ELSEVIER
Bull. Sci. math. 134 (2010) 391–404
BULLETIN DES SCIENCES MATHÉMATIQUES
www.elsevier.com/locate/bulsci

Exponential ergodicity of non-Lipschitz multivalued stochastic differential equations ☆

Jiagang Ren [a,*], Jing Wu [a], Xicheng Zhang [b,c]

[a] *Zhongshan University, Guangzhou, Guangdong 510275, PR China*
[b] *The University of New South Wales, Sydney 2052, Australia*
[c] *Huazhong University of Science and Technology, Wuhan, Hubei 430074, PR China*

Received 15 December 2008

Available online 10 July 2009

Abstract

Under the conditions of coefficients being non-Lipschitz and the diffusion coefficient being elliptic, we study the strong Feller property and irreducibility for the transition probability of solutions to general multivalued stochastic differential equations by using the coupling method, Girsanov's theorem and a stopping argument. Thus we can establish the exponential ergodicity and the spectral gap.

Keywords: Non-Lipschitz multivalued stochastic differential equation; Ellipticity; Girsanov's theorem; Stopping time; Strong Feller property; Irreducibility; Ergodicity

1. Introduction

Recently in [9] the third named author proved the exponential ergodicity of the solution of the following stochastic differential equation with non-Lipschitz coefficients:

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0, \tag{1}$$

where $b:\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma:\mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^n$ are continuous (but not necessarily Lipschitz continuous) functions, $(W_t)_{t\geqslant 0}$ is an n-dimensional standard Brownian motion defined on some complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The idea therein was to use the drift transform together

☆ Research supported by NSFC (Grant No. 10871215).
* Corresponding author.
E-mail addresses: renjg@mail.sysu.edu.cn (J. Ren), wjjosie@hotmail.com (J. Wu), XichengZhang@gmail.com (X. Zhang).

0007-4497/$ – see front matter
doi:10.1016/j.bulsci.2009.01.003

with an appropriate coupling function to deduce necessary estimates. To carry out this idea it was assumed that the diffusion coefficient, σ, was a square matrix and is uniformly elliptic.

On the other hand, Cépa and Jacquot in [3] proved the ergodicity for the solution of the following stochastic variational inequality (SVI in short):

$$\mathrm{d}X_t + \partial\varphi(X_t) \ni b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0 \in \overline{\mathrm{Dom}(\varphi)}, \tag{2}$$

where $\partial\varphi$ is the sub-differential of some convex function φ with $\mathrm{Dom}(\varphi) = \{x\colon \varphi(x) < \infty\}$. The main tool therein is the Bismut formula and it is also assumed, besides the C_b^2 regularity of σ and b, that σ is a square matrix and uniformly elliptic.

Consequently, a common drawback of the two papers mentioned above is the *uniform ellipticity assumption* of the diffusion coefficients. The purpose of the present paper is to remove this assumption and instead we assume only the *ellipticity*. Our main result is stated in Theorem 3.1 below which unifies and improves the main results in both [9] and [3]. At the same time, it turns out that we do not need to assume the diffusion matrix is square any more and that our method even works for general multivalued stochastic differential equations (MSDEs in abbreviation):

$$\mathrm{d}X_t + A(X_t) \ni b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \qquad X_0 = x_0 \in \overline{D(A)}, \tag{3}$$

where A is a multivalued maximal monotone operator on $\mathbb{R}^d$ with $\mathrm{Int}(D(A)) \neq \emptyset$. It is well known (see [1]) that these equations include stochastic variational inequalities as a special case. It is worthwhile mentioning here that SVIs generalize reflected stochastic differential equations in a convex domain.

2. Preliminaries

We begin with some notions and notations. Let $\{X_t(x),\ t \geqslant 0,\ x \in \mathbb{E}\}$ be a family of Markov processes with state space $\mathbb{E}$ being a Hausdorff topology space, and transition probability $P_t(x, E)$. Then

(i) P_t is called strong Feller if for each $t > 0$ and $E \in \mathscr{B}(\mathbb{E})$

$$\mathbb{E} \ni x \mapsto P_t(x, E) \in [0, 1] \text{ is continuous;}$$

(ii) P_t is called irreducible if for each $t > 0$ and $x \in \mathbb{E}$

$$P_t(x, E) > 0 \quad \text{for any non-empty open set } E \subset \mathbb{E}.$$

(iii) A measure μ on $(\mathbb{E}, \mathscr{B}(\mathbb{E}))$ is an invariant measure for P_t if

$$\int_{\mathbb{E}} P_t(x, E)\,\mu(\mathrm{d}x) = \mu(E), \qquad \forall t > 0,\ E \in \mathscr{B}(\mathbb{E}).$$

The transition probability $P_t(x, \cdot)$ determines a Markov semigroup $(P_t)_{t \geqslant 0}$. The theorem below is a classical result combining the concepts above (cf. [4]).

Theorem 2.1. *Assume a Markov semigroup $(P_t)_{t\geqslant 0}$ is strong Feller and irreducible. Then there exists at most one invariant measure for it. Moreover, if μ is the invariant measure, then μ is ergodic and equivalent to each $P_t(x, \cdot)$ and that as $t \to \infty$, $P_t(x, B) \to \mu(B)$ for any arbitrary Borel set B.*

In this paper the semigroup $(P_t)_{t\geqslant 0}$ is defined as

$$P_t f(x_0) := \mathbf{E} f\big(X_t(x_0)\big), \quad t > 0,\ f \in B_b\big(\mathbb{R}^d\big)$$

and the transition function is

$$P_t(x_0, E) := \mathbf{P}\big(X_t(x_0) \in E\big)$$

where $X_t(x_0)$ is the solution to (3) and $B_b(\mathbb{R}^d)$ denotes the set of all bounded measurable functions on $\mathbb{R}^d$.

We collect here definitions and some useful properties about the maximal monotone operator and solutions to MSDEs. For more details, we refer to Cépa [1].

