(\rho-1)} + \frac{2k\theta y^{-r}}{\sigma^2(1-2\rho)}\Big)\Big]dy. \end{aligned} \tag{40}$$

Since

$$\begin{aligned} &\lim_{y\to 0}\exp\Big[y^{-2\rho}\Big(-\frac{\beta}{\sigma^2\rho} + \frac{2k\theta y}{\sigma^2(1-2\rho)} + \frac{k\theta y^2}{\sigma^2(\rho-1)} - \frac{2\alpha y^{r+1}}{\sigma^2(r+1-2\rho)}\Big)\Big] \\ &\quad = \exp[-\infty] = 0, \end{aligned}$$

$$\lim_{y\to\infty} \exp\Big[y^{r+1-2\rho}\Big(-\frac{2\alpha}{\sigma^2(r+1-2\rho)} - \frac{\beta y^{-1-r}}{\sigma^2\rho} + \frac{k\theta y^{1-r}}{\sigma^2(\rho-1)} + \frac{2k\theta y^{-r}}{\sigma^2(1-2\rho)}\Big)\Big] = \exp[-\infty] = 0,$$

condition (7) is satisfied. Moreover, by choosing ϵ small enough, we can estimate the integrals in (40) as

$$\begin{aligned}
\int_0^\infty e^{-\psi(y)}dy &\le \frac{2}{\sigma^2}\int_0^\epsilon y^{-2\rho}\exp\Big[-\frac{\beta}{2\sigma^2\rho}y^{-2\rho}\Big]dy + \frac{2}{\sigma^2}\int_\epsilon^{\frac{1}{\epsilon}} \epsilon^{-2\rho}\exp\Big[\frac{k\epsilon^{2(1-\rho)}}{\sigma^2(\rho-1)}\Big]dy \\
&\quad + \frac{2}{\sigma^2}\int_{\frac{1}{\epsilon}}^\infty y^{-2\rho}\exp\Big[-\frac{\alpha}{\sigma^2(r+1-2\rho)}y^{r+1-2\rho}\Big]dy \\
&\le \frac{2}{\sigma^2}\int_0^\infty \frac{\sigma^2}{\beta}\Big(\frac{2\sigma^2\rho}{\beta}\Big)^{\frac{1}{2\rho}} z^{-\frac{1}{2\rho}}e^{-z}dz + \frac{2}{\sigma^2}\Big[\frac{1}{\epsilon}-\epsilon\Big]\epsilon^{-2\rho}\exp\Big[\frac{k\epsilon^{2(1-\rho)}}{\sigma^2(\rho-1)}\Big] \\
&\quad + \frac{2}{\sigma^2}\int_0^\infty \frac{\sigma^2}{\alpha}\Big(\frac{\sigma^2(r+1-2\rho)}{\alpha}\Big)^{-\frac{r}{r+1-2\rho}} z^{-\frac{r}{r+1-2\rho}}e^{-z}dz \\
&\le \frac{2}{\sigma^2}\frac{\sigma^2}{\beta}\Big(\frac{2\sigma^2\rho}{\beta}\Big)^{\frac{1}{2\rho}}\Gamma(1-\frac{1}{2\rho}) + \frac{2}{\sigma^2}\epsilon^{-1-2\rho}\exp\Big[\frac{k\epsilon^{2(1-\rho)}}{\sigma^2(\rho-1)}\Big] \\
&\quad + \frac{2}{\sigma^2}\frac{\sigma^2}{\alpha}\Big(\frac{\sigma^2(r+1-2\rho)}{\alpha}\Big)^{-\frac{r}{r+1-2\rho}}\Gamma(\frac{1-2\rho}{r+1-2\rho}),
\end{aligned}$$

which is finite, thus condition (5) is also satisfied. Therefore there exists a unique stationary measure μ_∞ for the corresponding Markov semigroup.

To check condition (18), observe that

$$\begin{aligned}
\frac{1}{2}bb_{yy} - a_y + \frac{ab_y}{b} &= k(1-\rho) + \frac{\beta(1+\rho)}{y^2} + \frac{k\theta\rho}{y} + \frac{\sigma^2\rho(\rho-1)}{2}y^{2\rho-2} + \alpha(r-\rho)y^{r-1} \\
&> k(1-\rho) + \beta(\rho-1)y^{-2} + k\theta(\rho-1)y^{-1} \\
&\quad + \frac{\sigma^2\rho(\rho-1)}{2}y^{2\rho-2} + \alpha(\rho-1)y^{r-1} \\
&= (\rho-1)\Big[-k + \beta y^{-2} + k\theta y^{-1} + \frac{\sigma^2\rho}{2}y^{2(\rho-1)} + \alpha y^{r-1}\Big] \\
&> (\rho-1)\Big[-k + k\theta y^{-1} + \frac{\sigma^2\rho}{2}y^{2(\rho-1)}\Big] = (\rho-1)h(y),
\end{aligned}$$

where $h : \mathbb{R}_+ \to \mathbb{R}$ is a convex function which has a minimum at $y_* = \Big[\frac{k\theta}{\rho(\rho-1)\sigma^2}\Big]^{\frac{1}{2\rho-1}}$. Therefore, the sufficient condition for an exponential rate of convergence to the stationary measure is

$$h(y_*) > 0 \Leftrightarrow y_* = \Big[\frac{k\theta}{\rho(\rho-1)\sigma^2}\Big]^{\frac{1}{2\rho-1}} < \frac{\theta(2\rho-1)}{2(\rho-1)}. \tag{41}$$

The sufficient condition (41) is independent of α, β. On the other hand, the Fisher–Lamperti transformation can be written explicitly as $Q(x) = \frac{1}{\sigma}x^{1-\rho}$, therefore we can also prove that there exists a random attractor for the random dynamical system generated by (39). The proof is similar as above and will be omitted here.