Definition 2.2. Given a multivalued operator A from $\mathbb{R}^d$ to $2^{\mathbb{R}^d}$, define:

$$D(A) := \big\{x \in \mathbb{R}^d\colon\ A(x) \neq \emptyset\big\},$$
$$\mathrm{Gr}(A) := \big\{(x, y) \in \mathbb{R}^{2d}\colon\ x \in \mathbb{R}^d,\ y \in A(x)\big\}.$$

(1) A multivalued operator A is called monotone if

$$\langle y_1 - y_2, x_1 - x_2\rangle \geqslant 0, \quad \forall (x_1, y_1), (x_2, y_2) \in \mathrm{Gr}(A).$$

(2) A monotone operator A is called maximal monotone if and only if

$$(x_1, y_1) \in \mathrm{Gr}(A) \quad \Leftrightarrow \quad \langle y_1 - y_2, x_1 - x_2\rangle \geqslant 0, \quad \forall (x_2, y_2) \in \mathrm{Gr}(A).$$

Proposition 2.3. *Let A be a maximal monotone operator on $\mathbb{R}^d$, then*:

(i) $\mathrm{Int}(D(A))$ *and* $\overline{D(A)}$ *are convex subsets of* $\mathbb{R}^d$ *and* $\mathrm{Int}(D(A)) = \mathrm{Int}(\overline{D(A)})$.

(ii) *For each $x \in D(A)$, $A(x)$ is a closed and convex subset of $\mathbb{R}^d$. Let $A^\circ(x) := \mathrm{proj}_{A(x)}(0)$ be the minimal section of A, where proj_D is designated as the projection on every closed and convex subset D on $\mathbb{R}^d$ and $\mathrm{proj}_\emptyset(0) = \infty$. Then*

$$x \in D(A) \quad \Leftrightarrow \quad \big|A^\circ(x)\big| < +\infty.$$

(iii) *The resolvent operator $J_n := (1 + \frac{1}{n}A)^{-1}$ is a single-valued and contractive operator defined on $\mathbb{R}^d$ and valued in $D(A)$.*

(iv) *The Yosida approximation $A_n := n(I - J_n)$ is a single-valued, maximal monotone and Lipschitz continuous function with Lipschitz constant n. Moreover, for every $x \in D(A)$, as $n \nearrow \infty$,*

$$A_n(x) \to A^\circ(x)$$

and

$$\big|A_n(x)\big| \nearrow \big|A^\circ(x)\big| \quad \text{if } x \in D(A),$$
$$\big|A_n(x)\big| \nearrow +\infty \quad \text{if } x \notin D(A).$$

Definition 2.4. A pair of continuous and $(\mathcal{F}_t)$-adapted processes (X, K) is called a solution of (3) if

(i) $X_0 = x_0$, $X_t \in \overline{D(A)}$ a.s.;

(ii) K is of locally finite variation and $K_0 = 0$ a.s.;

(iii) $\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t - \mathrm{d}K_t$, $0 \leqslant t < \infty$, a.s.;

(iv) For any continuous and $(\mathcal{F}_t)$-adapted functions (α, β) with

$$(\alpha_t, \beta_t) \in \mathrm{Gr}(A), \quad \forall t \in [0, +\infty),$$

the measure

$$\langle X_t - \alpha_t, \mathrm{d}K_t - \beta_t \,\mathrm{d}t \rangle \geqslant 0 \quad \text{a.s.}$$

Proposition 2.5. *Let A be a multivalued maximal monotone operator, (X, K) and (X', K') be continuous functions with $X, X' \in \overline{D(A)}$, K, K' being of finite variation. Let (α, β) be continuous functions satisfying*

$$(\alpha_t, \beta_t) \in \mathrm{Gr}(A), \quad \forall t \geqslant 0.$$

If

$$\langle X_t - \alpha_t, \mathrm{d}K_t - \beta_t \,\mathrm{d}t \rangle \geqslant 0,$$
$$\left\langle X'_t - \alpha_t, \mathrm{d}K'_t - \beta_t \,\mathrm{d}t \right\rangle \geqslant 0,$$

then

$$\left\langle X_t - X'_t, \mathrm{d}K_t - \mathrm{d}K'_t \right\rangle \geqslant 0.$$

The following generalization of the Gronwall–Bellman type inequality (cf. [7]) is useful in our proof.

Lemma 2.6 *(Bihari's inequality). Let $\rho_\eta : \mathbb{R}^+ \to \mathbb{R}^+$ be a concave function given by*

$$\rho_\eta(x) := \begin{cases} x \log x^{-1}, & x \leqslant \eta, \\ \eta \log \eta^{-1} + (\log \eta^{-1} - 1)(x - \eta), & x > \eta, \end{cases}$$

where $0 < \eta < e^{-1}$. If $g(s)$, $q(s)$ are two strictly positive functions on $\mathbb{R}^+$ such that

$$g(t) \leqslant g(0) + \int_0^t q(s) \rho_\eta\big(g(s)\big) \,\mathrm{d}s, \quad t \geqslant 0.$$

Then

$$g(t) \leqslant \big(g(0)\big)^{\exp\{-\int_0^t q(s)\,\mathrm{d}s\}}. \tag{4}$$

Let $\rho_{1,\eta}$, $\rho_{2,\eta}$ be two concave functions defined by

$$\rho_{j,\eta}(x) := \begin{cases} x[\log(\frac{1}{x})]^{1/j}, & 0 < x \leqslant \eta, \\ ([\log(\frac{1}{\eta})]^{1/j} - \frac{1}{j}[\log(\frac{1}{\eta})]^{(1/j)-1})x + \frac{1}{j}[\log(\frac{1}{\eta})]^{(1/j)-1}\eta, & x > \eta, \end{cases}$$

where $j = 1, 2$ and $0 < \eta < 1/e$.

The following lemma has been proved in [7].

Lemma 2.7.

1^o: *For $j = 1, 2$, $\rho^j_{j,\eta}$ is decreasing in η, i.e. $\rho^j_{j,\eta_1} \leqslant \rho^j_{j,\eta_2}$ if $e^{-1} > \eta_1 > \eta_2$.*

2^o: *For any $k, \varepsilon > 0$, $n > 1$ and $e^{-1} > \eta > 0$, there is a sufficiently small $\delta > 0$ such that*

$$k(x \wedge n) \log(x \wedge n) \leqslant -k\rho_{1,\eta}(x) + \varepsilon \rho_{1,\delta}(x).$$

We make the following assumptions in the present paper:

(**H**1) (Monotonicity) There exists $\lambda_0 \in \mathbb{R}$ such that for all $x, y \in \mathbb{R}^d$

$$2\langle x-y, b(x)-b(y)\rangle + \|\sigma(x)-\sigma(y)\|_{\mathrm{HS}}^2 \leqslant \lambda_0 |x-y|^2 \big(1 \vee \log|x-y|^{-1}\big).$$

(**H**2) (Growth of σ) There exists $\lambda_1 > 0$ such that for all $x \in \mathbb{R}^d$

$$\|\sigma(x)\|_{\mathrm{HS}} \leqslant \lambda_1 \big(1+|x|\big).$$

(**H**3) (Ellipticity of σ)

$$\sigma\sigma^*(x) > 0, \quad \forall x \in \mathbb{R}^d.$$

(**H**4) (One side growth of b) There exist a $p \geqslant 2$ and constants $\lambda_3 > 0$, $\lambda_4 \geqslant 0$ such that for all $x \in \mathbb{R}^d$

$$2\langle x, b(x)\rangle + \|\sigma(x)\|_{\mathrm{HS}}^2 \leqslant -\lambda_3 |x|^p + \lambda_4.$$

The theorem below presents existence and uniqueness of solution to the MSDE.

Theorem 2.8. *Assume* (**H**1) *and* (**H**2) *hold. Then* (3) *has a unique strong solution.*

Proof. Suppose that (X_t, K_t) and $(\tilde{X}_t, \tilde{K}_t)$ are two solutions. Set $Z_t := X_t - \tilde{X}_t$, then by Itô's formula,

$$|X_t - \tilde{X}_t|^2 = 2\int_0^t \langle Z_s, \big(\sigma(X_s) - \sigma(\tilde{X}_s)\big)\,\mathrm{d}W_s\rangle - 2\int_0^t \langle Z_s, \mathrm{d}K_s - \mathrm{d}\tilde{K}_s\rangle$$
$$+ \int_0^t \big[2\langle Z_s, b(X_s) - b(\tilde{X}_s)\rangle + \|\sigma(X_s) - \sigma(\tilde{X}_s)\|_{\mathrm{HS}}^2\big]\,\mathrm{d}s.$$

Note that there exists an $\eta > 0$ such that

$$r^2\big(1 \vee \log r^{-1}\big) \leqslant \rho_\eta\big(r^2\big), \quad \forall r > 0,$$

by (**H**1) and Proposition 2.5,

$$\mathbf{E}|X_t - \tilde{X}_t|^2 \leqslant \lambda_0 \mathbf{E}\int_0^t |Z_s|^2\big(1 \vee \log|Z_s|^{-1}\big)\,\mathrm{d}s$$
$$\leqslant \lambda_0 \mathbf{E}\int_0^t \rho_\eta\big(|Z_s|^2\big)\,\mathrm{d}s \leqslant \lambda_0 \int_0^t \rho_\eta\big(\mathbf{E}|Z_s|^2\big)\,\mathrm{d}s,$$

where the last inequality is due to Jensen's inequality.

By the Bihari inequality (4), we get

$$\mathbf{E}|Z_t|^2 = 0,$$

which yields the pathwise uniqueness.

The existence of a weak solution has been proved in [3] and therefore by Yamada–Watanabe's theorem (cf. [6]), (3) has a unique strong solution. □

Remark 2.9. When b is monotone, the existence and uniqueness of solutions to (3) has been proved in [8].

3. Main result and the proof

At the beginning of this section, we present our main result of the paper.

Theorem 3.1. *Assume* (**H**1)–(**H**3)*. Then the transition probability P_t of the solution to* (3) *is irreducible and strong Feller. If in addition,* (**H**4) *holds, then there exists a unique invariant probability measure μ of P_t having full support in $\overline{D(A)}$ such that*:

(i) *If $p \geqslant 2$ in* (**H**4)*, then for all $t > 0$ and $x_0 \in \overline{D(A)}$, μ is equivalent to $P_t(x_0, \cdot)$, and*

$$\lim_{t\to\infty} \big\| P_t(x_0, \cdot) - \mu \big\|_{\mathrm{Var}} = 0,$$

where $\|\cdot\|_{\mathrm{Var}}$ denotes the total variation of a signed measure.

(ii) *If $p > 2$ in* (**H**4)*, then for some $\alpha, C > 0$ independent of x_0 and t,*

$$\big\| P_t(x_0, \cdot) - \mu \big\|_{\mathrm{Var}} \leqslant C \cdot e^{-\alpha t}.$$

Moreover, for any $q > 1$ and each $\varphi \in L^q(\overline{D(A)}, \mu)$

$$\big\| P_t\varphi - \mu(\varphi) \big\|_q \leqslant C_q \cdot e^{-\alpha t/q} \|\varphi\|_q, \quad \forall t > 0,$$

where α is the same as above and $\mu(\varphi) := \int_{\overline{D(A)}} \varphi(x)\, \mu(\mathrm{d}x)$. In particular, let L_q be the generator of P_t in $L^q(\overline{D(A)}, \mu)$, then L_q has a spectral gap (*greater than α/q*) *in $L^q(\overline{D(A)}, \mu)$.*

From Theorem 2.1 we know that the key point of the proof lies in irreducibility and strong Feller property of the transition probability of the solution.