Acknowledgments

We would like to thank the anonymous referees for their careful reading and insightful remarks which lead to improvement of our manuscript. This work was carried out at the Max Planck Institute for Mathematics in the Sciences (MPI MIS Leipzig). L.H.D. and T.D.T. are partially supported by Vietnam Institute for advanced studies in mathematics (VIASM). L.H.D. is also partially supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 101.03-2014.42.

References

[1] Y. Ait-Sahalia, Testing continuous time models of the spot interest rate, Rev. Financ. Stud. 9 (2) (1996) 385–426.
[2] L. Arnold, Random Dynamical Systems, Springer, Berlin Heidelberg New York, 1998.
[3] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, in: Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 38, Springer, Cham, 2014.
[4] L. Bertini, L. Passalacqua, Modelling interest rates by correlated multi-factor CIR-like processes, 2008. Available at SSRN 1175702.
[5] J.C. Cox, J.E. Ingersoll, S.A. Ross, A theory of the term structure of interest rates, Econometrica 53 (1985) 385–407.
[6] H. Crauel, Measure attractors and Markov attractors, Stoch. Dyn. 23 (1) (2008) 75–107.
[7] H. Crauel, P. Imkeller, M. Steinkamp, Bifurcations of one-dimensional stochastic differential equations, Stoch. Dynam. (1999) 27–47.
[8] I. Csiszár, J. Körner, Information Theory: Coding Theorem for Discrete Memoryless Systems, Cambridge University Press, 2011.
[9] G. Da Prato, J. Zabczyk, Ergodicity for infinite dimensional systems, Vol. 229, Cambridge University Press, 1996.
[10] G. Deelstra, Long term returns in stochastic interest models: applications, Astin Bull. 30 (1) (2000) 123–140.
[11] G. Deelstra, F. Delbaen, Existence of solutons of stochastic differential equations related to the Bessel process, Dept. Mathematics, ETH, Zürich, Switzerland, 1994, preprint.
[12] G. Deelstra, F. Delbaen, Long term returns in stochastic interest rate models, Insur. Math. Econ. 17 (1995) 163–169.
[13] G. Deelstra, F. Delbaen, Long term returns in stochastic interest rate models: Convergence in law, Stoch. Stoch. Rep. 55 (1995) 253–277.
[14] D. Duffie, D. Filipovic, W. Schachermayer, Affine processes and applications in finance, Ann. Appl. Probab. 13 (3) (2003) 984–1053.
[15] W.J. Ewens, Mathematical Population Genetics. I, Springer-Verlag, New York, 2004.
[16] W. Feller, Two singular diffusion problems, Ann. of Math. 54 (1951) 173–182.
[17] W. Feller, Diffusion processes in one dimension, Trans. Amer. Math. Soc. 77 (1954) 1–31.
[18] M. Fukushima, Dirichlet Forms and Markov Process, Vol. 9, North Holland Publishing, Amsterdam, 1980, pp. 107–123.
[19] M. Fukushima, D. Stroock, Reversibility of solutions to martingale problems, Adv. Math. Suppl. Stud. 9 (1986) 107–123.
[20] M. Garrido-Atienza, P. Kloeden, A. Neuenkirch, Discretization of stationary solution of stochastic systems driven by fractional brownian motion, Appl. Math. Optim. 69 (2009) 151–172.
[21] L. Gross, Logarithmic Sobolev inequalities, Amer. J. Math. 97 (4) (1975) 1061–1083.
[22] M. Hairer, Convergence of Markov Processes, in: Lecture Notes, 2016.
[23] T. Harris, The existence of stationary measures for certain markov pr ocesses, in: Proceedings of the Third Berkeley Symposium on Mathematica l Statistics and Probability (1954–1955), vol. II, University of California Press, Berkeley and LosAngeles, 1956, pp. 113–124.
[24] N. Ikeda, S. Watanabe, Stochastic differential equations and diffusion processes, Amsterdam-Oxford-New York, 1981.
[25] R. Jordan, D. Kinderlehrer, F. Otto, Free energy and the Fokker–Planck equation, in: Landscape Paradigms in Physics and Biology, Los Alamos, NM, 1996, Physica D 2–4 (1997) 265–271.
[26] A. Neuenkirch, L. Szpruch, First order strong approximations of scalar SDEs defined in a domain, Numer. Math. 128 (1) (2014) 103–136.
[27] B. Schmalfuß, Measure attractors and random attractors for stochastic partial differential equations, Stoch. Anal. Appl. 17 (6) (1999) 1075–1101.

[28] L. Szpruch, X. Mao, D.J. Higham, J. Pan, Numerical simulation of a strongly nonlinear Ait-Sahalia type interest rate model, BIT 51 (2) (2011) 405–425.
[29] T.D. Tran, J. Hofricher, J. Jost, The free energy method and the Wright-Fisher model with 2 alleles, Theory Biosci. 134 (2015) 83–92.
[30] T.D. Tran, J. Hofrichter, J. Jost, The free energy method for the Fokker–Planck equation of the Wright-Fisher model, In Review.
[31] T. Yamada, S. Watanabe, On the uniqueness of stochastic differential equations, J. Math. Kyoto Uni. 11 (1971) 155–167.
[32] T. Yamada, S. Watanabe, On the uniqueness of stochastic differential equations II, J. Math. Kyoto Uni. 11 (1971) 553–563.