3.1. Irreducibility

We shall need the following lemma whose proof is adapted from [3]:

Lemma 3.2. *Suppose $y_0 \in \mathrm{Int}(D(A))$, $m > 0$, and Y_t is the solution to the MSDE below*:

$$\mathrm{d}Y_t + A(Y_t)\,\mathrm{d}t \ni -m(Y_t - y_0)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t, \qquad Y_0 = x_0,$$

where σ is the diffusion coefficient of (3)*. Then under* (**H**1) *and* (**H**2) *we have*

$$\mathbf{E}|Y_t - y_0|^2 \leqslant e^{-C(m)t}|x_0 - y_0|^2 + \frac{C_0}{C(m)},$$

where $C(m) = 2(m - 2\lambda_1^2 - 1/2)$ and $C_0 = 2\lambda_1^2(1 + 2|y_0|^2) + |A^\circ(y_0)|^2$.

Proof. Consider the solution Y_t^n to the following equation:

$$\mathrm{d}Y_t^n + A_n\big(Y_t^n\big)\,\mathrm{d}t = -m\big(Y_t^n - y_0\big)\,\mathrm{d}t + \sigma\big(Y_t^n\big)\,\mathrm{d}W_t, \qquad Y_0^n = x_0,$$

where A_n is the Yosida approximation of A. Since the law of Y_t^n converges to that of Y_t (cf. [1,2]), it is enough to prove the inequality for Y_t^n. By (**H**2) and Proposition 2.3, we have

$$
\begin{aligned}
&-2m|x-y_0|^2+\|\sigma(x)\|_{\mathrm{HS}}^2-2\langle A_n(x),x-y_0\rangle\\
&\quad\leqslant -2m|x-y_0|^2+\lambda_1^2\big(1+|x|\big)^2-2\langle A_n(x)-A_n(y_0),x-y_0\rangle-2\langle A_n(y_0),x-y_0\rangle\\
&\quad\leqslant -2m|x-y_0|^2+\lambda_1^2\big(1+|x|\big)^2+|x-y_0|^2+\big|A^\circ(y_0)\big|^2\\
&\quad\leqslant -2m|x-y_0|^2+2\lambda_1^2\big(1+2|x-y_0|^2+2|y_0|^2\big)+|x-y_0|^2+\big|A^\circ(y_0)\big|^2\\
&\quad= -C(m)|x-y_0|^2+C_0.
\end{aligned}
$$

Thus, by Itô's formula we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}\big|Y_t^n-y_0\big|^2&=-2\mathbf{E}\big(\big\langle Y_t^n-y_0,A_n\big(Y_t^n\big)\big\rangle\big)+\mathbf{E}\big[\mathrm{Tr}\big(\sigma\sigma^*\big(Y_t^n\big)\big)\big]-2m\mathbf{E}\big|Y_t^n-y_0\big|^2\\
&\leqslant -C(m)\mathbf{E}\big|Y_t^n-y_0\big|^2+C_0.
\end{aligned}
$$

Therefore it is easy to deduce that

$$
\mathbf{E}\big|Y_t^n-y_0\big|^2\leqslant e^{-C(m)t}|x_0-y_0|^2+\frac{C_0}{C(m)}. \qquad \square
$$

Proposition 3.3. *Under* (**H**1)–(**H**3)*, the transition probability P_t is irreducible.*

Proof. To prove the irreducibility, it suffices to prove that for any $x_0\in\overline{D(A)}$, $T>0$, $y_0\in\mathrm{Int}(D(A))$ and $a>0$,

$$
P_T\big(x_0,B(y_0,a)\big)=\mathbf{P}\big(X_T(x_0)\in B(y_0,a)\big)=\mathbf{P}\big(\big|X_T(x_0)-y_0\big|\leqslant a\big)>0,
$$

or equivalently:

$$
\mathbf{P}\big(\big|X_T(x_0)-y_0\big|>a\big)<1.
$$

From now on, a, T and y_0 are fixed. By Lemma 3.2 and Chebyshev's inequality, we can choose an m large enough such that, denoting by $(Y_t,\tilde{K}_t)$ the unique solution to

$$
\mathrm{d}Y_t+A(Y_t)\,\mathrm{d}t\ni -m(Y_t-y_0)\,\mathrm{d}t+\sigma(Y_t)\,\mathrm{d}W_t,\qquad Y_0=x_0\in\overline{D(A)}, \tag{5}
$$

$$
\mathbf{P}\big(\big|Y_T(x_0)-y_0\big|>a\big)\leqslant\bigg(e^{-C(m)T}|x_0-y_0|^2+\frac{C_0}{C(m)}\bigg)\Big/a^2<1. \tag{6}
$$

Set

$$
\tau_N:=\inf\big\{t\colon |Y_t|\geqslant N\big\}.
$$

Note that by [3] for some constant C depending on x_0, y_0, λ_1, m and T,

$$
\mathbf{E}\Big[\sup_{t\in[0,T]}\big|Y_t(x_0)\big|\Big]\leqslant C
$$

holds.

Thus we may fix an N so that

$$
\mathbf{P}(\tau_N\leqslant T)+\mathbf{P}\big(\big|Y_T(x_0)-y_0\big|>a\big)<1. \tag{7}
$$

Define

$$
U_t:=\sigma(Y_t)^*\big[\sigma(Y_t)\sigma(Y_t)^*\big]^{-1}\big(-m(Y_t-y_0)-b(Y_t)\big)
$$

and

$$Z_T = \exp\left(\int_0^{T\wedge\tau_N} U_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^{T\wedge\tau_N} |U_s|^2\,\mathrm{d}s\right).$$

Since $|U_{t\wedge\tau_N}|^2$ is bounded, then $\mathbf{E}[Z_T] = 1$ by Novikov's criteria.

By Girsanov's theorem, $W_t^* := W_t + V_t$ is a Q-Brownian motion, where

$$V_t := \int_0^{t\wedge\tau_N} U_s\,\mathrm{d}s, \qquad Q := Z_T\mathbf{P}.$$

By (7) we have

$$Q\big(\{\tau_N \leqslant T\} \cup \big\{\big|Y_T(x_0) - y_0\big| > a\big\}\big) < 1. \tag{8}$$

Note that the solution of (5) $(Y_t, \tilde{K}_t)$ also solves the MSDE

$$Y_t + \int_0^t A(Y_s)\,\mathrm{d}s \ni \int_0^t \sigma(Y_s)\,\mathrm{d}W_s^* + \int_0^{t\wedge\tau_N} b(Y_s)\,\mathrm{d}s - \int_{t\wedge\tau_N}^t m(Y_s - y_0)\,\mathrm{d}s.$$

Set

$$\theta_N := \inf\big\{t\colon\ |X_t| \geqslant N\big\}.$$

Then the uniqueness in distribution for (5) yields that the law of $\{(X_t\mathbb{1}_{\{\theta_N\geqslant T\}})_{t\in[0,T]}, \theta_N\}$ under $\mathbf{P}$ is the same as that of $\{(Y_t\mathbb{1}_{\{\tau_N\geqslant T\}})_{t\in[0,T]}, \tau_N\}$ under Q. Hence

$$\begin{aligned}
\mathbf{P}\big(\big|X_T(x_0) - y_0\big| > a\big) &\leqslant \mathbf{P}\big(\{\theta_N \leqslant T\} \cup \big\{\theta_N \geqslant T,\ \big|X_T(x_0) - y_0\big| > a\big\}\big) \\
&= Q\big(\{\tau_N \leqslant T\} \cup \big\{\tau_N \geqslant T,\ \big|Y_T(x_0) - y_0\big| > a\big\}\big) \\
&\leqslant Q\big(\{\tau_N \leqslant T\} \cup \big\{\big|Y_T(x_0) - y_0\big| > a\big\}\big) \\
&< 1
\end{aligned}$$

as desired. □

3.2. Strong Feller property

To prove that P_t is strong Feller, we first need a lemma.

Lemma 3.4. *Denote by $(X_t(x), K_t(x))$ the solution of* (3) *with initial value x. Then for any $p > d$, there exists $t_p > 0$ such that for all $r > 0$*

$$\mathbf{E}\Big[\sup_{x\in D_r,\, s\leqslant t_p} \big|X_s(x)\big|^p\Big] < \infty,$$

where $D_r := \overline{D(A)} \cap \{|x| \leqslant r\}$.

Proof. For $x, y \in D_r$, set

$$Z_t := X_t(x) - X_t(y)$$

and

$$\Lambda_t := \sigma\big(X_t(x)\big) - \sigma\big(X_t(y)\big).$$

By Itô's formula and Proposition 2.5 we have

$$\begin{aligned}|Z_t|^p &= |x-y|^p + p\int_0^t |Z_s|^{p-2}\big\langle Z_s, \Lambda_s\,\mathrm{d}W(s)\big\rangle + \frac{1}{2}\int_0^t \sum_{i,j=1}^d f_{ij}(Z_s)\big(\Lambda_s^*\Lambda_s\big)_{ij}\,\mathrm{d}s \\ &\quad + p\int_0^t |Z_s|^{p-2}\big\langle b\big(X_s(x)\big) - b\big(X_s(y)\big), Z_s(x)\big\rangle\,\mathrm{d}s \\ &\quad - p\int_0^t |Z_s|^{p-2}\big\langle Z_s, \mathrm{d}K_s(x) - \mathrm{d}K_s(y)\big\rangle \\ &\leqslant |x-y|^p + p\int_0^t |Z_s|^{p-2}\big\langle Z_s, \Lambda_s\,\mathrm{d}W(s)\big\rangle + \frac{1}{2}\int_0^t \sum_{i,j=1}^d f_{ij}(Z_s)\big(\Lambda_s^*\Lambda_s\big)_{ij}\,\mathrm{d}s \\ &\quad + p\int_0^t |Z_s|^{p-2}\big\langle b\big(X_s(x)\big) - b\big(X_s(y)\big), Z_s(x)\big\rangle\,\mathrm{d}s,\end{aligned}$$

where

$$f_{ij}(x) := p(p-2)|x|^{p-4}x_i x_j + p|x|^{p-2}\delta_{ij}.$$

Taking expectations and using Burkhölder's inequality and Lemma 2.6 give

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}|Z_s|^p\Big] \leqslant |x-y|^p + C_p\int_0^t \rho_{1,\eta^p}\Big(\mathbf{E}\Big[\sup_{0\leqslant u\leqslant s}|Z_u|^p\Big]\Big)\,\mathrm{d}s,$$

where C_p is a constant depending on p. Hence by the Bihari inequality (4) we have

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}\big|X_s(x) - X_s(y)\big|^p\Big] = \mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}|Z_s|^p\Big] \leqslant |x-y|^{p\cdot\exp(-C_p t)}.$$

For $p > d$, choose $t_p > 0$ small enough such that $p\cdot\exp(-C_p t_p) > d$. Deduce by Kolmogorov's criterion that

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t_p}\sup_{x\in D_r}\big|X_s(x)\big|^p\Big] \leqslant C\Big(1 + \mathbf{E}\Big[\sup_{0\leqslant s\leqslant t_p}\big|X_s(x_0)\big|^p\Big]\Big),$$

where $x_0 \in D(A)$. By [1], we also know

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t_p}\big|X_s(x_0)\big|^p\Big] < \infty.$$

The proof is thus complete. □

Proposition 3.5. *Under* (**H**1)–(**H**3)*, the semigroup* P_t *is strong Feller.*

Proof. We divide the proof into two steps.

Step 1. Assume that the diffusion coefficient is uniformly elliptic with

$$\big\|[\sigma^*\sigma]^{-1}\big\|_{\mathrm{HS}} \leqslant \lambda_2$$

for some $\lambda_2 > 0$ and in this case we follow essentially the argument in [9].

Consider the following drift transformed MSDE:

$$\begin{cases} \mathrm{d}Y_t + A(Y_t)\,\mathrm{d}t \ni b(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t + |x_0 - y_0|^\alpha \dfrac{X_t - Y_t}{|X_t - Y_t|} \cdot 1_{\{X_t \neq Y_t\}} \cdot 1_{\{t<\tau\}}\,\mathrm{d}t, \\ Y_0 = y_0 \in \overline{D(A)}, \end{cases} \tag{9}$$

where $\alpha \in (0,1)$, X_t is the solution to (3) and τ is the coupling time given by

$$\tau := \inf\big\{t > 0\colon |X_t - Y_t| = 0\big\}.$$

We make a change to the coupling function:

$$\mathrm{d}Y_t^\delta + A\big(Y_t^\delta\big)\,\mathrm{d}t \ni b\big(Y_t^\delta\big)\,\mathrm{d}t + \sigma\big(Y_t^\delta\big)\,\mathrm{d}W_t + c_\delta\big(X_t - Y_t^\delta\big)\,\mathrm{d}t, \qquad Y_0^\delta = y_0, \tag{10}$$

where

$$c_\delta(z) := |x_0 - y_0|^\alpha \cdot f_\delta\big(|z|\big) \cdot \frac{z}{|z|},$$

and

$$f_\delta(r) := \begin{cases} 1, & r > \delta, \\ 0, & r \in [0, \delta/2] \end{cases}$$

is a smooth function from $\mathbb{R}_+$ to $[0,1]$.

It is easy to see that for any $z, z' \in \mathbb{R}^d$, there exists a constant C_δ such that

$$\big|c_\delta(z) - c_\delta(z')\big| \leqslant C_\delta \cdot |z - z'|.$$

Thus there is a unique solution, denoted here by $(Y_t^\delta, \tilde{K}_t^\delta)$, to (10).

Define

$$\tau_\delta := \inf\big\{t > 0\colon \big|X_t - Y_t^\delta\big| \leqslant \delta\big\}.$$

For $\delta' < \delta$, the uniqueness yields that $\tau_{\delta'} \geqslant \tau_\delta$ and

$$Y_t^\delta = Y_t^{\delta'}, \qquad \tilde{K}_t^\delta = \tilde{K}_t^{\delta'} \quad \text{on } \{t < \tau_\delta\}.$$

So $\tau = \lim_{\delta \downarrow 0} \tau_\delta$ is just the coupling time and Y_t is well defined on $[0, \tau]$. For all $t \geqslant \tau$, define

$$Y_t := X_t, \qquad \tilde{K}_t := K_t.$$

Then $(Y_t, \tilde{K}_t)$ solves (9).

Now fix a $T > 0$ and define

$$U_T := \exp\left[\int_0^{T\wedge\tau} \big\langle \mathrm{d}W_s, H(X_s, Y_s)\big\rangle - \frac{1}{2}\int_0^{T\wedge\tau} \big|H(X_s, Y_s)\big|^2\,\mathrm{d}s\right]$$

and

$$\tilde{W}_t := W_t + \int_0^{t\wedge\tau} H(X_s, Y_s)\,\mathrm{d}s,$$

where

$$H(x,y) := |x_0 - y_0|^\alpha \cdot \sigma^*(y)\big[\sigma\sigma^*(y)\big]^{-1}\frac{x-y}{|x-y|}.$$

Since $\|[\sigma\sigma^*(y)]^{-1}\|_{\mathrm{HS}} \leqslant \lambda_2$, we have

$$\big|H(x,y)\big|^2 \leqslant \lambda_2 \cdot |x_0 - y_0|^{2\alpha}.$$

Thus,

$$\mathbf{E}U_T = 1 \quad \text{and} \quad \mathbf{E}U_T^2 \leqslant \exp\big[\lambda_2 T \cdot |x_0 - y_0|^{2\alpha}\big].$$

By the elementary inequality $e^r - 1 \leqslant re^r$ for $r \geqslant 0$, we have for any $|x_0 - y_0| \leqslant \eta$,

$$\begin{aligned}\big(\mathbf{E}|1-U_T|\big)^2 &\leqslant \mathbf{E}|1-U_T|^2 = \mathbf{E}U_T^2 - 1\\ &\leqslant \exp\big[\lambda_2 T\cdot|x_0-y_0|^{2\alpha}\big] - 1\\ &\leqslant C_{T,\lambda_2,\eta}\cdot|x_0-y_0|^{2\alpha}\end{aligned} \tag{11}$$

and

$$\begin{aligned}\big(\mathbf{E}\big[(1+U_T)1_{\{\tau\geqslant T\}}\big]\big)^2 &\leqslant \big(3+\mathbf{E}U_T^2\big)\cdot\mathbf{P}(\tau\geqslant T)\\ &\leqslant C_{T,\lambda_2,\eta}\cdot\mathbf{P}\big((2T)\wedge\tau\geqslant T\big)\\ &\leqslant C_{T,\lambda_2,\eta}\cdot\mathbf{E}\big((2T)\wedge\tau\big)/T.\end{aligned} \tag{12}$$

Now apply Itô's formula to $\sqrt{|Z_{t\wedge\tau}|^2+\varepsilon}$ where $Z_s := X_s - Y_s$, and then let $\varepsilon \downarrow 0$, we obtain by (**H**1) and Proposition 2.5,

$$\begin{aligned}&|Z_{t\wedge\tau}| - |x_0-y_0| - \int_0^{t\wedge\tau}\big\langle\bar Z_s, \big(\sigma(X_s)-\sigma(Y_s)\big)\,\mathrm{d}W_s\big\rangle\\ &\quad= \int_0^{t\wedge\tau}\big(2|Z_s|\big)^{-1}\cdot\big(2\big\langle Z_s, b(X_s)-b(Y_s)\big\rangle + \big\|\sigma(X_s)-\sigma(Y_s)\big\|_{\mathrm{HS}}^2\big)\,\mathrm{d}s - \int_0^{t\wedge\tau}\big\langle\bar Z_s, a(Z_s)\big\rangle\,\mathrm{d}s\\ &\qquad - \int_0^{t\wedge\tau}\langle\bar Z_s, \mathrm{d}K_s - \mathrm{d}\tilde K_s\rangle - \int_0^{t\wedge\tau}\big(2|Z_s|\big)^{-1}\cdot\big|\big(\sigma(X_s)-\sigma(Y_s)\big)^*\bar Z_s\big|^2\,\mathrm{d}s\\ &\quad\leqslant \frac{\lambda_0}{2}\int_0^{t\wedge\tau}|Z_s|\big(1\vee\log|Z_s|^{-1}\big)\,\mathrm{d}s - |x_0-y_0|^\alpha(t\wedge\tau),\end{aligned}$$

where

$$\bar z = z/|z|$$

and

$$a(Z_t) = |x_0-y_0|^\alpha\,\bar Z_t.$$

Note that there exists an $\eta > 0$ such that

$$r\big(1\vee\log r^{-1}\big) \leqslant \rho_\eta(r), \quad \forall r > 0.$$

Taking expectations yields that

$$\begin{aligned}\mathbf{E}|X_{t\wedge\tau}-Y_{t\wedge\tau}| &\leqslant |x_0-y_0|-|x_0-y_0|^\alpha\cdot\mathbf{E}(t\wedge\tau)+\frac{\lambda_0}{2}\mathbf{E}\int_0^{t\wedge\tau}\rho_\eta\big(|X_s-Y_s|\big)\,\mathrm{d}s\\ &\leqslant |x_0-y_0|-|x_0-y_0|^\alpha\cdot\mathbf{E}(t\wedge\tau)+\frac{\lambda_0}{2}\int_0^{t}\rho_\eta\big(\mathbf{E}|X_{s\wedge\tau}-Y_{s\wedge\tau}|\big)\,\mathrm{d}s.\end{aligned}$$

By the Bihari inequality (4), we get that for any $t>0$ and $|x_0-y_0|<\eta$

$$\mathbf{E}|X_{t\wedge\tau}-Y_{t\wedge\tau}|\leqslant |x_0-y_0|^{\exp\{-\lambda_0 t/2\}}$$

and thus

$$\mathbf{E}(t\wedge\tau)\leqslant |x_0-y_0|^{1-\alpha}+\frac{\lambda_0 t}{2}\rho_\eta\big(|x_0-y_0|^{\exp\{-\lambda_0 t/2\}}\big)\cdot|x_0-y_0|^{-\alpha}. \tag{13}$$

Taking $\alpha=\exp\{-\lambda_0 T\}/2$, there exists a $0<\eta'<\eta$ such that for any $|x_0-y_0|<\eta'$

$$\mathbf{E}\big((2T)\wedge\tau\big)\leqslant C_{T,\lambda_0,\eta'}\cdot|x_0-y_0|^{\exp\{-\lambda_0 T\}/2}. \tag{14}$$

But by Girsanov's theorem, $(\tilde{W}_t)_{t\in[0,T]}$ is still an n-dimensional Brownian motion under the new probability measure $U_T\cdot\mathbf{P}$. Note that $(Y_t,\tilde{K}_t)$ also solves

$$\mathrm{d}Y_t+A(Y_t)\,\mathrm{d}t\ni b(Y_t)\,\mathrm{d}t+\sigma(Y_t)\,\mathrm{d}\tilde{W}_t,\qquad Y_0=y_0.$$

So, the law of $X_T(y_0)$ under $\mathbf{P}$ is the same as that of $Y_T(y_0)$ under $U_T\cdot\mathbf{P}$.

Thus by (11), (12) and (14), for any $f\in B_b(\mathbb{R}^d)$,

$$\begin{aligned}\big|P_Tf(x_0)-P_Tf(y_0)\big| &= \big|\mathbf{E}\big(f\big(X_T(x_0)\big)-U_T\cdot f\big(Y_T(y_0)\big)\big)\big|\\ &\leqslant \mathbf{E}\big|(1-U_T)\cdot f\big(X_T(x_0)\big)\cdot 1_{\{\tau\leqslant T\}}\big|\\ &\quad+\mathbf{E}\big|\big(f\big(X_T(x_0)\big)-U_T\cdot f\big(Y_T(y_0)\big)\big)\cdot 1_{\{\tau>T\}}\big|\\ &\leqslant \|f\|_0\cdot\mathbf{E}|1-U_T|+\|f\|_0\cdot\mathbf{E}\big[(1+U_T)1_{\{\tau>T\}}\big]\\ &\leqslant C_{T,\lambda_0,\lambda_2,\eta}\cdot\|f\|_0\cdot|x_0-y_0|^{\exp\{-\lambda_0 T\}/4}.\end{aligned}$$

Hence we have proved the strong Feller property of (P_t) when the diffusion coefficient is uniformly elliptic.

Step 2. Now we turn to the case under the assumption (**H**3), that is, the diffusion coefficient is only elliptic. By the Markov property of the solution, we only need to prove that for every $f\in B_b(\mathbb{R}^d)$, $x\mapsto P_tf(x)$ is continuous on D_r for all $t\leqslant t_p$, $p>d$ where p and t_p are specified in Lemma 3.4. Set

$$c_0:=\|f\|_\infty$$

and

$$\tau:=\inf\Big\{t>0\colon \sup_{x\in D_r}\big|X_t(x)\big|>N\Big\}.$$

Let $\varepsilon>0$ be given. By Lemma 3.4 and Chebyshev inequality, there exists $N>r$ such that

$$\mathbf{P}(\tau\leqslant t_p)=\mathbf{P}\Big(\sup_{x\in D_r,\,t\leqslant t_p}\big|X_t(x)\big|>N\Big)\leqslant\mathbf{E}\Big[\sup_{x\in D_r,\,t\leqslant t_p}\big|X_t(x)\big|^p\Big]\Big/N^p<\varepsilon. \tag{15}$$

Define

$$\tilde{\sigma}(x) := \sigma(x), \quad \forall |x| \leqslant N.$$

Extend $\tilde{\sigma}$ to the whole $\mathbb{R}^d$ such that it satisfies the condition (**H**1) to (**H**3). Denote by $\tilde{X}_t(x)$ the solution to (3) with σ replaced by $\tilde{\sigma}$. By Step 1, there exists a $\delta > 0$ such that if $|x-y| < \delta$ and $x, y \in D_r$,

$$\left|\mathbf{E}\left[f\left(\tilde{X}_t(x)\right)\right] - \mathbf{E}\left[f\left(\tilde{X}_t(y)\right)\right]\right| < \varepsilon. \tag{16}$$

Hence for $t \leqslant t_p$

$$\begin{aligned}
&\left|\mathbf{E}\left[f\left(X_t(x)\right)\right] - \mathbf{E}\left[f\left(X_t(y)\right)\right]\right| \\
&\quad \leqslant \left|\mathbf{E}\left[\left(f\left(X_t(x)\right) - f\left(X_t(y)\right)\right)1_{(\tau > t_p)}\right]\right| + \left|\mathbf{E}\left[\left(f\left(X_t(x)\right) - f\left(X_t(y)\right)\right)1_{(\tau \leqslant t_p)}\right]\right| \\
&\quad \leqslant \left|\mathbf{E}\left[\left(f\left(\tilde{X}_t(x)\right) - f\left(\tilde{X}_t(y)\right)\right)1_{(\tau > t_p)}\right]\right| + 2c_0\varepsilon \\
&\quad \leqslant \left|\mathbf{E}\left[f\left(\tilde{X}_t(x)\right) - f\left(\tilde{X}_t(y)\right)\right]\right| + \left|\mathbf{E}\left[\left(f\left(\tilde{X}_t(x)\right) - f\left(\tilde{X}_t(y)\right)\right)1_{(\tau \leqslant t_p)}\right]\right| + 2c_0\varepsilon \\
&\quad \leqslant (1+4c_0)\varepsilon,
\end{aligned}$$

and the proof is completed. □

Now we are in a position to complete the proof of Theorem 3.1.

Proof of Theorem 3.1. (i) By Itô's formula and (**H**4), we get

$$\begin{aligned}
\mathbf{E}|X_t|^2 &= |x_0|^2 + 2\int_0^t \mathbf{E}\langle X_s, b(X_s)\rangle \,\mathrm{d}s - 2\int_0^t \mathbf{E}\langle X_s, \mathrm{d}K_s\rangle + \int_0^t \mathbf{E}\left\|\sigma(X_s)\right\|_{\mathrm{HS}}^2 \mathrm{d}s \\
&\leqslant |x_0|^2 + \int_0^t \mathbf{E}\left(-\lambda_3|X_s|^p + \lambda_4\right) \mathrm{d}s.
\end{aligned}$$

Taking derivatives with respect to t and using Hölder's inequality give

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{E}|X_t|^2}{\mathrm{d}t} &\leqslant -\lambda_3 \mathbf{E}|X_t|^p + \lambda_4 \\
&\leqslant -\lambda_3\left(\mathbf{E}|X_t|^2\right)^{p/2} + \lambda_4.
\end{aligned}$$

Since $\lambda_3 > 0$ we have for all $t > 0$,

$$\frac{1}{t}\int_0^t \mathbf{E}|X_s|^2 \,\mathrm{d}s \leqslant \lambda_4/\lambda_3.$$

Therefore by Krylov–Bogoliubov's method (cf. [4]), there exists an invariant probability measure μ. As we have just proved, P_t is strong Feller and irreducible, then by Theorem 2.1, μ is equivalent to each $P_t(x,\cdot)$ with $x \in \overline{D(A)}$, $t > 0$ and consequently (i) holds.

(ii) If $p > 2$, consider the following ODE:

$$f'(x) = -\lambda_3 f(x)^{p/2} + \lambda_4, \qquad f(0) = |x_0|^2.$$

By the comparison theorem (cf. [4]), there exists some $C > 0$ such that

$$\mathbf{E}|X_t|^2 \leqslant f(t) \leqslant C\big(1 + t^{2/(2-p)}\big).$$

We also have

$$\inf_{x_0 \in B(0,r)} P_t\big(x_0, B(0,a)\big) > 0, \quad \forall r, a > 0,\ t > 0$$

because of the strong Feller property and irreducibility. Therefore (ii) holds due to Theorems 2.5(b) and 2.7 in [5]. □

References

[1] E. Cépa, Équations différentielles stochastiques multivoques, in: Sém. Prob. XXIX, in: Lecture Notes in Math., 1995, pp. 86–107.

[2] E. Cépa, Problème de Skorohod multivoque, Ann. Probab. 26 (2) (1998) 500–532.

[3] E. Cépa, S. Jacquot, Ergodicité d'inégalités variationnelles stochastiques, Stochastics Stochastics Rep. 63 (1997) 41–64.

[4] S. Cerrai, Second Order PDE's in Finite and Infinite Dimension. A Probabilistic Approach, Lecture Notes in Math., vol. 1762, Springer-Verlag, Berlin, 2001, x+330 pp.

[5] B. Goldys, B. Maslowski, Exponential ergodicity for stochastic reaction–diffusion equations, in: Stochastic Partial Differential Equations and Applications—VII, in: Lect. Notes Pure Appl. Math., vol. 245, Chapman Hall/CRC, Boca Raton, FL, 2006, pp. 115–131.

[6] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland Publishing Company, Tokyo, 2001, x+330 pp.

[7] J. Ren, X. Zhang, Stochastic flows for SDEs with non-Lipschitz coefficients, Bull. Sci. Math. 127 (2003) 739–754.

[8] X. Zhang, Skorohod problem and multivalued stochastic evolution equations in Banach spaces, Bull. Sci. Math. 131 (2) (2007) 175–217.

[9] X. Zhang, Exponential ergodicity of non-Lipschitz stochastic differential equations, Proc. Amer. Math. Soc. 137 (2009) 329–337.